Mini Review

proliferating until their telomeres become so short that the chromosomes undergo fusion-breakage-bridge cycles and the ensuing genomic instability results in culture crisis. Occasionally (at a rate of ~1 in $10^7$ cells), an immortalized cell will emerge from crisis and begin to divide again, yielding a continuous cell line.[1] The changes seen throughout this process have many parallels within cancer development, both for malignancy in general and when considering specific tumor types.[7,8]

Despite these advantages, numerous cautions have emerged from the literature regarding appropriate use of cell lines as model systems.[9,10] Even where cultures have been transformed through the introduction of specific genes, cell lines that have passed through replicative senescence and crisis are aneuploid, heteroploid and genotypically and phenotypically unstable, resulting in considerable heterogeneity within the culture.[10] This instability will cause changes in the characteristics of the cell line but a further consequence may result: alterations in a cell line can be accepted by the user as intrinsic to that culture when there is actually extrinsic contamination present.

## Cell Line Cross-contamination and Misidentification

Cell lines become contaminated when a foreign cell line or microorganism is introduced without the handler's knowledge. Although we do not wish to minimize the problem of microbial contamination, we will focus on cell line cross-contamination in this article. Cross-contamination may arise due to several causes, including poor technique (spread *via* aerosols or accidental contact), use of unplugged pipets, sharing media and reagents among cell lines and use of mitotically inactivated feeder layers or conditioned medium, which may carry contaminating cells if not properly eliminated, for example, by freeze-thaw and filtration.[11] In addition, a cell line can be replaced by another as a result of misidentification by confusing cultures during handling, mislabeling or poor freezer inventory control. Simple errors during labeling of culture flasks, truncation of the cell line name or typographic errors in a published manuscript, can result in significant confusion for years after the event when another researcher attempts to use the same cell line for ongoing experimental work.[12]

Cross-contamination may occur "early," in which case the original cell line has probably never existed independently, or "late," where the tested sample has been overgrown but other stocks of the original may still exist.[13] Unfortunately, cell lines generally become cross-contaminated early, while still within the originating laboratory.[14] This is not surprising: cultures can remain in crisis for a prolonged period of time before emergence of an immortalized population and this is a time when a single cell, if introduced from a separate cell line, would rapidly take over the culture.

There are now a number of studies pointing out the severity of this problem and the need to take urgent action to minimize cross-contamination and its consequences.[9,15–17] Ten years ago, the German Collection of Microorganisms and Cell Cultures (DSMZ) published data from its identification testing of cancer cell lines submitted by various laboratories for deposit at the cell bank.[14] They found that 18% of 252 submitted cell lines were cross-contaminated with more than half of cases arising within only 6 laboratories. Subsequent work by the DSMZ, extending the number of cell lines tested (Fig. 1), shows that of 598 leukemia–lymphoma cell lines (the group provided with the most complete genetic data), 187 (31%) were contaminated with Mycoplasma and/or a second cell line with 38 (6%) of cell lines contaminated with both. These data suggest that poor practice within some laboratories results in contamination of multiple cell lines with multiple contaminants, which can then be disseminated more widely if these cultures are used by others.

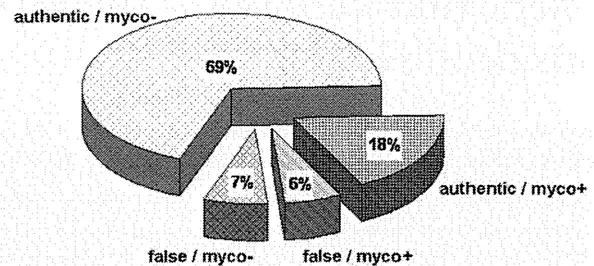



Figure 1. Rates of contamination for leukemia–lymphoma cell lines. Percentages of cross-contaminated and Mycoplasma-contaminated cell lines from a dataset of 598 leukemia and lymphoma cell lines analyzed by the German cell line bank DSMZ. "False/authentic" refers to the presence or absence of cross-contamination; "myco+/myco−" refers to the presence or absence of Mycoplasma contamination. Cell lines fall into the following categories: authentic/myco− ($n = 411$, 69%); authentic/myco+ ($n = 108$, 18%); false/myco− ($n = 41$, 7%) and false/myco+ ($n = 38$, 6%). (Courtesy of Hans Drexler, DSMZ.)

Other studies have pointed out that testing of cell lines is often infrequent, resulting in the failure to detect contaminated samples. John Ryan of Corning Life Sciences conducted surveys of seminar attendees in 1990, asking about Mycoplasma contamination; 50% were not currently performing testing and only 18% said they tested their cultures regularly. Almost 1 in 4 respondents (23%) had experienced Mycoplasma contamination, but with such a low level of testing, it is likely that the real figure was much higher.[18] Other data on cross-contamination were published in 2004 by researchers at the University of California, Berkeley, where Walter Nelson-Rees worked on this problem in the 1970s, focusing on the HeLa cell line.[19] Of 483 respondents to a questionnaire on cell line usage, 35% were using cell lines obtained from another laboratory rather than a cell line repository, but almost half of all respondents performed no testing for cross-contamination.[20]

A practical example of the consequences of cell line contamination can be found in a recent study published by Berglind *et al.*[21] The authors analyzed data within the UMD_p53 (2007) database, which includes information on the p53 status of 1,211 cell lines. Discrepancies were found in p53 status for 23% (88/384) of cell lines where data have been published by 2

independent laboratories. It is likely that many of these discrepancies arose due to work with cross-contaminated samples; the authors noted that many groups rely on previously published reports of a cell line's p53 status,[21] resulting in further confusion when interpreting results from these cell lines.

Cell banks have the expertise to detect such cross-contamination, and have been proactive in publishing reports of cross-contaminated cell lines,[22,23] in publishing test results online[24] and in developing new detection methods.[25–27] Unfortunately, however, cell banks have also reported reluctance from many researchers to deposit cell lines for distribution.[28] Such repositories specialize in the detection of cross-contamination and it is unlikely that most laboratories have comparable resources in this regard. In addition, many researchers obtain cell lines from one another, rather than approaching the originator or purchasing the cell line from a cell bank performing quality control testing. This may be faster or cheaper than obtaining cultures from a reputable source but the practice makes contamination more prevalent and harder to detect.

## Practical Responses

Having defined the problems, it is time to focus on what can be done. Several cancer-related journals, including the International Journal of Cancer, have recently responded to these issues by changing their policies to require evidence of authentication with all submitted manuscripts using continuous cell lines.[29,30] Their response underscores the need for laboratories to come to grips with cell line cross-contamination and misidentification. Every researcher involved in cell culture will have cell lines currently in culture, stored in liquid nitrogen or may be commencing work on a new cell line. Put practically, how can you know if your cell lines are cross-contaminated?

There are 2 important answers to this question:

1. Check the literature, for example, by searching the PubMed database using the cell line name and "cross-contamination."
2. Check your cultured cells. Unless a cell line has come directly from a repository or other laboratory performing identification testing, it should be tested on arrival, and all cultures should be periodically tested while in use, before cryopreservation and when thawed from liquid nitrogen.[31] A variety of methods are available for authentication; for human cell lines, short tandem repeat (STR) profiling is the current international reference standard and is recommended as an easy and economical way to confirm cell line identity by comparison to donor tissue or to other samples of the cell line held by laboratories worldwide.[26]

## Checking the Literature: A List of Cross-Contaminated Cell Lines

A 2004 survey of abstracts within the PubMed database would suggest that inappropriate usage of cross-contaminated cell lines is increasing,[20] despite many years of publication on this issue. It is possible that many researchers simply cannot find existing references to cross-contamination so, to make this already published work more accessible, we have surveyed the literature and other online resources for references to cell line contamination. The resulting list of cross-contaminated cell lines is included as Electronic Supporting Information.

To generate this list, the authors examined the PubMed database, references within other articles relating to this topic and the websites of 5 cell banks: the American Type Culture Collection (ATCC), DSMZ, European Collection of Cell Cultures (ECACC), Japanese Collection of Research Bioresources and the RIKEN Bioresource Center Cell Bank. A Wikipedia list of contaminated cell lines was also accessed (http://en.wikipedia.org/wiki/List_of_contaminated_cell_lines). Cross-contaminated cell lines are listed by name along with their species and cell type (both claimed and actual), the name of the contaminating cell line where identified, the reference in which this was reported and the PubMed ID number where available. Notes are also included for some cell lines. The list is made available in Excel spreadsheet or PDF format for easy accessibility.

The cell lines listed within this database are divided into 2 tables. Supporting Information Table 1 contains those cell lines where cross-contamination occurred as an early event, and thus where there is no original material remaining. Supporting Information Table 2 contains those cell lines where it is thought cross-contamination occurred as a late event and where original stocks may still exist. A full list of references is also given.

The current list of cross-contaminated cell lines (version 6.4) contains 360 cell lines, 346 in Supporting Information Table 1 and 14 in Supporting Information Table 2, drawn from 68 references. Cell lines affected are primarily human, although cultures from at least 8 other species are included, and come from a wide spectrum of tissue types. The cell or tumor type is given within the list where known; extensive work has been done by some cell banks and laboratories in this area to characterize the actual cell type or tumor type.[22,32] In some cases, this work has shown that a cell line carries the correct name but its cell or tumor type has been incorrectly identified, for example, the cell line RPMI-6666 was initially thought to have come from Hodgkin lymphoma but is now known to be an EBV-positive B-lymphoblastoid cell line.[22]

Common features for cross-contaminating cell lines within the current list are summarized in Table 1. It can be seen that most cross-contamination events have arisen from within the same species but a substantial minority (9%, 33/360) involved cross-contamination from a second species. For the intraspecies contaminants, all of those detected were human but it is likely that this relates to the difficulty of detecting intraspecies contaminants for nonhuman species. The commonest contaminant remains the HeLa cell line

Mini Review

Mini Review

**Table 1.** Cross-contaminating cell lines

| Type of contaminant | Number of cell lines affected |
|---|---|
| **Intraspecies** | |
| Human | 324 |
| Nonhuman | 0 |
| Interspecies | 33 |
| Correct name—incorrect cell type (misidentified)[1] | 3 |
| Total | 360 |
| **Contaminating cell line—12 most frequent** | **Number of cell lines affected** |
| HeLa (human cervical adenocarcinoma) | 106 |
| T-24 (human bladder carcinoma) | 18 |
| HT-29 (human colon carcinoma) | 12 |
| CCRF-CEM (human acute lymphoblastic leukemia) | 9 |
| K-562 (human chronic myeloid leukemia) | 9 |
| U-937 (human lymphoma) | 8 |
| OCI/AML2 (human acute myeloid leukemia) | 8 |
| Hcu-10 (human esophageal carcinoma)[2] | 7 |
| M14 (human melanoma) | 7 |
| HL-60 (human acute myeloid leukemia) | 6 |
| PC3 (human prostate carcinoma) | 6 |
| SW-480, SW620 (human colon carcinoma)[3] | 6 |

[1]For additional misidentified cell lines see Drexler *et al.*[22] [2]Hcu-10 carries the same genetic identity as Hcu-18, Hcu-22, Hcu-27, Hcu-33, Hcu-37 and Hcu-39; it is unclear which is the correct identity (see Electronic Supporting Information for reference). [3]SW480 and SW620 come from the same donor and therefore carry the same genetic identity (see Electronic Supporting Information for reference).

(29%, 106/360), followed by T-24 (5%, 18/360) and HT-29 (3%, 12/360).

It is important for such a list to be continually updated and feedback is welcome for this purpose. An earlier version of the database was released online by ECACC[31]; 6 cell banks have now agreed to make the database available online and to update this information where necessary. Current website addresses for access to the list of cross-contaminated cell lines are given in Table 2. In future, it is envisaged that the current list of misidentified cell lines will be included in a new initiative improving access to authentication data. The Standard Development Organization at the ATCC is in the process of producing an international standard for human cell line identification based on STR profiling (ATCC SDO Workgroup ASN-0002, manuscript submitted). Strict criteria for STR profiles derived from cancer cell lines are being developed. One consequence of this initiative is that funding is being sought for a quality controlled and curated cell line database with free access into which the database described here will be incorporated.

**Table 2.** Websites for ongoing access to the list of cross-contaminated cell lines

| Cell bank | Website address |
|---|---|
| ATCC | http://www.atcc.org/ |
| CellBank Australia | http://www.cellbankaustralia.com/ |
| DSMZ | http://www.dsmz.de/ |
| ECACC | http://www.hpacultures.org.uk/collections/ecacc.jsp |
| JCRB | http://cellbank.nibio.go.jp/ |
| RIKEN Bioresource Center Cell Bank | http://www.brc.riken.go.jp/lab/cell/english/guide.shtml |

## Checking Your Cultures: Authentication of Cell Lines

Even if a search of the literature shows no indication that a cell line is contaminated, it is still essential to test the sample that you are working with. Authentication testing should be considered in a positive light, as an essential part of good cell culture practice[33] and as an assurance for researchers, funding bodies and journals that the cell line used is a valid experimental model.[17]

There are a number of methods for testing cell line identity. When the issue of cross-contamination was first identified, HeLa contaminants were detected through a combination of isoenzyme and chromosomal analysis.[19,34] Both techniques continue to be used but there are also many newer molecular approaches. Commonly used authentication methods are summarized in Table 3; what factors should be considered when choosing between these methods?

The expertise of the laboratory holding the cell line is an important factor. For example, laboratories with experience in cytogenetics would have the skills to identify species through karyotype analysis and cell lines through the presence or absence of appropriate markers.[35] Although this is an older approach, it still allows clear identification of cell lines, and many cell banks have published karyotypic information on their cell lines to allow comparison to well-characterized stocks. It should be noted that tumor-derived cell lines can be surprisingly difficult to harvest for cytogenetic analysis[35] and are typically heteroploid making interpretation difficult: the experience of the operator is important for success.

The species of cell lines held within the laboratory is also important. Although some authentication methods can be used on more than 1 species, molecular methods such as STR profiling are only successful for a single species; other species will simply fail to amplify.[26] This may not be an issue for laboratories working only with human samples but clearly is a significant factor for groups working with rodent cell lines. In this regard, multilocus DNA fingerprint analysis has a clear advantage, since probes are able to hybridize to a wide variety of species.[25] Unfortunately, although successful within a single laboratory, it can be challenging to compare DNA fingerprints across several experimental runs, and it is difficult to exchange data among laboratories or for cell

Table 3. Commonly used methods for authenticating cell lines

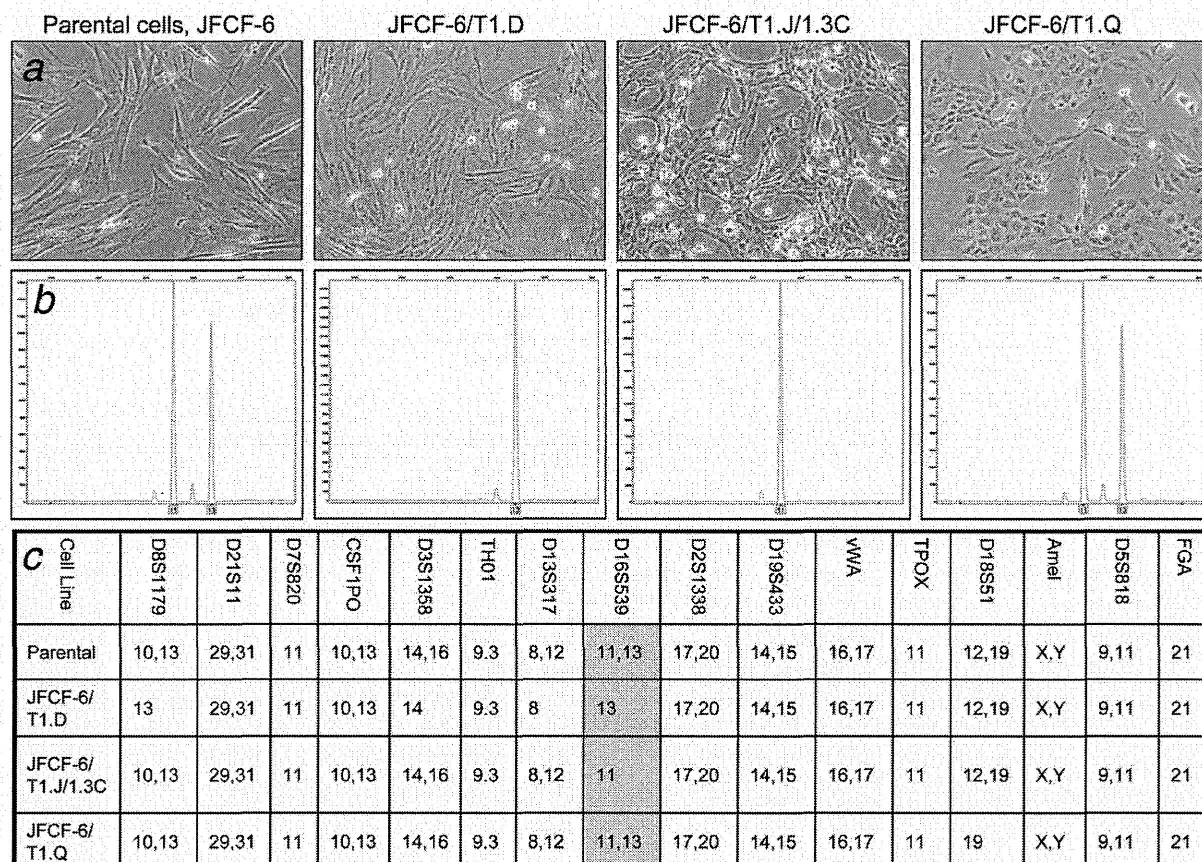
| Name | Description | Purpose | References |
|---|---|---|---|
| Chromosomal analysis/karyotyping | Involves preparation of a metaphase spread with chromosome banding and painting to identify chromosome number and markers | Separates species, plus individual cell lines if detailed analysis performed | Ref. 35 |
| Isoenzyme analysis | Biochemical method separating isoenzymes by electrophoresis; isoenzyme mobility may vary within or across species. Kits available include the Authentikit gel electrophoresis system | Separates species, sometimes individuals | Refs. 36,37 |
| Multilocus DNA fingerprint analysis | Molecular method detecting variation in length within minisatellite DNA containing variable numbers of tandem repeat sequences. Analysis is by Southern blot hybridization using probes 33.6 and 33.15, M13 phage DNA, or oligonucleotide sequence | Separates individual cell lines across multiple species | Refs. 25,38 |
| Short tandem repeat (STR) profiling | Molecular method detecting variation in length within microsatellite DNA containing variable numbers of short tandem repeat sequences. Analysis is by PCR with comparison to set size standards; usually available in a kit format allowing amplification of up to 16 loci | Separates individual cell lines within a single species | Refs. 26,39 |
| Polymerase chain reaction (PCR) fragment analysis | Molecular method involving amplification of specific genes or gene families, aiming to detect variations in exon/intron sequence, transcript splicing, or the presence of pseudogenes. Genes examined include the aldolase gene family and the beta-globin gene | Separates species only | Refs. 40,41 |
| Sequencing of "DNA barcode" regions | Involves sequencing of a DNA fragment from the mitochondrial gene cytochrome *c* oxidase subunit I, with comparison to sequence obtained from online databases. This "DNA barcode" has been shown in practice to distinguish a broad range of animal species | Separates species only | Refs. 27,42 |

banks to publish such fingerprints online. It is advisable to always compare the test sample to a known sample within the same experiment, ideally using DNA from the blood or tissue of the original donor.

The obvious advantage of STR profiling lies in the use of control samples to generate a numerical code for each sample, which precisely identifies that cell line and which can be readily shared and published online. It is primarily for this reason that STR profiling is recommended as an international reference standard for human cell lines[26] and accepted within the legal system for human identity testing.[39] STR profiling is based on the presence of STRs within the human genome that exist at variable lengths throughout the population. Each of the repeat regions to be analyzed (usually tetra or pentanucleotide repeats in noncoding sequence) is amplified by PCR using primers carrying fluorescent tags and electrophoresed in a sequencing gel; the precise length of each allele is determined and compared with size standards and controls. This allows identification software to assign a number to each allele at that locus (see, *e.g.*, Fig. 2). The combination of multiple loci—classically 13, as used in the FBI Laboratory's Combined DNA Index System (CODIS)—gives sufficient data to uniquely identify that individual.

STR profiles for individual cell lines and panels have now been reported by many laboratories (*e.g.*, Ref. 44) and are published online by several cell banks. However, there are some cautions to be aware of when using this approach. It is accepted within the forensic field that tumor samples are not as genetically stable as other tissue sources for STR profiling, because of loss of heterozygosity and microsatellite instability.[45,46] This is even more evident in tumor-derived cell lines, where evolution or genetic drift continues to occur with passage.[47] When searching an online database of STR profiles from cell lines, the user needs to look for close matches and not just identical matches; most studies would agree that 80% similarity is an appropriate threshold for declaring a match when comparing cell line profiles.[26,44] There may also be a significant start-up cost if testing in-house; in addition to an STR kit, access to methods for DNA extraction, precise quantitation, fragment analysis and software for STR profile identification is required.

The fact that STR profiling is only suitable for distinguishing cell lines of a single species has led to the need to re-examine authentication of nonhuman cell lines. Laboratory rodent samples will always be difficult to identify precisely due to inbreeding; laboratories working with rat or mouse cultures may wish to examine strain identity rather than authentication of individual cell lines, particularly if they have expertise in single nucleotide polymorphism (SNP) or single sequence length polymorphism (SSLP) analysis,




| Cell Line | D8S1179 | D21S11 | D7S820 | CSF1PO | D3S1358 | TH01 | D13S317 | D16S539 | D2S1338 | D19S433 | vWA | TPOX | D18S51 | Amel | D5S818 | FGA |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Parental | 10,13 | 29,31 | 11 | 10,13 | 14,16 | 9.3 | 8,12 | 11,13 | 17,20 | 14,15 | 16,17 | 11 | 12,19 | X,Y | 9,11 | 21 |
| JFCF-6/ T1.D | 13 | 29,31 | 11 | 10,13 | 14 | 9.3 | 8 | 13 | 17,20 | 14,15 | 16,17 | 11 | 12,19 | X,Y | 9,11 | 21 |
| JFCF-6/ T1.J/1.3C | 10,13 | 29,31 | 11 | 10,13 | 14,16 | 9.3 | 8,12 | 11 | 17,20 | 14,15 | 16,17 | 11 | 12,19 | X,Y | 9,11 | 21 |
| JFCF-6/ T1.Q | 10,13 | 29,31 | 11 | 10,13 | 14,16 | 9.3 | 8,12 | 11,13 | 17,20 | 14,15 | 16,17 | 11 | 19 | X,Y | 9,11 | 21 |

**Figure 2.** Example of STR profile generation and interpretation. An example of STR profiling is given for the JFCF-6 cell fibroblast strain and 3 of its immortalized derivatives, JFCF-6/T1.D, JFCF-6/T1.J/1.3C and JFCF-6/T1.Q.[43] Derivatives were established after transfection with SV40 early region DNA and were handled by CellBank Australia through its Culture and Return service. DNA from each culture was amplified using the AmpFlSTR Identifiler PCR Amplification Kit (Applied Biosystems, Mulgrave, Australia), which includes primers for 16 STR loci. Amplified sequence was analyzed using an ABI PRISM 3100 Genetic Analyzer and data files were assessed using GeneMapper ID software (Applied Biosystems). (*a*) Photographs taken of each culture, comparing parental cells to the morphology of each derived cell line (scale bar = 100 μm). Each derivative has a markedly different morphology, showing the need for authentication testing to confirm that derivatives correspond to the parental strain. (*b*) Examples of STR peak amplification for the D16S539 locus of each culture. Amplification varies at this locus due to genetic drift during establishment of the 3 JFCF-6–derived cell lines. The peaks shown correspond to specific allele sizes known to exist at this locus and confirmed using size standards and controls supplied with the kit (data not shown). (*c*). STR profiles for JFCF-6 and derived cell lines; the locus shown in B, D16S539, is highlighted in grey. Despite the differences seen due to genetic drift, the profiles for derived lines closely match the parental cell strain and all of these cultures are correctly identified.

which can be used for strain identification.[48,49] SNP analysis can also be used to identify individual samples[50] and has been used for cell line authentication,[51] making it a method of great promise for application to human and nonhuman samples alike. Laboratories working on specific cell types may be able to use expressed markers for identification, as 1 laboratory has done recently, publishing a technique for identification of hybridomas based on sequencing of light-chain variable regions.[52]

A simple method has recently emerged to help detect interspecies contamination. The term DNA barcoding here refers to amplifying a specific 648 bp fragment of the mitochondrial gene, cytochrome *C* oxidase subunit I (COI), using primers developed by Folmer *et al.*[53] Sequence divergences within this fragment allow species discrimination across almost all animal phyla.[42] Although debate is ongoing as to whether DNA barcoding is sufficient for assignment of species in taxonomic terms,[54] it is clear that the technique can readily identify the species of an unknown specimen if compared with previously sequenced reference material in online databases.[55] DNA barcoding has been tested for species identification of cell lines[27] and its use would reduce the incidence of interspecies cell line

contamination, found here to cause almost 1 in 10 of all published cross-contamination events.

Whatever the authentication method used, it should be clearly recorded within the researcher's experimental notes, and the result should be linked if possible to the laboratory's liquid nitrogen records, so that quality control for frozen vials is clearly evident. When publishing experimental work, the Material and Methods section should include the correct and full name of the cell line used, its origin (with appropriate references), the source of the cultures used and details of authentication testing.

## Conclusions

Cell line contamination is a serious issue that detracts from the use of cell lines as model systems to help us understand a broad range of diseases, including cancer. Responding practically by checking each cell line before it is used, searching for previous references and authenticating the sample itself is worthwhile and will reduce the risk and subsequent consequences of contamination long-term.

## Acknowledgements

The authors gratefully acknowledge the work of many cell banks and laboratories working in this area, and those responsible for compiling the list in Wikipedia, and regret that there is insufficient space to include all references here. A complete list of publications on cross-contamination can be found in the Electronic Supporting Information. Elsa Moy is thanked for her work in handling the cell lines shown in Figure 2. CellBank Australia was established by a joint venture of the Children's Medical Research Institute, Cure Cancer Australia Foundation and National Breast Cancer Foundation, and by an Enabling Grant of the National Health and Medical Research Council of Australia.

Mini Review

## References

1. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell* 2000;100:57–70.
2. Wistuba II, Behrens C, Milchgrub S, Syed S, Ahmadian M, Virmani AK, Kurvari V, Cunningham TH, Ashfaq R, Minna JD, Gazdar AF. Comparison of features of human breast cancer cell lines and their corresponding tumors. *Clin Cancer Res* 1998;4:2931–8.
3. Ross DT, Scherf U, Eisen MB, Perou CM, Rees C, Spellman P, Iyer V, Jeffrey SS, Van de RM, Waltham M, Pergamenschikov A, Lee JC, et al. Systematic variation in gene expression patterns in human cancer cell lines. *Nat Genet* 2000;24: 227–35.
4. Jones S, Chen WD, Parmigiani G, Diehl F, Beerenwinkel N, Antal T, Traulsen A, Nowak MA, Siegel C, Velculescu VE, Kinzler KW, Vogelstein B, et al. Comparative lesion sequencing provides insights into tumor evolution. *Proc Natl Acad Sci USA* 2008;105:4283–8.
5. Reddel RR. The role of senescence and immortalization in carcinogenesis. *Carcinogenesis* 2000;21:477–84.
6. Boehm JS, Hahn WC. Immortalized cells as experimental models to study cancer. *Cytotechnology* 2004;45:47–59.
7. Masters JR. Human cancer cell lines: fact and fantasy. *Nat Rev Mol Cell Biol* 2000;1: 233–6.
8. MacLeod RA, Nagel S, Scherr M, Schneider B, Dirks WG, Uphoff CC, Quentmeier H, Drexler HG. Human leukemia and lymphoma cell lines as models and resources. *Curr Med Chem* 2008;15:339–59.
9. Hughes P, Marshall D, Reid Y, Parkes H, Gelber C. The costs of using unauthenticated, over-passaged cell lines: how much more data do we need? *Biotechniques* 2007;43:575, 577–2.
10. van Staveren WC, Solis DY, Hebrant A, Detours V, Dumont JE, Maenhaut C. Human cancer cell lines: experimental models for cancer cells in situ? For cancer stem cells? *Biochim Biophys Acta* 2009; 1795:92–103.
11. van Pelt JF, Decorte R, Yap PS, Fevery J. Identification of HepG2 variant cell lines by short tandem repeat (STR) analysis. *Mol Cell Biochem* 2003;243:49–54.
12. Shimada Y. Researchers should have respect for the originator of the cell lines. *Clin Cancer Res* 2005;11:4634.
13. Drexler HG, Uphoff CC, Dirks WG, MacLeod RA. Mix-ups and mycoplasma: the enemies within. *Leuk Res* 2002;26: 329–33.
14. MacLeod RA, Dirks WG, Matsuo Y, Kaufmann M, Milch H, Drexler HG. Widespread intraspecies cross-contamination of human tumor cell lines arising at source. *Int J Cancer* 1999;83: 555–63.
15. Stacey GN. Cell contamination leads to inaccurate data: we must take action now. *Nature* 2000;403:356.
16. Masters JR. False cell lines: the problem and a solution. *Cytotechnology* 2002;39:69–74.
17. Nardone RM. Eradication of cross-contaminated cell lines: a call for action. *Cell Biol Toxicol* 2007;23:367–72.
18. Ryan JA. Understanding and managing cell culture contamination. Corning Technical Bulletin 1994. Available at: http://catalog2.corning.com/Lifesciences/media/pdf/cccontamination.pdf accessed on 18 August 2009.
19. Nelson-Rees WA, Flandermeyer RR. HeLa cultures defined. *Science* 1976;191: 96–8.
20. Buehring GC, Eby EA, Eby MJ. Cell line cross-contamination: how aware are mammalian cell culturists of the problem and how to monitor it? *In Vitro Cell Dev Biol Anim* 2004;40:211–5.
21. Berglind H, Pawitan Y, Kato S, Ishioka C, Soussi T. Analysis of p53 mutation status in human cancer cell lines: a paradigm for cell line cross-contamination. *Cancer Biol Ther* 2008;7:699–708.
22. Drexler HG, Dirks WG, Matsuo Y, MacLeod RA. False leukemia-lymphoma cell lines: an update on over 500 cell lines. *Leukemia* 2003;17:416–26.
23. Yoshino K, Iimura E, Saijo K, Iwase S, Fukami K, Ohno T, Obata Y, Nakamura Y. Essential role for gene profiling analysis in the authentication of human cell lines. *Hum Cell* 2006;19:43–8.
24. Dirks WG, MacLeod RA, Nakamura Y, Kohara A, Reid Y, Milch H, Drexler HG, Mizusawa H. Cell line cross-contamination initiative: an interactive reference database of STR profiles covering common cancer cell lines. *Int J Cancer* 2010;126:303–4.
25. Stacey GN, Bolton BJ, Morgan D, Clark SA, Doyle A. Multilocus DNA fingerprint analysis of cell banks: stability studies and culture identification in human B-lymphoblastoid and mammalian cell lines. *Cytotechnology* 1992;8:13–20.
26. Masters JR, Thomson JA, ly-Burns B, Reid YA, Dirks WG, Packer P, Toji LH, Ohno T, Tanabe H, Arlett CF, Kelland LR, Harrison M, et al. Short tandem repeat profiling provides an international reference standard for human cell lines. *Proc Natl Acad Sci USA* 2001;98: 8012–17.
27. Cooper JK, Sykes G, King S, Cottrill K, Ivanova NV, Hanner R, Ikonomi P. Species identification in cell culture: a two-pronged molecular approach. *In Vitro Cell Dev Biol Anim* 2007;43:344–51.
28. MacLeod RA, Drexler HG. Public repositories: users reluctant to give materials. *Nature* 2006;439:912.

Mini Review

29. Potash J, Anderson KC. What's your line? *Clin Cancer Res* 2009;15:4251.
30. Lichter P, Allgayer H, Bartsch H, Fusenig N, Hemminki K, von Knebel DM, Kyewski B, Miller AB, zur HH. Obligation for cell line authentication: appeal for concerted action. *Int J Cancer* 2010;126:1.
31. Freshney RI. Database of misidentified cell lines. *Int J Cancer* 2010;126:302.
32. Schweppe RE, Klopper JP, Korch C, Pugazhenthi U, Benezra M, Knauf JA, Fagin JA, Marlow LA, Copland JA, Smallridge RC, Haugen BR. Deoxyribonucleic acid profiling analysis of 40 human thyroid cancer cell lines reveals cross-contamination resulting in cell line redundancy and misidentification. *J Clin Endocrinol Metab* 2008;93:4331–41.
33. Balls M, Coecke S, Bowe G, Davis J, Gstraunthaler G, Hartung T, Hay R, Merten OW, Price A, Schechtman LM, Stacey G, Stokes W. The importance of Good Cell Culture Practice (GCCP). *ALTEX* 2006;23 Suppl.:270–3.
34. Gartler SM. Genetic markers as tracers in cell culture. *Natl Cancer Inst Monogr* 1967; 26:167–95.
35. MacLeod RA, Kaufmann M, Drexler HG. Cytogenetic harvesting of commonly used tumor cell lines. *Nat Protoc* 2007;2: 372–82.
36. O'Brien SJ, Shannon JE, Gail MH. A molecular approach to the identification and individualization of human and animal cells in culture: isozyme and allozyme genetic signatures. *In Vitro* 1980;16: 119–35.
37. Stacey GN, Hoelzl H, Stephenson JR, Doyle A. Authentication of animal cell cultures by direct visualization of repetitive DNA, aldolase gene PCR and isoenzyme analysis. *Biologicals* 1997;25:75–85.
38. Jeffreys AJ, Wilson V, Thein SL. Hypervariable "minisatellite" regions in human DNA. *Nature* 1985;314:67–73.
39. Butler JM. Genetics and genomics of core short tandem repeat loci used in human identity testing. *J Forensic Sci* 2006;51:253–65.
40. Liu M, Liu H, Tang X, Vafai A. Rapid identification and authentication of closely related animal cell culture by polymerase chain reaction. *In Vitro Cell Dev Biol Anim* 2008;44:224–7.
41. Steube KG, Koelz AL, Drexler HG. Identification and verification of rodent cell lines by polymerase chain reaction. *Cytotechnology* 2008;56:49–56.
42. Hebert PD, Cywinska A, Ball SL, deWaard JR. Biological identifications through DNA barcodes. *Proc Biol Sci* 2003;270:313–21.
43. Jiang WQ, Zhong ZH, Nguyen A, Henson JD, Toouli CD, Braithwaite AW, Reddel RR. Induction of alternative lengthening of telomeres-associated PML bodies by p53/ p21 requires HP1 proteins. *J Cell Biol* 2009;185:797–810.
44. Lorenzi PL, Reinhold WC, Varma S, Hutchinson AA, Pommier Y, Chanock SJ, Weinstein JN. DNA fingerprinting of the NCI-60 cell line panel. *Mol Cancer Ther* 2009;8:713–24.
45. Poetsch M, Petersmann A, Woenckhaus C, Protzel C, Dittberner T, Lignitz E, Kleist B. Evaluation of allelic alterations in short tandem repeats in different kinds of solid tumors–possible pitfalls in forensic casework. *Forensic Sci Int* 2004;145:1–6.
46. Vauhkonen H, Hedman M, Vauhkonen M, Kataja M, Sipponen P, Sajantila A. Evaluation of gastrointestinal cancer tissues as a source of genetic information for forensic investigations by using STRs. *Forensic Sci Int* 2004;139:159–67.
47. Parson W, Kirchebner R, Muhlmann R, Renner K, Kofler A, Schmidt S, Kofler R. Cancer cell line identification by short tandem repeat profiling: power and limitations. *FASEB J* 2005;19:434–6.
48. Witmer PD, Doheny KF, Adams MK, Boehm CD, Dizon JS, Goldstein JL, Templeton TM, Wheaton AM, Dong PN, Pugh EW, Nussbaum RL, Hunter K, et al. The development of a highly informative mouse Simple Sequence Length Polymorphism (SSLP) marker set and construction of a mouse family tree using parsimony analysis. *Genome Res* 2003;13: 485–91.
49. Petkov PM, Cassell MA, Sargent EE, Donnelly CJ, Robinson P, Crew V, Asquith S, Haar RV, Wiles MV. Development of a SNP genotyping panel for genetic monitoring of the laboratory mouse. *Genomics* 2004;83: 902–11.
50. Pakstis AJ, Speed WC, Fang R, Hyland FC, Furtado MR, Kidd JR, Kidd KK. SNPs for a universal individual identification panel. *Hum Genet* 2010;127:315–24.
51. Demichelis F, Greulich H, Macoska JA, Beroukhim R, Sellers WR, Garraway L, Rubin MA. SNP panel identification assay (SPIA): a genetic-based assay for the identification of cell lines. *Nucleic Acids Res* 2008;36:2446–56.
52. Koren S, Kosmac M, Colja VA, Montanic S, Curin SV. Antibody variable-region sequencing as a method for hybridoma cell-line authentication. *Appl Microbiol Biotechnol* 2008;78:1071–8.
53. Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol Mar Biol Biotechnol* 1994;3:294–9.
54. Linares MC, Soto-Calderon ID, Lees DC, Anthony NM. High mitochondrial diversity in geographically widespread butterflies of Madagascar: a test of the DNA barcoding approach. *Mol Phylogenet Evol* 2009;50:485–95.
55. Dawnay N, Ogden R, McEwing R, Carvalho GR, Thorpe RS. Validation of the barcoding gene COI for use in forensic genetic species identification. *Forensic Sci Int* 2007;173:1–6.

*Nature Reviews Cancer* | AOP, published online 7 May 2010; doi:10.1038/nrc2852



MODELS OF CANCER SERIES — SCIENCE AND SOCIETY

# Cell line misidentification: the beginning of the end

*American Type Culture Collection Standards Development Organization Workgroup ASN-0002*

Abstract | Cell lines are used extensively in research and drug development as models of normal and cancer tissues. However, a substantial proportion of cell lines is mislabelled or replaced by cells derived from a different individual, tissue or species. The scientific community has failed to tackle this problem and consequently thousands of misleading and potentially erroneous papers have been published using cell lines that are incorrectly identified. Recent efforts to develop a standard for the authentication of human cell lines using short tandem repeat profiling is an important step to eradicate this problem.

Cell lines are used extensively in biomedical research as *in vitro* models. The validity of the data obtained often depends on the identity of the cell line, particularly when it is being used as a surrogate for the tissue of origin. Surprisingly, the frequency of cell line misidentification is high, and consequently the ascribed origin of a cell line is often incorrect. This problem has been known for over 50 years and has been described as the most compelling quality-control issue confronting the scientific community[1]. Based on analyses of cell lines submitted to international cell banks, the incidence of misidentification in 1977 was 16%[2] and in 1999 was 18%[3]. Until recently, the authenticity of cell lines used in biomedical research has received little attention. This Science and Society article has been written by the members of the American Type Culture Collection (ATCC) Standards Development Organization (SDO) Workgroup ASN-0002 (BOX 1), a working group currently developing a standard for human cell line authentication. The ATCC SDO was formed in 2007 to develop best practices (standards) for use in the life sciences and to promote their use globally, using a consensus-driven process that balances the viewpoints of industry, government, regulatory agencies and academia. We expect that the draft standard (BOX 2) will be available for public review and comment in 2010 and subsequently the final draft will be approved by the American National Standards Institute (ANSI).

Here we describe the causes and scientific effects of cell line misidentification, its history and the efforts taken to solve the problem. The various methods currently available for authenticating cell lines are discussed and a recommendation is made for the use of short tandem repeat (STR) profiling for authenticating human cell lines. Perhaps of the greatest importance, a universal database of human cell line STR profiles is under construction.

### Discovery of cell line misidentification

Misidentification of human and animal cell cultures is a long-standing problem, and awareness of this problem dates back to the 1950s (TIMELINE). Karyotyping and immunological approaches were first used for cell line authentication[4–6]. Extensive species misidentification was reported, leading to the establishment of a bank of authenticated cell lines at the ATCC in 1962.

Misidentification within species could not be detected in 1962, but in 1966 Stanley Gartler (FIG. 1a) introduced the concept of biochemical polymorphisms to distinguish human cell lines on the basis of their isozyme expression. At the Second Decennial Review Conference on Cell, Tissue and Organ Culture in 1966, Gartler reported that 18 human cell lines supposedly of independent origins were all HeLa cells[7], the first human cancer cell line to be established in culture[8]. The examples included cells claimed to be derived from normal intestinal epithelium (Int-407), normal amnion (WISH), normal liver (Chang liver), laryngeal cancer (Hep-2) and oral cancer (KB). The HeLa cell line was derived from a glandular cervical cancer in a female patient named Henrietta Lacks and, because of its celebrated status, was distributed internationally and passed from laboratory to laboratory. Then, as today, many scientists were oblivious to the possibility of cross-contamination. HeLa cells are particularly robust and fast-growing and consequently can rapidly overgrow other cells.

### Denial and complacency

There was resistance and some hostility to Gartler's findings — even among scientists "who should have known better", according to Gartler — but one scientist, Walter Nelson-Rees (FIG. 1b), took particular note of Gartler's talk. Nelson-Rees ran a cell bank at Berkeley under contract for the National Cancer Institute. With his colleagues he developed karyotyping methods for authenticating cell lines and in a series of papers he showed there was extensive cross-contamination among the supposedly unique cultures sent to the bank (for example, see REF. 9). Nelson-Rees's work showed widespread cross-contamination by HeLa cells and for some years all cell lines were under suspicion of being HeLa cells until proven otherwise. He developed methods for cell identification and raised awareness of the problem in the scientific literature and through correspondence with individual scientists affected by the problem. Nelson-Rees's last contribution to the subject was published in 2009, soon after his death[10].

When Nelson-Rees first published his findings, some scientists ignored or denied the evidence and continued to publish papers containing false information[11]. As a consequence, Nelson-Rees felt that he had no option but to highlight the papers (and consequently the individuals) using cross-contaminated cell lines. At that time (and possibly today), Nelson-Rees's behaviour was



© 2010 Macmillan Publishers Limited. All rights reserved

regarded as unscientific and he was attacked by many colleagues. He was branded a self-appointed vigilante and his contract terminated by the National Institutes of Health (NIH) in 1981. After this, cell line misidentification went largely unchecked and the problem escalated. For the next 10–20 years, cell banks distributed many cell lines under their false names.

Estimating how much misleading and erroneous research is attributable to cross-contamination or misidentification of cell lines has been difficult. The use of misidentified cultures increased about 10-fold in the PubMed database (see Further information for a link) between 1969 and 2004, and the papers that used cultured cells increased only 2–2.5-fold during the same time period[12,13]. By 2004, HeLa was just the tip of the iceberg, and many other cell lines masqueraded under various guises in laboratories worldwide.

A survey that profiled active cell culture workers found that of 483 respondents, 32% used HeLa cells, 9% unwittingly were using HeLa contaminants, only 33% of the investigators tested their cell lines for authenticity and 35% obtained their cell lines from other laboratories rather than from a major repository[12].

Although complacency and, in some cases, denial have been the primary responses to cell line misidentification over the past five decades, a few individuals have devoted a great deal of personal effort into remediation of the problem. Among the largely independent efforts were letters to editors from concerned individuals requesting that readers be alerted about the problem, and that authors be required to provide evidence that the cell lines used in their studies were neither cross-contaminated nor misidentified. These efforts were largely ignored in the period after Nelson-Rees's contract was terminated, despite the development of DNA-fingerprinting techniques, which brought new and more reproducible methods that once again revealed the extent of cell line misidentification in the early 1990s[14].

Roland Nardone (FIG. 1c) started the second crusade in 2004. He gained the support of Joseph B. Perrone, who was then Vice President for Standards at ATCC and provided ideas and the matching outrage needed to fuel the crusade. Together with other concerned scientists, Nardone developed a comprehensive and coordinated initiative that simultaneously sought to raise awareness of the nature and magnitude of the problem and canvassed the involvement of individuals and organizations concerned or affected by the problem[1,15]. Such organizations included the NIH, the Howard Hughes

Box 1 | **Authors and members of workgroup ASN-0002**


- Christine Alston-Roberts, Standards Specialist, ATCC, 10801 University Boulevard, Manassas, VA 20110, USA
- Rita Barallon, Ph.D., Service Business Manager, Life and Food Sciences Life Sciences, LGC, Queens Road, Teddington, Middlesex, TW11 0LY, UK
- Steven R. Bauer*, Ph.D., FDA/Center for Biologics Evaluation and Research, Chief, Cell and Tissue Therapy Branch, Division of Cellular and Gene Therapies, Office of Cellular, Tissue and Gene Therapies, NIH Building 29B 2NN10 HFM-740, 8800 Rockville Pike, Bethesda, MD 20892, USA
- John Butler, Ph.D., Biochemical Science Division (831), Advanced Chemical Science Laboratory (227), Room B226, NIST, 100 Bureau Drive, Stop 8312, Gaithersburg, MD 20899-8312, USA
- Amanda Capes-Davis, Ph.D., CellBank Australia, Children's Medical Research Institute, Westmead, New South Wales, Australia
- Wilhelm G. Dirks, Ph.D., Molecular Biology, DSMZ — German Collection of Microorganisms and Cell Cultures, Inhoffenstr. 7b, 38124 Braunschweig, Germany
- Eugene Elmore, Ph.D., Project Scientist, Department of Radiation Oncology, University of California, Medical Sciences I, B146D, Irvine, CA 92697, USA
- Manohar Furtado, Ph.D., Vice President, R & D, Applied Markets Division, Applied Biosystems, 850 Lincoln Centre Drive, MS404-1, Foster City, CA 94404, USA
- Liz Kerrigan, Director, Standards and Certification, ATCC, 10801 University Boulevard, Manassas, VA 20110, USA
- Margaret C. Kline, Research Biologist, Biochemical Science Division (831), Advanced Chemical Science Laboratory (227), Room B226, National Institutes of Standards and Technology,100 Bureau Drive, Stop 8312, Gaithersburg, MD 20899-8312, USA
- Arihiro Kohara, Ph.D., Scientist, National Institute of Biomedical Innovation, Department Biomedical Services, Laboratory of Cell Cultures, 7-6-8 Saito-Asagi, Ibaraki, Osaka, Japan 567-0085
- Georgyi V. Los, M.D., Ph.D., Honorary Fellow, Neuroscience Training Program, University of Wisconsin-Madison, 1300 University Avenue, Madison, WI 53706, USA
- Roderick A.F. MacLeod, Ph.D., Cytogenetics Laboratory, DSMZ — German Collection of Microorganisms and Cell Cultures, Inhoffenstr. 7b, 38124 Braunschweig, Germany
- John R. W. Masters, Ph.D., FCRPath, Professor of Experimental Pathology, University College London, 67 Riding House Street, London, W1W 7EJ, UK
- Mark Nardone, Director, Bio-Trac Program, The Foundation for the Advanced Education in the Sciences at the National Institutes of Health, Bethesda, MD 20892, USA
- Roland M. Nardone, Ph.D., Professor Emeritus, Catholic University of America, Cell and Molecular Biology, 620 Michigan Avenue NE, Washington, DC 20064, USA
- Raymond W. Nims, Ph.D., Consultant, RMC Pharmaceutical Solutions Inc., 2150 Miller Drive, Suite A, Longmont, CO 80501, USA
- Paul J. Price, Ph.D., CSO, Research and Development, Room B-33, D-Finitive Cell Technology, 1023 Wappoo Rd, Charleston, SC 29407, USA
- Yvonne A. Reid, Ph.D., Collection Scientist, Cell Biology Collection, ATCC, 10801 University Boulevard, Manassas, VA 20110, USA
- Jaiprakash Shewale, Ph.D., Director, Biology, Applied Markets/Genetic Systems, Life Technologies, 850 Lincoln Centre Drive, Foster City, CA 94404, USA
- Anton F. Steuer, Ph.D., Principal Scientist, BioReliance, 14920 Brochart Road, Rockville, MD 20850, USA
- Douglas R. Storts, Ph.D., Head of Research, Nucleic Acid Technologies, Promega Corporation, 2800 Woods Hollow Road, Madison, WI 53711, USA
- Gregory Sykes, Biologist, ATCC, 10801 University Blvd., Manassas, VA 20110, USA
- Zenobia Taraporewala*, Ph.D., FDA/Center for Biologics Evaluation and Research, Reviewer, Division of Cellular and Gene Therapies, Office of Cellular, Tissue, and Gene Therapies, 1401 Rockville Pike, Room 200N, Rockville, MD 20892, USA
- Jim Thomson, Innovation and Support Team, LGC, Queens Rd, Teddington, TW11 0LY, UK

*S. R. B. and Z. T. did not contribute as authors to this Perspective.




© 2010 Macmillan Publishers Limited. All rights reserved

Medical Institute, heads of funding organizations and their attorney generals, leaders of professional societies and editors of science journals.

Copies of a white paper, 'Eradication of cross-contaminated cell lines: a call for action' (subsequently published by Nardone in 2007 (REF. 15)) were distributed to thousands of scientists. The white paper presented what seemed to be a straightforward solution: funding agencies would require cell line authentication as a condition for the receipt of funds and journals would have a similar requirement for manuscripts submitted for publication. This approach was met initially with indifference. Nevertheless, over a period of 4 years, several substantial milestones were reached[1]. An open letter[16] to Michael O. Leavitt, Secretary of Health and Human Resources, led the NIH to re-examine the situation. On November 28 2007, the NIH published an addition to its guidelines for research in the form of a notice regarding authentication of cultured cell lines calling for diligence and more careful peer review[17].

Two factors have driven this progress. One is heightened awareness. The other is the outrage of scientists angered by the failure of funding agencies and journals to address the problem and allowing it to fester and amplify for 50 years. Many scientists now accept the need for a standardized method of human cell line authentication to satisfy the new requirements. ASN-0002 will be the first step towards a universally adopted standard.

**Box 2 | ATCC SDO standards development process**

- American Type Culture Collection (ATCC) Standard Development Office (SDO) Consensus Standards Partnership (CSP) members recommend a new standard.
- Recommendation forwarded to ATCC SDO steering committee for review and vote.
- Project Initiation Notification System (PINS) published in *American National Standards Institute (ANSI) Standards Action* for 30-day public comment period, concurrent with CSP (ATCC SDO members) review.
- Recommendation for workgroup chair(s) sent to ATCC SDO steering committee for vote.
- Workgroup established; (ASN-0002), which includes stakeholders from academia, industry and government, and proceeds to draft the standard (see BOX 1 for members of the workgroup).
- ASN-0002 workgroup forwards draft standard to steering committee for internal review. Workgroup edits draft standard and forwards to ANSI and CSP (ATCC SDO membership) for concurrent 45-day public review.
- ASN-0002 workgroup responds to all comments and resolves any differences. If there are no substantial changes to the standard, the standard is submitted to the ANSI board of standards review for final action and publication as an ANSI-approved standard.

**Examples and impact**

Cross-contamination and misidentification have a long history with many examples, but it is difficult to judge which have been the most substantial and costly.

The classic case already described is contamination by HeLa cells, of which there are several examples (see REFS 7,9 for example). It is astonishing that many of these cell lines have continued to be used under their false descriptions in respected journals for over 40 years after they were first shown to be HeLa cells (BOX 3).

T24 is another fast-growing cell line that has contaminated many supposedly distinct bladder cancer cell lines (BOX 3). ECV304 was originally claimed to be a spontaneously transformed human normal endothelial cell line, but later shown to be T24 bladder cancer cells[18]. Surprisingly, the demonstration that ECV304 cells are not endothelial cells had little effect on its use as a model for endothelial cells in publications (FIG. 2).

The putative human prostate cancer cell lines TSU-Pr1 and JCA-1 are also derived from T24 bladder cancer cells[19]. These findings were published in *Cancer Research*, but that did not prevent TSU-Pr1 cells being used as a prostate cancer cell model in a later paper in *Cancer Research* (BOX 3).

DNA-fingerprinting analysis revealed that the NCI/ADR-RES cell line was actually an ovarian tumour cell line, OVCAR-8, rather than a breast cancer cell line. Around 300 papers have been published using the incorrect identification of the NCI/ADR-RES cell line[20]. NCI/ADR-RES is included in the NCI60 panel of cell lines, which has been subject to STR profiling (discussed below)[21].

A paper describing misidentification of oesophageal cell lines stated "Experimental results based on these contaminated cell lines have led to ongoing clinical trials recruiting EAC [oesophageal adenocarcinoma] patients, to more than 100 scientific publications, and to at least three National Institutes of Health cancer research grants and 11 US patents"(REF. 22).

The consequences of widespread misidentification and cross-contamination of cell lines are immeasurable. In addition to the waste of millions of dollars of public money, time and intellectual resources, there is the loss of confidence in published work, and the integrity of science suffers.

**Over 50 years of suppression. Why?**

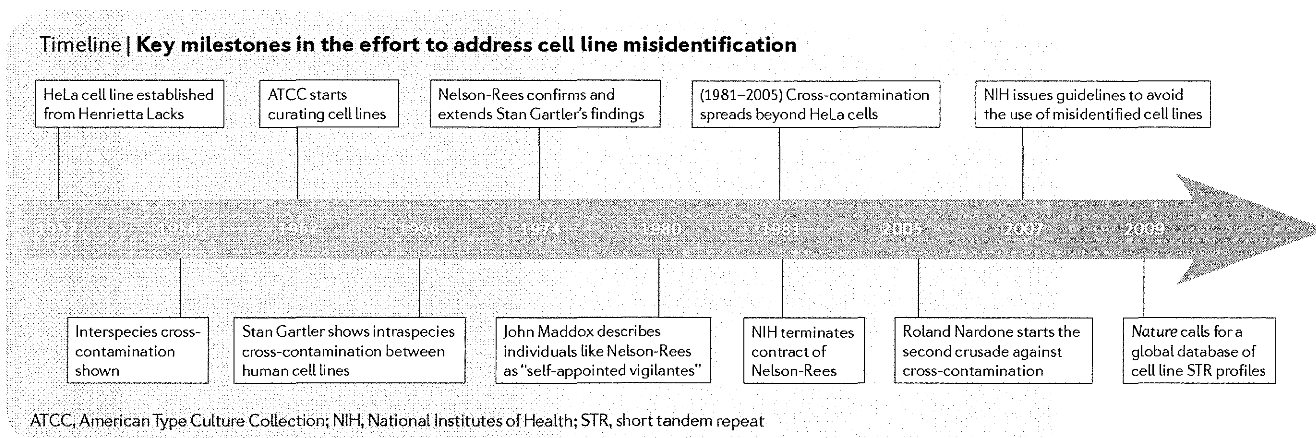
Three constituencies share responsibility for cell line misidentification — individual scientists, scientific journals and funding agencies. For most of the past 50 years it is only individual scientists who have addressed the issue. Nevertheless, it is hard to escape the conclusion that many scientists have knowingly used misidentified cell lines in publications (for example, the evidence in FIG. 2). Furthermore, authors are often reluctant to publish corrections to the literature based on cell line misidentification.

John Maddox, the editor of *Nature* in 1980, wrote an editorial about a high-profile case of cross-contamination entitled 'Responsibility for trust in research' (REF. 23). With an almost complete lack of insight into the problem he suggested that "there is no reason to suppose that the few cases [of cross-contamination] that have come to light are in any sense the tip of the iceberg". In the same editorial, scientists like Nelson-Rees were vilified, as the article made the point that it would be tragic if these civilized habits (that is, truth in research) "were to be corrupted by the activities of self-appointed vigilantes". The history of cell line cross-contamination indicates that truth and trust are not as universal among the scientific community as many scientists wish to believe.

The responses of editors of scientific journals to the problem continue to be illuminating. There have been hundreds of papers in scientific journals describing examples of misidentification and, until recently, no remedial action has been taken to eradicate the problem by journals or funding agencies. The editor of an influential tissue culture journal was asked to consider introducing authentication as a requirement for publication and replied that it would be financial suicide. Editors of other journals also refused to consider such quality-control measures on the basis that introducing such a hurdle to publication would substantially reduce the number of authors willing to submit manuscripts to their journal.



© 2010 Macmillan Publishers Limited. All rights reserved




Timeline | **Key milestones in the effort to address cell line misidentification**

ATCC, American Type Culture Collection; NIH, National Institutes of Health; STR, short tandem repeat

Over the past 2 years attitudes have begun to change, with journals, such as *In Vitro Cellular and Developmental Biology*, *International Journal of Cancer*, *Cell Biochemistry and Biophysics* and the American Association for Cancer Research (AACR) journals, demanding that all cell lines are authenticated before publication. *Nature* has indicated that first the funding organizations have to demand authentication and provide the necessary funds. Once they do, *Nature* will require cell line identification prior to publication[24]. In the meantime, the funding organizations continue to ignore the problem.

The constituency with the most power to maintain standards in science is the funding agencies. Surprisingly, these have been resistant to addressing or even acknowledging the problem of cell line misidentification. For example, the NIH advisory note issued in 2007 ignores the fact that individual scientists and reviewers have failed to overcome this problem. As an editorial in *Nature* pointed out, the advisory note merely enforces the status quo[24].

Attempts to address the problem by individual scientists have met with unhelpful responses from funding bodies, which have tended to deny or belittle the problem. A recent public statement by a senior scientist from Cancer Research UK made light of cell line misidentification, saying that "this issue raises its head every few years". Funding bodies seem to be threatened by the issue and are resistant to engaging with scientists who try to address the problem and often attempt to disparage and discredit those who try to find a solution.

Any of the major funding organizations that support biomedical research in the United States or United Kingdom could have eradicated cell line misidentification during the past 10 years for less than the cost of the average project grant by funding the measures outlined in this Science and Society article. Yet, these funding agencies have repeatedly ignored and in some cases suppressed debate, and continued to provide grants for research using false cell lines. There could be wider implications concerning the role of funding agencies in the control of scientific misrepresentation and fraud.

Zero tolerance of cell line misidentification is needed from both journals and funding agencies. There are signs that Nardone's crusade is gaining influence and the standard for human cell line identification will be tangible evidence of Nardone's legacy.

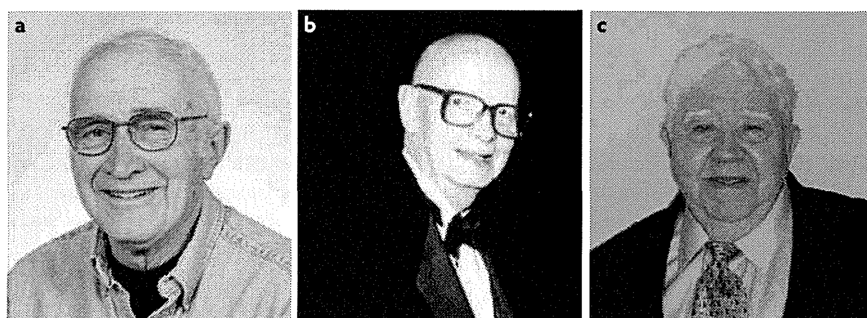



Figure 1 | **Pioneers of awareness of cell line misidentification. a** | Stanley Gartler **b** | Walter Nelson-Rees **c** | Roland Nardone

### Causes of cell line misidentification

Most cell lines are established in academic environments in which tissue culture is often regarded as a technique requiring little skill and essential facilities, such as flow cabinets and incubators, are used without restriction. In these circumstances, it is not surprising that attempts to establish new cell lines often lead to cross-contamination. Among 550 leukaemia and lymphoma cell lines submitted to the Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ; German Collection of Microorganisms and Cell Cultures; please see Further information for a link) cell bank, 59/395 (15%) submitted by originators and 23/155 (15%) submitted by secondary sources were false[25]. Presumably most of the cell lines submitted by the secondary sources had also been cross-contaminated or misidentified by the originators.

There are many causes of cell culture misidentification and every laboratory is at risk. Perhaps the most straightforward cause is mislabelling of a cell culture vessel during routine manipulation. Factors contributing to this error include operator workload, lack of attention, or distractions during manipulation of cell lines.

Cross-contamination of a culture and subsequent overgrowth by the contaminating cell type is another frequent cause of cell line misidentification. The chances of this occurring are increased by the use of shared reagents, repeated use of the same pipette during re-feeding operations and manipulation of multiple cultures at the same time without adequate isolation of one cell type from another. When cross-contamination happens, one cell type may



© 2010 Macmillan Publishers Limited. All rights reserved

rapidly outgrow the other, leading to a pure culture of the contaminating cells in four or five passages[26].

Intentional co-cultivation during propagation of human stem or primary cells using a feeder layer derived from another species (such as mouse 3T3 cells) can result in cross-contamination and overgrowth of the human cell line. Normally, feeder cells are rendered incapable of proliferating, but if the growth arrest procedure is inadequate, the feeder cells can proliferate and displace the human cells. Somatic cell hybridization is unusual but can occur, as found in the human mantle cell lymphoma line NCEB-1, which carries seven mouse chromosomes[27].

Xenografting can also lead to cell line cross-contamination and misidentification[28]. Recovered cell lines from xenografts can be replaced by cells derived from the host animal.

In general, cross-contamination results in the complete and rapid displacement of the less fit cell type. Two cell lines cannot co-exist in the same culture environment for extended periods unless there is a symbiotic relationship, which as far as we know has never been reported. Consequently, cell mixtures are discovered rarely. The only known situation in which a cell population contains a stable mixture of genomes over many passages is following somatic cell hybridization.

Simple, cheap quality-control measures can prevent or at least minimize the consequences of misidentification. Misidentification is rife because of a combination of lack of awareness and the failure to include quality-control measures. The extensive quality-control measures demanded of the biopharmaceutical industry and mandated in the applicable regulatory documents are believed to have contributed to the relatively low frequency of cell line misidentification reported in this industry[29].

### Detection of cross-contamination

Many methods have been used to detect cross-contamination, including isoenzyme analysis, karyotyping, human leukocyte antigen (HLA)-typing, immunotyping and DNA fingerprinting. These methods can authenticate a cell line, but with differing levels of ambiguity and powers of discrimination (Supplementary information S1 (table)). However, the data produced by these methods are not sufficiently reproducible between laboratories to allow any of them to be used for a standardized reference database.

Many laboratories have adopted STR profiling to identify human cell lines. STR profiling is the method used by forensic analysts and depends on the simultaneous amplification of multiple stretches of polymorphic DNA in a single tube. STR loci consist of repetitive DNA sequences that have varying numbers of repeats. Each STR locus can be amplified and the amplified products labelled with fluorophores of different colours, making the products easy to distinguish by size and colour (FIG. 3). STR analysis is rapid, inexpensive, amenable to automation and generates reproducible data in a format suitable for a standard reference database. For the quick, unambiguous authentication of cell lines, STR analysis has the greatest value.

### STR profiling — potential and limitations

DNA repeat sequences of 3–5 bases have been used routinely for paternity testing, forensic casework, and the identification of victims of mass disaster for more than two decades[30–33]. Consequently, STR profiling was applied to cell line identification[34–36]. There are several advantages to using STRs for the authentication of human cell lines (Supplementary information S2 (box)).

Cancer cell lines contain many genetic alterations, and therefore the criteria used to compare them using STR profiling must be different to those used for normal tissue (Supplementary information S3 (box)). Cancer cells often show loss of heterozygosity (that is, loss of an allele, which cannot be distinguished easily from homozygosity) and can contain multiple copies of alleles owing to DNA duplication. Similarly, during culture, cancer cell lines can lose or more rarely gain a copy of an allele (for examples, see REF. 34). Consequently, sub-lines of the same cell line may not have identical STR profiles.

Box 3 | **Examples of the use of cell lines under false descriptions**

The examples discussed below were picked at random from PubMed searches. The impact of the false descriptions ranges from minor to invalidation of the conclusions. The individual authors have been failed by peer review. The papers indicate that the editors and some of the reviewers of these journals (and by inference most scientific journals) are unaware of the extent of cell line misidentification, and indicate a general lack of awareness throughout the scientific community.

**HeLa cervical cancer cells**

- Int-407 (described as "non-transformed intestinal epithelial cells") in *Br. J. Cancer* **101**, 1596 (2009), *EMBO J.* **22**, 5003 (2003) and *J. Biol. Chem.* **280**, 13538 (2005)
- WISH (described as "non-transformed amniotic epithelial cells") in *Mol. Pharmacol.* **69**, 796 (2006), *Endocrinology* **147**, 2490 (2006) and *J. Biol. Chem.* **278**, 31731 (2003)
- Chang liver (described as "normal liver cells") in *Oncogene* **28**, 3526 (2009), *Proteomics* **14**, 2885 (2008) and *J. Biol. Chem.* **279**, 28106 (2000)
- HEp-2 (described as "laryngeal cancer") in *Investig. New Drugs* **26**, 111–118 (2008), *Carcinogenesis* **29**, 1519 (2008) and *J. Biol. Chem.* **283**, 36272 (2008)
- KB (described as "oral cancer") in *Biochem. Pharmacol.* **73**,1901–1909 (2007), *Clin. Cancer Res.* **14**, 8161(2008) and *J. Biol. Chem.* **280**, 23829 (2005)
- HeLa, Int-407 and HEp-2 cells were used as three distinct cell lines in the same study in *Cancer Res.* **69**, 632 (2009)

The scientists that use these cell lines sometimes use them under their false descriptions in many publications. For example, one group has used Int-407 as a model of normal intestinal cells since 1988 and during the past 10 years has published in the *Biochemical Journal* (2 papers), *Biochemical Society Transactions*, *British Journal of Cancer*, *Cancer Research* (2 papers), *Carcinogenesis*, *Experimental Cell Research* (3 papers), *Gastroenterology* (2 papers), *Journal of Biological Chemistry* (3 papers), *Journal of Cell Physiology*, *Journal of Cell Science* (3 papers), *Oncogene*, *PLoS One* and several other journals.

**T24 bladder cancer cells**

In 1999, ECV304 cells (originally described as spontaneously immortalized normal endothelial cells) were shown to be T24 cells[18].

Yet, many papers continue to describe ECV304 cells as endothelial, for example *Nature Immunol.* **6**, 497 (2005) and *Nature Biotechnol.* **25**, 921 (2007). Some studies use ECV304 cells in endothelial research without claiming that they are endothelial cells, but not stating that they are T24 bladder cancer cells, such as *Proc. Natl Acad. Sci. USA* **106**, 6849 (2009). Some studies have used T24 and one or more of its cross-contaminants as distinct bladder cancer cell lines, for example *J. Urol.* **181**, 1372–1380 (2009). Some studies describe ECV304 as bladder cancer cells, but fail to state that they are T24 cells, such as *J. Biol. Chem.* **285**, 555–564 (2010).

In *Cancer Research* in 2001, it was shown that TSU-Pr1 are T24 bladder cancer cells (*Cancer Res.* **61**, 6340–6344 (2001)). In the same journal, less than 3 years later, TSU-Pr1 cells were used as a prostate cancer model (*Cancer Res.* **64**, 1058–1066 (2004)). TSU-Pr1 continue to be used in some studies as a model for prostate cancer, such as *Endocrinology* **147**, 530–542 (2006) and *Cancer Cell* **5**, 67 (2004).



© 2010 Macmillan Publishers Limited. All rights reserved

Comparing identical alleles, a threshold of 75% identity has distinguished all known cross-contaminated cell lines in published datasets, and no two cell lines thought to be derived from different individuals showed more than 50% identity[21,34]. Consequently, there is a comfortable cushion of 25% between cell lines that are unique and those that show evidence of cross-contamination. Any cell line found with an identity level between 50 and 75% should be regarded with suspicion.

Major issues in the interpretation of genotypes from human cell lines include heterozygote peak height imbalance (that is the peak height or area of one allele is much larger than the peak height or area of the second allele), multiple alleles at a locus, and allele dropout (no amplification product of the expected size). Cancer cell lines are aneuploid and consequently STR profiles typically show heterozygous peak height imbalances and/or multiple alleles at one or more loci.

The cost of genotyping is a major concern, but trivial in relation to the cost of the work being done with the cell line. The cost of STR profiling includes DNA extraction, polymerase chain reaction (PCR) amplification of STR loci, separation of amplified products by capillary electrophoresis and data analysis. Increasing the number of STR loci, for example, from 6 to 15 would achieve a much higher power of discrimination (Supplementary information S4 (table)).

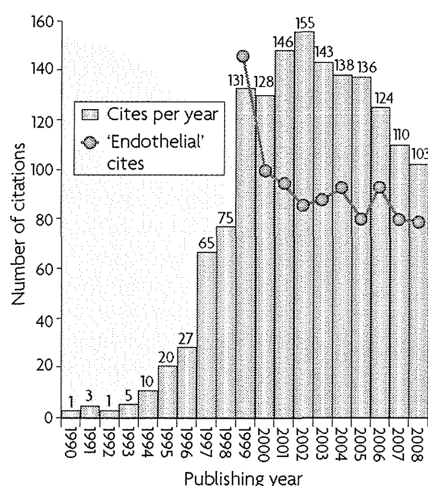



Figure 2 | **Citations of T24 bladder cancer cells referred to as normal endothelial cells.** The demonstration that ECV304 cells are not endothelial cells had little effect on its use as a model for endothelial cells in publications, as shown by the graph. Data generated courtesy of R.A.F. MacLeod, National Institute of Standards and Technology.

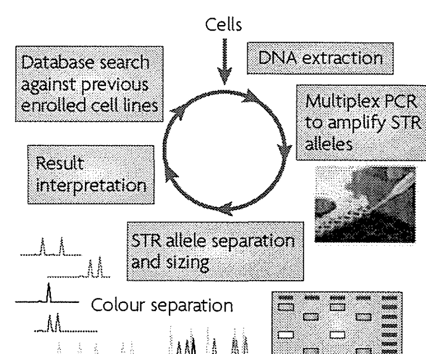



Figure 3 | **Short tandem repeat profiling methodology.** Short tandem repeat (STR) loci consist of repetitive DNA sequences with varying numbers of repeats. Each STR locus can be polymerase chain reaction (PCR) amplified and the amplified products labelled with fluorophores of different colours, making the products easy to distinguish by size and colour. Images courtesy of J. Butler, National Institute of Standards and Technology.

A major limitation of STR profiling is that it will not detect contaminating cells of another species, although if human cells are overgrown by cells of another species, the DNA will not amplify using human or higher primate-specific STR primers. PCR using species-specific primers can be used to detect contaminating cells from other species. If STR profiles have been established for the other species (currently restricted to a few commercially important species), STR can be carried out to definitively identify the contaminating cells.

For most of the established cell lines, donor tissue is not available and many originators of widely used cell lines are retired or deceased. In these cases, an assumption has to be made, based on the oldest possible cell stocks in repositories. These profiles will need to be labelled as provisional to indicate the absence of authentication back to the original donor tissue.

Until the database described below is available, there are limited resources available for comparing STR profiles. The ATCC and DSMZ cell bank websites and Cell Line Integrated Molecular Authentication (CLIMA; see Further information for links) database[37] provide some information, and at least two series of STR profiles have been published[21,34]. Currently, one of the most useful resources is the list of misidentified cell lines collected by Amanda Capes-Davis and Ian Freshney (supplementary table in REF. 38), which can also be seen on the European Collection of Cell Cultures (ECACC; see Further information for a link) website. All scientists should check the names of the cell lines they are using against this list.

### The interactive database

It is proposed that a database will be established to exploit available STR data to validate the identity of human cell lines. The interactive database will be accessible to everyone, but only the database administrators can make changes or additions. The database will provide DNA profiles and will allow laboratories to compare the STR profiles of their lines, thereby facilitating the validation of experimental data.

Universal criteria are needed for what constitutes a good database. The standard for cell line authentication will establish an interactive database of validated DNA profiles for each unique cell line and will also put in place requirements for carrying out and interpreting the STR assays. The members of the standard committee in conjunction with the National Center for Biotechnology Information (NCBI) will develop the requirements for the database and the database will be maintained by NCBI. The database will initially contain around 500 validated cell lines frequently used by scientists and banked in major cell repositories. The profile of each cell line will be validated before it is submitted to the database.

The most effective database to compare cell line STR-profiling data would consist of a common set of markers. However, not all data have been collected for the same STR loci or using the same generation of sequencing instruments. The use of different primer sets for the same STR markers is a common practice for the forensic and human identity community, which in the United States uses a core set of 13 STR markers for data input into the Federal Bureau of Investigation-maintained Combined DNA Index System (CODIS). To maintain the integrity of the data entered into CODIS, laboratories must use CODIS-approved STR-typing kits and instrumentation, and follow strict quality assurance standards[39]. Approved CODIS STR kits have undergone extensive validation studies that include concordance studies designed to elucidate STR-typing differences that may be seen with the use of different primer sets. Similar protocols will be needed for STR profiling of cell lines.

### The future

Cell line verification by STR profiling will have a substantial effect on scientific research in terms of increased data credibility and less time, money and effort spent studying misidentified cell lines. Accurate identification of



© 2010 Macmillan Publishers Limited. All rights reserved

cell lines is crucial during the development of cell-based medical products to avoid the risks of exposing human subjects to misidentified cells. Although such misidentification can largely be avoided by adherence to quality-control measures, such as proper labelling and tracking schemes during manufacture of a cell-based product, the availability of a standardized method for unambiguous cell and tissue identification could contribute to safety assurance when used to confirm that a cell product came from the intended donor and was not inadvertently mixed with cells from other donors. This issue is of great importance to personalized medicine and the application of stem cell-based technologies, including induced pluripotent stem cells.

No single method is available that provides all the information needed to authenticate a human cell line. STR profiling represents the optimal candidate at this time. Consequently, the standard is intended to evolve as new information becomes available. The interactive, searchable database openly available to everyone will largely eradicate the use of misidentified cell lines. Funding bodies and journals are encouraged to adopt a policy of zero tolerance and demand proof that all cell lines are as claimed.

*For members of the ATCC Standards Development Organization (SDO) Workgroup ASN-0002 see BOX 1*

*John R. W. Masters is at University College London, 67 Riding House Street, London W1W 7EJ, UK.*
*Correspondence to J.R.W.M.*
*e-mail: j.masters@ucl.ac.uk*

doi:10.1038/nrc2852
Published online 7 May 2010

1. Nardone, R. M. Curbing rampant cross-contamination and misidentification of cell lines. *Biotechniques* **45**, 221–227 (2008).
2. Nelson-Rees, W. A. & Flandermeyer, R. R. Inter- and intraspecies contamination of human breast tumor cell lines HBC and BrCa5 and other cell cultures. *Science* **195**, 1343–1344 (1977).
3. MacLeod, R. A. F. *et al.* Widespread intraspecies cross-contamination of human tumor cell lines arising at source. *Int. J. Cancer* **83**, 555–563 (1999).
4. Rothfels, K. H., Axelrad, A. A., Siminovitch, L., McCulloch, E. A. & Parker, R. C. in *Proc. 3rd Canadian Cancer Conf.* (ed. Begg, R. W.) 189–214 (Academic Press, New York, 1958).
5. Defendi, V., Billimgham, R. E., Silvers, W. K. & Moorhead, P. Immunological and karyological criteria for identification of cell lines. *J. Natl Cancer Inst.* **25**, 359–385 (1960).
6. Brand, K. G. & Syverton, J. T. Results of species-specific hemagglutination tests on "transformed", nontransformed, and primary cell cultures. *J. Natl Cancer Inst.* **28**, 147–157 (1962).
7. Gartler, S. M. Genetic markers as tracers in cell culture. *Natl Cancer Inst. Monogr.* **26**, 167–195 (1967).
8. Gey, G. O., Coffman, W. D. & Kubicek, M. T. Tissue culture studies of the proliferative capacity of cervical carcinoma and normal epithelium. *Cancer Res.* **12**, 264–265 (1952).
9. Nelson-Rees, W. A., Flandermeyer, R. R. & Hawthorne, P. K. Banded marker chromosomes as indicators of intraspecies cellular contamination. *Science* **184**, 1093–1096 (1974).
10. Lucey, B. P., Nelson-Rees, W. A. & Hutchins, G. M. Henrietta Lacks, HeLa cells, and cell culture contamination. *Arch. Pathol. Lab. Med.* **133**, 1463–1467 (2009).
11. Nelson-Rees, W. A. Responsibility for truth in research. *Phil. Trans. R. Soc. Lond. B* **356**, 849–851 (2001).
12. Buehring, G. C., Eby, E. A. & Eby, M. J. Cell line cross-contamination: how aware are mammalian cell culturists of the problem and how to monitor it? *In Vitro Cell Dev. Biol. Anim.* **40**, 211–215 (2004).
13. Chatterjee, R. Cell biology. Cases of mistaken identity. *Science* **315**, 928–931 (2007).
14. Gilbert, D. A. *et al.* Application of DNA fingerprints for cell-line individualization. *Am. J. Hum. Genet.* **47**, 499–514 (1990).
15. Nardone, R. M. Eradication of cross-contaminated cell lines: a call for action. *Cell Biol. Toxicol.* **23**, 367–372 (2007).
16. Nardone, R. M. *et al.* An open letter regarding the misidentification and cross-contamination of cell lines: significance and recommendations for correction. *Japanese Collection of Research Bioresources Cell Line Catalogue* [online] http://cellbank.nibio.go.jp/cellbank/qualitycontrol/OL7-11-07.pdf (2007).
17. National Institutes of Health. Notice Regarding Authentication of Cultured Cell Lines. *NIH Guide for Grants and Contracts* [online] http://grants.nih.gov/grants/guide/notice-files/NOT-OD-08-017.html (2007).
18. Dirks, W. G., Drexler, H. G. & MacLeod, R. A. F. ECV304 (endothelial) is really T24 (bladder carcinoma): cell line cross- contamination at source. *In Vitro Cell Dev. Biol.* **35**, 558–559 (1999).
19. van Bokhoven, A., Varella-Garcia, M., Korch C & Miller, G. J. TSU-Pr1 and JCA-1 cells are derivatives of T24 bladder carcinoma cells and are not of prostatic origin. *Cancer Res.* **61**, 6340–6344 (2001).
20. Liscovitch, M. & Ravid, D. A case study in misidentification of cancer cell lines: MCF-7/AdrR cells (re-designated NCI/ADR-RES) are derived from OVCAR-8 human ovarian carcinoma cells. *Cancer Lett.* **245**, 350–352 (2006).
21. Lorenzi, P. L. *et al.* DNA fingerprinting of the NCI-60 cell line panel. *Mol. Cancer Ther.* **8**, 713–724 (2009).
22. Boonstra, J. J. *et al.* Verification and unmasking of widely used human esophageal adenocarcinoma cell lines. *J. Natl Cancer Inst.* **102**, 1–4 (2010).
23. Maddox, J. Responsibility for trust in research. *Nature* **289**, 211–212 (1981).
24. Miller, L. J. Identity crisis. *Nature* **457**, 935–936 (2009).
25. Drexler, H. G., Dirks, W. G., Matsuo, Y. & MacLeod, R. A. F. False leukemia-lymphoma cell lines: an update on over 500 cell lines. *Leukemia* **17**, 416–426 (2003).
26. Nims, R. W., Shoemaker, A. P., Bauernschub, M. A., Rec, L. J. & Harbell, J. W. Sensitivity of isoenzyme analysis for the detection of interspecies cell line cross-contamination. *In Vitro Cell Dev. Biol. Anim.* **34**, 35–39 (1998).
27. Drexler, H. G. & MacLeod, R. A. F. Mantle cell lymphoma-derived cell lines: unique research tools. *Leukemia Res.* **30**, 911–913 (2006).
28. Pathak, S., Nemeth, M. & Multani, A. S. Human tumor xenografts in nude mice are not always of human origin: a warning signal. *Cancer* **86**, 898–900 (1999).
29. Nims, R. W. & Herbstritt, C. J. Cell line authentication using isoenzyme analysis: strategies for accurate speciation and case studies for detection of cell line cross-contamination using a commercial kit. *BioPharm Int.* **18**, 76–82 (2005).
30. Debenham, P. G. & Webb, M. B. Cell line characterization by DNA fingerprinting; a review. *Dev. Biol. Stand.* **76**, 39–42 (1992).
31. Moretti, T. R. *et al.* Validation of short tandem repeats (STRs) for forensic usage: performance testing of fluorescent multiplex STR systems and analysis of authentic and simulated forensic samples. *J. Forensic Sci.* **46**, 647–660 (2001).
32. Budowle, B., Shea, B., Niezgoda, S. & Chakraborty, R. CODIS STR loci data from 41 sample populations. *J. Forensic Sci.* **46**, 453–489 (2001).
33. Butler, J. M. *Forensic DNA Typing. Biology, Technology, and Genetics of STR Markers*, 2nd edn (Academic Press, Burlington, USA, 2005).
34. Masters, J. R. *et al.* Short tandem repeat profiling provides an international reference standard for human cell lines. *Proc. Natl Acad. Sci. USA* **98**, 8012–8017 (2001).
35. Parson, W. *et al.* Cancer cell line identification by short tandem repeat profiling: powers and limitations. *FASEB J.* **19**, 434–436 (2005).
36. Schweppe, R. E. *et al.* DNA profiling analysis of 40 human thyroid cancer cell lines reveals cross-contamination resulting in cell line redundancy and misidentification. *J. Clin. Endocrin. Metab.* **93**, 4331–4341 (2008).
37. Romano, P. *et al.* Cell Line Data Base: structure and recent improvements towards molecular authentication of human cell lines. *Nucleic Acids Res.* **37**, D925–D932 (2009).
38. Capes-Davis, A. *et al.* Check your cultures! A list of cross-contaminated or misidentified cell lines. *Int. J. Cancer* 8 Feb 2010 [epub ahead of print].
39. Federal Bureau of Investigation. Standards and Guidelines Quality Assurance Standards for Forensic DNA Testing Laboratories. *Federal Bureau of Investigation Forensic Science Communications* [online] http://www.fbi.gov/hq/lab/fsc/backissu/oct2008/standards/2008_10_standards01.htm (2008).

**Competing interests statement**
The authors declare competing financial interests: see Web version for details.

DATABASES
**CLIMA database:** http://bioinformatics.istge.it/clima
**PubMed:** http://www.ncbi.nlm.nih.gov/pubmed

FURTHER INFORMATION
**ATCC SDO homepage:** http://www.atccsdo.org
**ATCC cell bank:** http://www.atcc.org
**DSMZ cell culture collection:** http://www.dsmz.de/human_and_animal_cell_lines
**ECACC cell culture collection:** http://www.hpacultures.org.uk/collections/ecacc.jsp

SUPPLEMENTARY INFORMATION
**See online article:** S1 (table) | S2 (box) | S3 (box) | S4 (table)
**ALL LINKS ARE ACTIVE IN THE ONLINE PDF**



© 2010 Macmillan Publishers Limited. All rights reserved

# ONLINE ONLY

**Author biographies**

The members of the American Type Culture Collection Standards Development Organization Workgroup ASN-0002 are Christine Alston-Roberts, Rita Barallon, Steven R. Bauer, John Butler, Amanda Capes-Davis, Wilhelm G. Dirks, Eugene Elmore, Manohar Furtado, Liz Kerrigan, Margaret C. Kline, Arihiro Kohara, Georgyi V. Los, Roderick A. F. MacLeod, John R. W. Masters, Mark Nardone, Roland M. Nardone, Raymond W. Nims, Paul J. Price, Yvonne A. Reid, Jaiprakash Shewale, Anton F. Steuer, Douglas R. Storts, Gregory Sykes, Zenobia Taraporewala and Jim Thomson. This workgroup is currently developing a standard for human cell line authentication and the standard development process is outlined in Box 2 of the article.

**TOC:**

000 **Cell line misidentification: the beginning of the end**

*American Type Culture Collection Standards Development Organization Workgroup ASN-0002: Christine Alston-Roberts, Rita Barallon, John Butler, Amanda Capes-Davis, Wilhelm G. Dirks, Eugene Elmore, Manohar Furtado, Liz Kerrigan, Margaret C. Kline, Arihiro Kohara, Georgyi V. Los, Roderick A. F. MacLeod, John R. W. Masters, Mark Nardone, Roland M. Nardone, Raymond W. Nims, Paul J. Price, Yvonne A. Reid, Jaiprakash Shewale, Anton F. Steuer, Douglas R. Storts, Gregory Sykes and Jim Thomson*

That a substantial proportion of cell lines is mislabelled or replaced by cells derived from a different individual, tissue or species has been a long known, but largely ignored problem. The history of cell line misidentification and recent efforts to develop a standard for the authentication of human cell lines using short tandem repeat profiling is discussed in this article.

**CFI statement**

Rita Barallon (LGC), Manohar Furtado (Applied Biosystems), Jaiprakash Shewale (Life Technologies) Douglas Storts (Promega Corporation) and Jim Thomson (LGC) work for organizations that provide reagents for or undertake cell line authentication using STR profiling. Manohar Furtado is a shareholder in Applied Biosystems.

**Box 3**

Please code all text as normal. Do not code any of the references in this text. Thank you!

© 2010 Macmillan Publishers Limited. All rights reserved

In Vitro Cell.Dev.Biol.—Animal (2010) 46:727–732
DOI 10.1007/s11626-010-9333-z

INVITED REVIEW

# Recommendation of short tandem repeat profiling for authenticating human cell lines, stem cells, and tissues

**Rita Barallon · Steven R. Bauer · John Butler · Amanda Capes-Davis · Wilhelm G. Dirks · Eugene Elmore · Manohar Furtado · Margaret C. Kline · Arihiro Kohara · Georgyi V. Los · Roderick A. F. MacLeod · John R. W. Masters · Mark Nardone · Roland M. Nardone · Raymond W. Nims · Paul J. Price · Yvonne A. Reid · Jaiprakash Shewale · Gregory Sykes · Anton F. Steuer · Douglas R. Storts · Jim Thomson · Zenobia Taraporewala · Christine Alston-Roberts · Liz Kerrigan**

Received: 30 March 2010 / Accepted: 16 June 2010 / Published online: 8 July 2010 / Editor: J. Denry Sato
© The Author(s) 2010. This article is published with open access at Springerlink.com

**Abstract** Cell misidentification and cross-contamination have plagued biomedical research for as long as cells have been employed as research tools. Examples of misidentified cell lines continue to surface to this day. Efforts to eradicate the problem by raising awareness of the issue and by asking scientists voluntarily to take appropriate actions have not been successful. Unambiguous cell authentication is an essential step in the scientific process and should be an inherent consideration during peer review of papers submitted for publication or during review of grants submitted for funding. In order to facilitate proper identity testing, accurate, reliable, inexpensive, and standardized methods for authentication of cells and cell lines must be made available. To this end, an international team of

R. Barallon · J. Thomson
LGC,
Queens Road,
Teddington, Middlesex TW11 0LY, UK

S. R. Bauer
Division of Cellular and Gene Therapies, Office of Cellular, Tissue and Gene Therapies, FDA/Center for Biologics Evaluation & Research, NIH Bldg 29B 2NN10 HFM-740,
8800 Rockville Pike,
Bethesda, MD 20892, USA

J. Butler · M. C. Kline
Biochemical Science Division (831),
Advanced Chemical Science Laboratory (227),
National Institutes of Standards and Technology, Room B226,
100 Bureau Drive, Stop 8312,
Gaithersburg, MD 20899-8312, USA

A. Capes-Davis
CellBank Australia, Children's Medical Research Institute,
Westmead, New South Wales, Australia

W. G. Dirks · R. A. F. MacLeod
DSMZ–German Collection of Microorganisms and Cell Cultures,
Inhoffenstr. 7b,
38124 Braunschweig, Germany

E. Elmore (✉)
Department of Radiation Oncology, University of California,
Medical Sciences I,
B146D,
Irvine, CA 92697, USA
e-mail: eelmore@uci.edu

M. Furtado · J. Shewale
Applied Markets/Genetic Systems Life Technologies,
850 Lincoln Centre Drive,
Foster City, CA 94404, USA

A. Kohara
Department Biomedical Resources, Laboratory of Cell Cultures,
National Institute of Biomedical Innovation,
7-6-8 Saito-Asagi,
Ibaraki, Osaka 567-0085, Japan

G. V. Los
Neuroscience Training Program,
University of Wisconsin-Madison,
1300 University Av.,
Madison WI53706, USA

J. R. W. Masters
Institute of Urology, University College London,
67 Riding House Street,
London W1W 7EJ, UK

scientists is, at this time, preparing a consensus standard on the authentication of human cells using short tandem repeat (STR) profiling. This standard, which will be submitted for review and approval as an American National Standard by the American National Standards Institute, will provide investigators guidance on the use of STR profiling for authenticating human cell lines. Such guidance will include methodological detail on the preparation of the DNA sample, the appropriate numbers and types of loci to be evaluated, and the interpretation and quality control of the results. Associated with the standard itself will be the establishment and maintenance of a public STR profile database under the auspices of the National Center for Biotechnology Information. The consensus standard is anticipated to be adopted by granting agencies and scientific journals as appropriate methodology for authenticating human cell lines, stem cells, and tissues.

**Keywords** Cell authentication · STR profiling · Consensus standard · Quality control

## Introduction

Animal and human primary cell cultures, continuous (immortalized) cell lines, and tissues are of overwhelming importance to the biopharmaceutical industry and to biomedical research as reagents, therapeutic modalities, and as proxy materials for the study of more complex physiological systems. Cell cultures have, from the beginning, been at risk for misidentification due to labeling errors, incorrect classification by pathologists, and cross-contamination with other cell types. Continuous cell lines are potentially jeopardized due to the extended time these are in culture and the frequent manipulations involved in the course of feeding and subculturing. Human stem cell preparations which are propagated in the presence of non-human feeder cell layers are at risk of cross-contamination with the feeder cells. Tumor cells propagated by xenografting onto host animals are at risk of cross-contamination with the host cells.

We know the risks involved in establishing and maintaining cell cultures. We know that periodic identity testing (authentication) is the only way to prove that the cell we are studying is the cell that we believe it to be, and not a contaminating tumor cell line such as HeLa. Why then are many investigators blindly assuming that they are using correctly identified cells? Recent publications appear to indicate that the problem of cell misidentification is not going away. For instance, Berglind et al. (2008) evaluated the p53 status of 1,211 cell lines published between 1989 and 2007 and found discrepancies in the p53 status for 23% of the cell lines. Schweppe et al. (2008) evaluated 40 human thyroid cancer cell lines and found that only 23 of these actually had unique genetic profiles, as determined using short tandem repeat (STR) profiling and single nucleotide polymorphism analysis. Certain of the presumed thyroid cancer cell lines were found to have profiles matching colon cancer or melanoma cells. Another recent revelation was that of Boonstra et al. (2010) indicating that three widely used esophageal cancer cell lines are, in fact, derived from other tumor types. Dittmar et al. (2010) have reported two new cases of misidentification of supposed human cells. Their work clearly demonstrates that phenotypic evaluation alone cannot provide adequate assurance of the authenticity of a cell line. More extensive lists of misidentified cells are available from a number of sources (e.g., ATCC SDO Workgroup ASN-0002 2010; Capes-Davis et al. 2010).

Within the highly regulated biopharmaceutical industry, cell lines used as production substrates must be characterized for identity through phenotypic analysis and confirmation of

M. Nardone
Bio-Trac Program, The Foundation for the Advanced Education in the Sciences at the National Institutes of Health,
Bethesda, MD 20892, USA

R. M. Nardone
Catholic University of America, Cell and Molecular Biology,
620 Michigan Ave NE,
Washington, DC 20064, USA

R. W. Nims
RMC Pharmaceutical Solutions, Inc.,
2150 Miller Drive, Suite A,
Longmont, CO 80501, USA

P. J. Price
D-Finitive Cell Technology, Room B-33,
1023 Wappoo Rd,
Charleston, SC 29407, USA

A. F. Steuer
BioReliance,
14920 Broschart Road,
Rockville, MD 20850, USA

D. R. Storts
Nucleic Acid Technologies, Promega Corporation,
2800 Woods Hollow Road,
Madison, WI 53711, USA

Z. Taraporewala
Division of Cellular and Gene Therapies, Office of Cellular, Tissue, and Gene Therapies,
FDA/Center for Biologics Evaluation and Research,
1401 Rockville Pike, Room 200N,
Rockville, MD 20852, USA

Y. A. Reid · G. Sykes · C. Alston-Roberts · L. Kerrigan
American Type Culture Collection,
10801 University Blvd.,
Manassas, VA 20110, USA

animal species of origin (US FDA 1993). This, together with implementation of current good manufacturing practices, is believed to have contributed to the relatively low frequency of cell line misidentification reported in this industry (Nims and Herbstritt 2005).

Remediation of the problem of cell line misidentification within the biomedical research community may eventually need to be driven by requirements for authentication from granting agencies and journal editors. An international group of scientists is now preparing a consensus standard which will provide investigators with guidance on the appropriate methodology for authenticating human cells. In this article, we describe the rationale for and the process involved in preparing this standard.

## Efforts to Remediate the Problem of Cell Misidentification

The earliest efforts toward tackling the problem of cell misidentification centered on disclosure of the issue through conference presentations and publications. Gartler (e.g., Gartler 1967) and Nelson-Rees (e.g., Nelson-Rees et al. 1974) were among the first and most vocal of those attempting to convince the scientific community of the seriousness of the issue. They hoped that such disclosures would motivate scientists to voluntarily take actions to remediate the problem.

More recently, Roland Nardone championed a series of efforts intended to reemphasize the seriousness of the cell misidentification problem and take any required steps to begin remediating the various causes for the continuing issue. His efforts began with the authoring of a white paper, entitled "Eradication of Cross-Contaminated Cell Lines: A Call for Action" (Nardone 2007). This paper presented recommendations for strict compliance measures in addition to continuing efforts to educate scientists. Nardone believed the time had come for granting agencies to demand cell line authentication as a condition for the receipt of funds and for journals to add a similar requirement to their instructions for authors for manuscripts submitted for publication.

As part of his efforts to convince granting agencies of the need for their participation in his overall remediation strategy, Nardone and a group of prominent cell scientists composed and signed an open letter to Michael O. Leavitt, Secretary of Health and Human Resources (Nardone et al. 2007), beseeching the NIH to take appropriate actions. On November 28, 2007, the NIH published an addition to their Guidelines for Research—Notice Regarding Authentication of Cultured Cell Lines (National Institutes of Health 2007) calling for diligence and more careful peer review.

Communications between Nardone (and others) and journal editors have achieved the desired result as slowly and surely, journals are beginning to add the requirement for cell authentication to their instructions for authors (e.g., *Cell Biochemistry and Biophysics*, *In Vitro Cellular & Developmental Biology—Animal*, *International Journal of Cancer*, and the journals of the American Association for Cancer Research).

Attempts to educate scientists in general of the need for cell authentication must go beyond simply raising the level of awareness of the problem. In his white paper, Nardone also stressed the need for training in cell authentication to be added to conference agendas. He recommended that societies sponsor conferences, workshops, and/or training activities to facilitate the adoption of cell line authentication standards (Nardone 2007).

As the requirement for cell authentication is adopted by granting agencies and scientific journals, the need for standardized methods and expectations regarding authentication itself to be defined becomes more critical. Recognizing this, an effort to prepare a consensus standard on authentication of human cells was initiated.

## The Concept of the Consensus Standard

The idea of the consensus standard is to allow a greater input from the overall international biomedical community into standards. The derivation of a standard through the consensus process improves the chance of universal voluntary acceptance. In turn, that acceptance will foster reproducibility and comparability of research employing human cells. Such a consensus-driven standard, if universally adopted, should ultimately lead to a marked decrease in the misidentification of human cells used by the biomedical community.

## The ATCC® Standards Development Organization

The mission of the ATCC® Standards Development Organization (SDO) is to develop best practices (standards) for use in the life science industry and to promote their global use, using a consensus-driven process that balances the viewpoints of the stakeholder community. Membership is free and open to all stakeholders in the biomedical community, including those involved in the development, production, application, and regulation of life science products. Stakeholders include, but are not limited to, members from academia, government, regulatory, and industry. All members are participants in the consensus review, comment, and voting process.

In 2007, the SDO became the first biological resource organization to become an American National Standards Institute (ANSI)-accredited standards development organi-

zation. ANSI accreditation ensures that procedures used by standards developers meet requirements for openness, balance, consensus, and due process.

The standard development process employed by the ATCC® SDO is shown in Fig. 1.

## ATCC SDO Workgroup ASN-0002

ATCC® SDO workgroup ASN-0002 "Development of a consensus standard for the authentication of human cells: standardization of STR profiling" was formally assembled in early 2009 as a result of a proposal submitted in 2008 by John Masters and Roland Nardone. The workgroup constitutes an international group of concerned and experienced scientists. Chaired by Masters, the workgroup includes individuals with relevant and current experience in DNA profiling technologies, as well as "stakeholders" or representatives from major cell repositories, industry, academia, and government agencies.

*Preparation of the Standard.* Working under the auspices of the ATCC SDO, the ASN-0002 workgroup has met monthly since early 2009. The overall effort was divided between two subgroups which have met independently at monthly or more frequent intervals.

The first subgroup is charged with drafting the introduction to the Standard, defining what is meant by "human cell line authentication," describing the historical aspects, from early discovery of cell line misidentification through to the present efforts encouraging remediation of the problem. The subgroup also is delineating the causes of cell line misidentification, surveying the existing technologies for cell line authentication, and providing the rationale for selection of STR profiling for the Standard. The subgroup is chaired by Raymond Nims.

The second subgroup, chaired by Yvonne Reid, is fleshing out the procedural details of the general protocol to be recommended for STR profiling. This subgroup is also responsible for determining the format and structure of an associated public database of STR profiles of

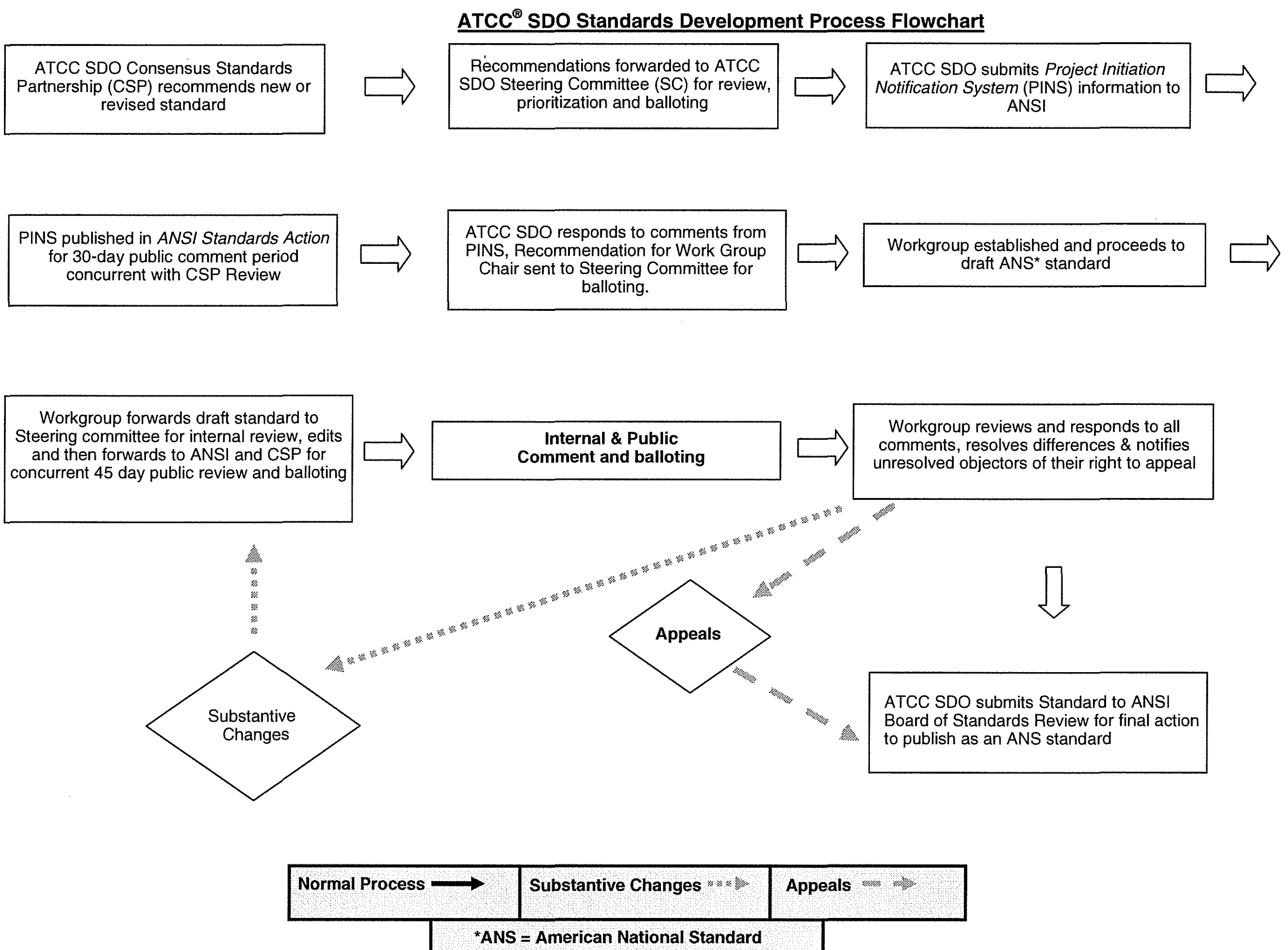



**Figure 1.** ATCC SDO standards development process flowchart.

human cell lines. Such a database is indispensable with regard to the establishment of a set of global reference STR profiles for human cell lines and critical in support of the Standard.

*Methodology.* STR profiling was selected as the recommended authentication technology for inclusion in this Standard primarily because it is capable of resolving human cells to the individual level. In contrast, historically important authentication technologies such as karyotyping, isoenzyme analysis, immunotyping, and human leukocyte antigen typing do not have sufficient resolving (discriminating) power to enable unambiguous authentication of human cells to the individual level (see ATCC SDO Workgroup ASN-0002 2010 for a more detailed discussion of the relative discriminating power of these technologies). In addition, the STR profiling technology is commercially available in kit form and is rapid and economical. Masters et al. (2001) demonstrated that the technology can provide a universal reference standard for human cell lines. The STR profiling technology, as normally used, detects only human cells, and therefore additional methods may need to be used to detect contamination with non-human cells. There are a number of different commercial kits now available for STR profiling, and the users will be encouraged to follow the protocol specific to the kit being used and to refer to the Standard for additional methodological information. The most important aspects of the Standard will be the discussions on the numbers and types of loci to be evaluated, quality control of the data, interpretation of the results (matching criteria, loss of alleles, etc.), and implementation of a universal STR database.

*Associated database.* Associated with the issuance of the Standard will be the construction of a comprehensive and continuously updatable public database of STR profiles based on results subject to agreed-upon interpretation guidelines and quality control parameters. Comparison of STR profiles generated from individual cell stocks to such a database will help reduce the frequency of misidentification of human cells, enhance confidence in results, assure the user's ability to compare scientific results between laboratories, and verify that important data originated from intended samples. STR profiles submitted to the database may, at the request of the user, be verified by staff at the National Institutes of Standards and Technology (NIST). The user also will have the option of submitting STR profiles to the database without verification. The database will indicate which profiles have been verified by NIST. The STR database will be established and maintained under the auspices of the National Center for Biotechnology Information and NIST.

*Timeline for completion of the Standard.* The Standard, once drafted, will be submitted to the SDO Steering Committee for initial review (Fig. 1). After a nominal 14-d review period, the ASN-0002 workgroup will have a chance to respond to any comments provided by the Steering Committee. At this point, the Standard will be submitted for public review and comment. ANSI notifies the public via its weekly publication "ANSI Standards Action." Concurrent with this, the Standard will be sent to all ATCC SDO members for review and comment. At the end of the 45-d public review and comment period, the workgroup will review and respond to all negative comments, resolve differences, and notify unresolved objectors of their right to appeal, if necessary.

The final document will be submitted for review and approval as an American National Standard by the ANSI.

Once the consensus standard has been approved and published by ANSI, the workgroup will take appropriate actions to raise awareness throughout the biomedical community of the existence of the new standard.

## Anticipated Flow and Impact

For newly developed human primary cell cultures and cell lines, including feeder layer-dependent human stem cells and tumor cells propagated through xenografting, an initial STR profile should be established on the earliest material, for example biopsy, frozen tissue, or formalin-fixed paraffin-embedded tissue. For feeder layer-dependent human stem cell preparations, DNA amplification and barcoding (e.g., Cooper et al. 2007), an isoenzyme analysis assay or an alternative species identification method may need to be performed to demonstrate that there are no cross-contaminating mouse feeder cells in the preparation. Additional testing may also be necessary in the case of tumor cell isolation from xenografts to demonstrate that there are no host cells remaining in the recovered tumor cell line.

For existing human cell lines, investigators will be encouraged to: (1) check the public database to see if the cell line is represented within the STR database; (2) perform an STR profile and compare the results to those within the STR database; and (3) ensure that the STR database indicates that this cell line is not misidentified. The Standard will provide the necessary matching criteria.

Continuous human cell lines which are manipulated within laboratories employing non-human cell lines may need to be monitored periodically for non-human cell cross-contamination using one of the cell species identification methods mentioned above (isoenzyme analysis or DNA amplification and barcoding).