

Cell-line authentication demystified

Vivien Marx

Some researchers run and hide from the task of authenticating cell lines. A few simple steps save time and money.

Scientists are not easily fooled, but they can unwittingly fool themselves. The identity of a cell line might not be what they assume it is. It is “very risky” for an investigator to begin a project and spend several years working with a cell line and yet be unsure of its identity, says Karina Eterovic, who directs the characterized cell-line core at MD Anderson Cancer Center’s institute of personalized cancer therapy.

Skipping cell-line authentication can lead to lost time, money and publications owing to cross-contamination and misidentification. Last year, a team had to retract work published in *Nature Methods* because the cells they studied were contaminated with other cells¹. It was not the first case^{2,3}, nor is it likely to be the last.

Eterovic helped colleagues clarify that the widely used KU7 line, derived from bladder cancer cells, is in fact the HeLa cell line⁴. The contamination occurred before 1984 at the source institution, which has an effect on all KU7-based experimental results and underscores the impact of misidentification.

Cell-line authentication can seem too time-consuming or daunting. Avoidance and denial behaviors are widespread, say researchers interviewed by *Nature Methods*. They understand why their colleagues skip essentials to pursue biological questions without delay, but they say it is time to put a few common misconceptions to rest.

Standards

Some sweet songs of praise for the scientific community can be heard on the subject of cell-line authentication. There has been “a lot of progress,” says Yvonne



Researchers should know the true identities of their cell lines.

Reid, a cell biologist at American Type Culture Collection (ATCC), a nonprofit biological resource center and repository that offers cell-line authentication services and is active in the area of standards in the life sciences. Reid has co-authored reports on best practices in cell-line authentication and the recommended standards⁵.

The standards have brought about more change than the numerous warnings and pleas in past decades, Reid says. The standards are used in many labs and core facilities around the world. Workshops abound on best practices in tissue culture;

many journals require authentication of cell lines, and grant-funding agencies recommend them, all of which makes many researchers optimistic that cell-line misidentification will one day be vanquished—or at least diminished.

That day is not quite nigh. The pollution of labs and scientific literature with misidentified cell lines may well be due to a continued lack of awareness about the extent of the problem, says Douglas Storts, who directs research and development for nucleic acid research technologies at Promega. He is a co-author with Reid of guidelines on cell-line authentication⁵ and was interviewed jointly with her.

Among the important resources is the Database of Cross-contaminated or Misidentified Cell Lines⁶, maintained by the International Cell Line Authentication Committee (ICLAC), a group founded in 2012 and run by volunteer scientists to establish best practices for the authentication and testing of human cell lines. According to ICLAC, approximately 400 cell lines are misidentified⁶.

The oldest cell line, HeLa, which is one of the most frequent contaminants of other

BOX 1 ADVICE FOR THE PASSAGE

- Cells must be transferred from flask to flask so they keep growing. But too much passaging can lead to genetic drift from the parental cell line.
- There is no global right or wrong number of passages. The number varies across cell and tissue types and culture conditions.
- Begin with high-quality, authenticated and characterized cells.
- Watch the cells carefully. Media and environmental conditions such as temperature and humidity need to be optimized for cells’ well-being.
- Growth-curve analysis should be routine to keep track of cell-doubling times.
- Freeze batches of cells at early passages.
- Use the lowest-passage aliquot to start fresh cultures.

Sources: ATCC; Promega



MD Anderson Cancer Center's Karina Eterovic offers in-house repository services for cell lines, including gene-edited ones.

cell lines, was derived without consent in 1951 from Henrietta Lacks, who had cervical cancer. The cells and the way they were acquired have been the subject of controversy over the years. Last year, the HeLa genome was sequenced, rekindling ethics issues because the family had not consented to the sequencing. Meetings between members of the Lacks family and leadership at the US National Institutes of Health (NIH) led to the establishment of a rule that all publications based on work with HeLa cells acknowledge the Lacks family, who also help to oversee requests to use the genome data.

It is still too early to tell how the discussions about these famous cells are influencing cell-line authentication trends. Having the NIH director address issues relating to cell-line identity is bound to be helpful, says Reid.

The published characterizations of the HeLa cell line show a large number of mutations and rearrangements and help explain why HeLa cells are able to displace other cell lines, says Storts.

Passaging

Researchers must 'passage' their cells—by changing the growth medium so the cells have fresh nutrients and moving cells to other flasks to maintain their exponential growth. The longer a cell line is kept in culture, the higher the chances it will acquire genetic changes⁷. With more passages, genetic drift is more likely, says Reid, which can lead to dropouts of

certain alleles or loss of heterozygosity (**Box 1**).

As cells adapt to environmental conditions they can change so much that they are no longer the cells they once were, says Darrin Disley, chief executive officer of gene-editing company Horizon, which sells cell lines and offers cell-line repository services for scientists.

A lab might send cells from the same parental cell line to 20 labs. But drift can lead to such large differences that it becomes hard to compare experimental results from these cells. This situation could be especially likely to arise, Disley says, if the cells were taken from a late-passaged clone that had "just been hanging on for grim life". Cell culture medium along with many other factors can also have a role.

Short tandem repeats

The main test for authenticating cell lines is short tandem repeat (STR) profiling, the analysis of microsatellite regions of DNA that have variable numbers of repeats and are located throughout the genome. The approach originated with Alec Jeffreys's discovery of variable genomic regions in 1984, and development has continued since then.

STR profiling is the test that most journals require, says Eterovic, and the assay is relatively simple. PCR amplicons are generated using primers for regions that flank these polymorphic sites. After amplification, PCR products are resolved using capillary electrophoresis, and a profile for these repetitive regions is created.

Investigators can do the analysis on their own, using kits from vendors with primers and reagents for well-known STRs. Buying a kit avoids the need to design primers

and optimize PCR reactions, Eterovic says. Scientists still need a capillary electrophoresis system, which may not be worth the investment if they are only working with a few cell lines.

Analyzing these regions is becoming easier thanks to STR profile databases. "Now individuals can blast against that database and see whether they have misidentified cell lines or not," says Reid.

Cancer cell lines have many mutations and deletions, so STR analysis requires carefully interpreting the DNA profile report, which presents alleles graphically as peaks. Expert help might be necessary for this analysis, says Gabriela Saldanha, a marketing manager at Promega, which offers cell-line authentication kits and services. She wonders whether some cell biologists avoid STR analysis because they have less experience with the molecular biology methods, she says. "Maybe they are not comfortable with having to perform a test themselves."

Scientists at core facilities might help in various ways; for example, they might be able to do the protocol more quickly, Saldanha says. But they may also leave the interpretation of results up to the scientists, which could pose a challenge for those unfamiliar with the territory (**Box 2**).

One misconception in some labs is that cell-line authentication calls for an identical STR profile match. But STR profiling looks for an 80% match or higher in profiled loci, says Donald Finley, Sigma-Aldrich's marketing manager for products relating to *Mycoplasma*. Sigma-Aldrich also sells kits for STR profiling.

"With 80% you can be reasonably confident that you're working with the cell line that you think you're working with,"

BOX 2 TROUBLESHOOTING STR ANALYSIS

- With cancer cell lines it can be difficult to distinguish low-level STR peaks from noise. Using the recommended amount of DNA can help minimize noise.
- Using positive and negative controls helps troubleshoot STR profiles.
- Artifacts may require a repeat of the STR protocol.
- 'Stutter peaks' are usually lower than true allele peaks, but using too much DNA can lead to extremely high stutter peaks, making them harder to distinguish from true allele peaks.
- Dye blobs can appear as wide peaks, which the software might assign as regular peaks. The blobs are formed when fluorescent dye tags detach from their primers. Sometimes this can happen if kits are not stored as vendors recommend.
- STR profiling software can identify atypically sized allele microvariants. Vendor kits might not cover all microvariants.

Source: ATCC

says Storts. The reason for this 80–100% range is that cell lines can be unstable owing to a loss of heterozygosity or issues relating to their ploidy.

According to guidelines published by the American National Standards



Donald Finley recommends that cell lines are tested regularly for contaminants.

Institute⁵, cell lines that match 55–80% of STR loci need further investigation. For many scientists, that broad range might seem to pose a quandary, says Finley. But it is important to interpret results in that range as inconclusive and

to perform further authentication steps before beginning an experiment with such cells.

Eterovic confirms that many types of stresses can result in loss of heterozygosity or lead to genomic rearrangements, which alter an STR profile. The stresses include extensive passaging, contamination with microorganisms, passage through mice and exposure to drugs.

MD Anderson has its own STR database and an algorithm to match STR profiles from unstable cell lines⁸. Many users can match their cells' profiles with the help of a public database. But STR results from unstable cell lines or potential mixes of cell lines may also require special algorithms and interpretation by scientists in a core facility, Eterovic says.

Mycoplasma

Most researchers are aware of potential contamination by *Mycoplasma* bacteria, but they probably underestimate its frequency and the degree to which it can influence their experiments, says Andrea Toell, product manager at Lonza Pharma&Biotech Bioscience Solutions.

An estimated 15–35% of cell cultures are contaminated by *Mycoplasma*. “The problem is that you cannot see a contamination just by looking at your cells,” Toell says.

Mycoplasma contamination can affect cell metabolism, increase sensitivity to inducers of apoptosis or inhibit cell growth. From having helped researchers troubleshoot transfection experiments, Toell has seen that *Mycoplasma* contamination decreases transfection efficiency. When scientists do not check for

potential contamination, “there is always some uncertainty if results gained from an experiment are really generated by the cells or artificially influenced by *Mycoplasma*,” Toell says. *Mycoplasma* bacteria are visible only under an electron microscope and they lack cell walls, so antibiotics may have no effect.

Lonza sells a biochemical test that looks for the presence of two enzymes found in the majority of the ~180 species of *Mycoplasma*. The test does not require PCR and is performed on cell culture supernatant in around 20 minutes. There are other tests; for example, the most sensitive one, Toell says, is a test in which cell culture supernatant is plated. *Mycoplasma* colonies will show up if they were present in the supernatant, but only after 2–3 weeks. PCR-based tests take around 4–5 hours to deliver a result, she says.

Non-PCR approaches involve the use of indicator cells, which gives results within 16–24 hours, or Hoechst staining, which takes about two hours and requires expertise to interpret the results. Sigma-Aldrich's Finley agrees that using the Hoechst stain, which his company also sells, is “a little tricky.”

Many contaminants can kill cell lines, but *Mycoplasma* is not one of them. *Mycoplasma* can affect cell performance and render



At the Coriell Institute for Medical Research, cells are kept in ampules with cryopreservative; scientists practice cell-line authentication on a large scale.



Cell lines are stored in liquid nitrogen, which brings biological activity to a halt. The Coriell Institute for Medical Research ships around 8,000 cell cultures annually.

metabolism data or gene-expression results “murky,” Finley says. Sigma-Aldrich sells kits for PCR-based *Mycoplasma* detection and for cleaning cells of *Mycoplasma*.

Overall, Finley recommends that labs set up a regular schedule to test cell lines for contaminants. *Mycoplasma* is ubiquitous, which is why aseptic technique is advisable.

Repositories

When obtaining cell lines from a repository or vendor, scientists receive a certificate of analysis to show that the cell lines have been authenticated and are free of contamination, says Finley.

A repository can be an in-house facility at a research institute or a separate entity. It can also be organized as a collection of many repositories under one roof. In the cases of grouped repositories, each distinct collection of biological specimens or cell lines belongs to a different university group, hospital or company. The Coriell Institute for Medical Research has repositories for government, academic and nonprofit organizations and company labs. For example, they hold the cell lines and DNA of the US Human Genetic Cell Repository of the National Institute of General Medical Sciences.

At Coriell, Dorit Berlin oversees cell-line repositories for several different institutions. She and her team practice cell-line authentication on a large scale.

And, for example, when they culture cells to replenish depleted inventory, they compare the results for loci in the new expansion to those of the original culture.

The team also uses these profile matches to confirm that different cell types attributed to the same donor are from that donor and that biomaterials collected from the same person but submitted at different times match one another. The tests also look for gender match using PCR-based techniques and a sex-linked gene.

To establish that a cell line is of human origin, the scientists perform a long-interspersed nuclear element (LINE) assay, which is sensitive enough to differentiate human DNA from closely related chimpanzee DNA, Berlin says.

Sigma-Aldrich partners with European Collection of Cell Cultures (ECACC) and sells cell lines that the ECACC has authenticated. ECACC is a cell and tissue repository run by Public Health England, which is part of the UK Department of Health.

“Moving cell lines across borders is difficult,” Finley says. If a US researcher wants a European cell line, Sigma-Aldrich’s distribution agreement will help with cross-border travel. For the cell lines developed at Sigma-Aldrich, such as those edited with zinc-finger nucleases, the company handles authentication, he says.

Using the services of an in-house repository can be more practical for researchers than culturing cell lines on their own in their lab.

At MD Anderson, Eterovic says her core facility offers repository services to expand, profile, store and distribute cell lines developed by in-house investigators, including gene-edited cell lines. She and her team want to reduce the burden on departments that have interesting cell lines but cannot keep and grow them. In some cases, cell lines need a home after the original developer leaves the institution.

Depositing cell lines with her team helps investigators eliminate the cost of expanding cell lines for requestors, and Eterovic says she strongly encourages them to do so, preferably at early passages. The core can then expand the cell lines, “but most importantly, it guarantees that the cell line is validated when requested by another investigator,” she says.

Gene editing

As gene-editing techniques become more widely used, creating cell lines with desired genetic changes is becoming easier. Sharing these gene-edited cell lines with other labs is good research etiquette, and generally required once the cell line has been reported in a published research article. Researchers should authenticate cell lines as soon as they receive them but they often do not, says Disley. As a researcher, he has received gene-edited cell lines that, upon authentication, turned out to be misidentified. The genetic change that the sender said had been edited into the cell line was not present. The originating lab may have passaged the cells too often, leading to genetic drift. “That’s how you end up with a mess across the whole sector,” he says.

Starting a research project with a misidentified human cell line and then layering on gene-editing techniques exacerbates the problem of misidentification, because a lab might launch directly into studying gene function with these cells. “You need a characterized cell line from the beginning,” says Disley, to analyze follow-on experiments that look at deletions, insertions, substitutions, rearrangements or to be able to credibly perform gain-of-function experiments.

Horizon was originally established to focus on a gene-editing technology

that uses adeno-associated viruses to stimulate homologous recombination, but it has now licensed other tools, such as zinc-finger nucleases or clustered regularly interspaced short palindromic repeats (CRISPR) for gene editing. The company sells cell lines that have been modified with these technologies and has just begun offering shares on the Alternative Investment Market, a London Stock Exchange market for emerging companies. Scientists who deposit cell lines with the company can land royalties if these cells find use in drug discovery through Horizon-mediated business.

Research with gene-edited cell lines resembles a do-it-yourself project in that both require adequate preparation, says Disley. One cannot build a chair without the right wood or without knowing whether it will be able to support a certain weight. On the topic of cell-line authentication, Coriell's Berlin is reminded of something her grandmother used to say: "You can't make gourmet food out of rotten ingredients."

1. Clément, V. *et al. Nat. Methods* **7**, 224–228 (2010); retracted **10**, 1035 (2013).
2. Gartler, S.M. *et al. Natl. Cancer Inst. Monogr.* **26**, 167–195 (1967).
3. Nelson-Rees, W.A. *et al. Science* **212**, 446–452 (1981).
4. Jäger, W. *et al. J. Urol.* **190**, 1404–1409 (2013).
5. American Type Culture Collection Standards Development Organization Workgroup. Authentication of human cell lines: standardization of STR profiling (American National Standards Institute, 2011).
6. Capes-Davis, A. *et al. Int. J. Cancer* **127**, 1–8 (2010).
7. Masters, J.R. & Stacey, G.N. *Nat. Protoc.* **2**, 2276–2284 (2007).
8. Eltonsy, N. *et al.* Detection algorithm for the validation of human cell lines. *Int. J. Cancer* **131**, E1024–E1030 (2012).

Vivien Marx is technology editor for *Nature and Nature Methods* (v.marx@us.nature.com).