

# Check your cultures! A list of cross-contaminated or misidentified cell lines

Amanda Capes-Davis<sup>1</sup>, George Theodosopoulos<sup>1</sup>, Isobel Atkin<sup>2</sup>, Hans G. Drexler<sup>3</sup>, Arihiro Kohara<sup>4</sup>, Roderick A.F. MacLeod<sup>3</sup>, John R. Masters<sup>5</sup>, Yukio Nakamura<sup>6</sup>, Yvonne A. Reid<sup>7</sup>, Roger R. Reddel<sup>1</sup> and R. Ian Freshney<sup>8</sup>

<sup>1</sup> CellBank Australia – Children’s Medical Research Institute, Westmead, NSW, Australia

<sup>2</sup> European Collection of Cell Cultures (ECACC) – Health Protection Agency, Porton Down, Salisbury, Wiltshire, United Kingdom

<sup>3</sup> DSMZ – German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany

<sup>4</sup> JCRB – Japanese Collection of Research Bioresources, Osaka, Japan

<sup>5</sup> Institute of Urology, University College London, London, United Kingdom

<sup>6</sup> RIKEN – BioResource Center Cell Engineering Division, Tsukuba, Japan

<sup>7</sup> ATCC – American Type Culture Collections, Manassas, VA

<sup>8</sup> Centre for Oncology and Applied Pharmacology, Glasgow University, Glasgow, United Kingdom

Continuous cell lines consist of cultured cells derived from a specific donor and tissue of origin that have acquired the ability to proliferate indefinitely. These cell lines are well-recognized models for the study of health and disease, particularly for cancer. However, there are cautions to be aware of when using continuous cell lines, including the possibility of contamination, in which a foreign cell line or microorganism is introduced without the handler’s knowledge. Cross-contamination, in which the contaminant is another cell line, was first recognized in the 1950s but, disturbingly, remains a serious issue today. Many cell lines become cross-contaminated early, so that subsequent experimental work has been performed only on the contaminant, masquerading under a different name. What can be done in response—how can a researcher know if their own cell lines are cross-contaminated? Two practical responses are suggested here. First, it is important to check the literature, looking for previous work on cross-contamination. Some reports may be difficult to find and to make these more accessible, we have compiled a list of known cross-contaminated cell lines. The list currently contains 360 cell lines, drawn from 68 references. Most contaminants arise within the same species, with HeLa still the most frequently encountered (29%, 106/360) among human cell lines, but interspecies contaminants account for a small but substantial minority of cases (9%, 33/360). Second, even if there are no previous publications on cross-contamination for that cell line, it is essential to check the sample itself by performing authentication testing.

**Key words:** authentication, cell culture, cell lines, cross-contamination, DNA profiling, misidentification

Additional Supporting Information may be found in the online version of this article.

**Novelty and Impact:** This manuscript reviews the literature relating to cross-contamination of cell lines. Its novelty comes from the inclusion of a list of known cross-contaminated cell lines (over 300 lines named), allowing researchers to check their own cell lines with reference to the article. Recent developments in this field, including methods of authentication testing, are also discussed.

**Grant sponsor:** National Health and Medical Research Council of Australia

**DOI:** 10.1002/ijc.25242

**History:** Received 24 Nov 2009; Accepted 18 Jan 2010; Online 8 Feb 2010

**Correspondence to:** Amanda Capes-Davis, CellBank Australia, Children’s Medical Research Institute, Locked Bag 23, Wentworthville, NSW 2145, Australia, Fax: +61 2 9687 2120, E-mail: acapdav@gmail.com

## Cell Lines as Model Systems

Continuous cell lines represent a readily accessible and easily studied resource for research into health and disease. These cell lines have acquired the ability to proliferate indefinitely if grown in the appropriate culture conditions; usually this is a rare event, since the majority of cells even in tumor tissue will cease proliferation after a limited number of cell divisions.<sup>1</sup> However, once established, a continuous cell line can be repeatedly passaged, reliably recovers from cryopreservation and retains many of the properties of its cell type or tissue of origin.<sup>2,3</sup> These advantages make continuous cell lines effective, and widely used, model systems for normal cellular processes and for a variety of disease states.

Cell lines are particularly attractive models for studying malignant disease. The genetic changes in tumor-derived cell lines closely resemble those of the tumors of origin.<sup>4</sup> Moreover, the genetic changes required to establish continuous cell lines from normal cells recapitulate many of the genetic changes occurring in cancer.<sup>5,6</sup> These genetic changes are required to overcome replicative senescence, in which normal cells continue to be metabolically active but are restricted from further division.<sup>1</sup> Cells able to overcome senescence continue

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  $\sim 1$  in  $10^7$  cells), an immortalized cell will emerge from crisis and begin to divide again, yielding a continuous cell line.<sup>1</sup> The changes seen throughout this process have many parallels within cancer development, both for malignancy in general and when considering specific tumor types.<sup>7,8</sup>

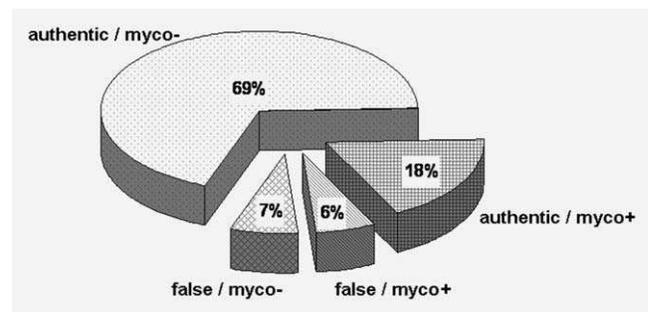
Despite these advantages, numerous cautions have emerged from the literature regarding appropriate use of cell lines as model systems.<sup>9,10</sup> Even where cultures have been transformed through the introduction of specific genes, cell lines that have passed through replicative senescence and crisis are aneuploid, heteroploid and genotypically and phenotypically unstable, resulting in considerable heterogeneity within the culture.<sup>10</sup> 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.<sup>11</sup> 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.<sup>12</sup>

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.<sup>13</sup> Unfortunately, cell lines generally become cross-contaminated early, while still within the originating laboratory.<sup>14</sup> 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.<sup>9,15-17</sup> 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 de-



**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.)

posit at the cell bank.<sup>14</sup> 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.

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.<sup>18</sup> 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.<sup>19</sup> 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.<sup>20</sup>

A practical example of the consequences of cell line contamination can be found in a recent study published by Berglind *et al.*<sup>21</sup> 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,<sup>21</sup> 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,<sup>22,23</sup> in publishing test results online<sup>24</sup> and in developing new detection methods.<sup>25–27</sup> Unfortunately, however, cell banks have also reported reluctance from many researchers to deposit cell lines for distribution.<sup>28</sup> 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.<sup>29,30</sup> 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.<sup>31</sup> 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.<sup>26</sup>

### 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,<sup>20</sup> 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](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.<sup>22,32</sup> 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.<sup>22</sup>

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

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

Type of contaminant	Number of cell lines affected
<b>Intraspecies</b>	
Human	324
Nonhuman	0
<b>Interspecies</b>	
Correct name—incorrect cell type (misidentified) <sup>1</sup>	3
Total	360
<b>Contaminating cell line—12 most frequent</b>	
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) <sup>2</sup>	7
M14 (human melanoma)	7
HL-60 (human acute myeloid leukemia)	6
PC3 (human prostate carcinoma)	6
SW-480, SW620 (human colon carcinoma) <sup>3</sup>	6

<sup>1</sup>For additional misidentified cell lines see Drexler *et al.*<sup>22</sup> <sup>2</sup>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). <sup>3</sup>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<sup>31</sup>; 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	<a href="http://www.atcc.org/">http://www.atcc.org/</a>
CellBank Australia	<a href="http://www.cellbankaustralia.com/">http://www.cellbankaustralia.com/</a>
DSMZ	<a href="http://www.dsmz.de/">http://www.dsmz.de/</a>
ECACC	<a href="http://www.hpacultures.org.uk/collections/ecacc.jsp">http://www.hpacultures.org.uk/collections/ecacc.jsp</a>
JCRB	<a href="http://cellbank.nibio.go.jp/">http://cellbank.nibio.go.jp/</a>
RIKEN Bioresource Center Cell Bank	<a href="http://www.brc.riken.go.jp/lab/cell/english/guide.shtml">http://www.brc.riken.go.jp/lab/cell/english/guide.shtml</a>

### 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<sup>33</sup> and as an assurance for researchers, funding bodies and journals that the cell line used is a valid experimental model.<sup>17</sup>

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.<sup>19,34</sup> 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.<sup>35</sup> 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<sup>35</sup> 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.<sup>26</sup> 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.<sup>25</sup> 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 <i>c</i> 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

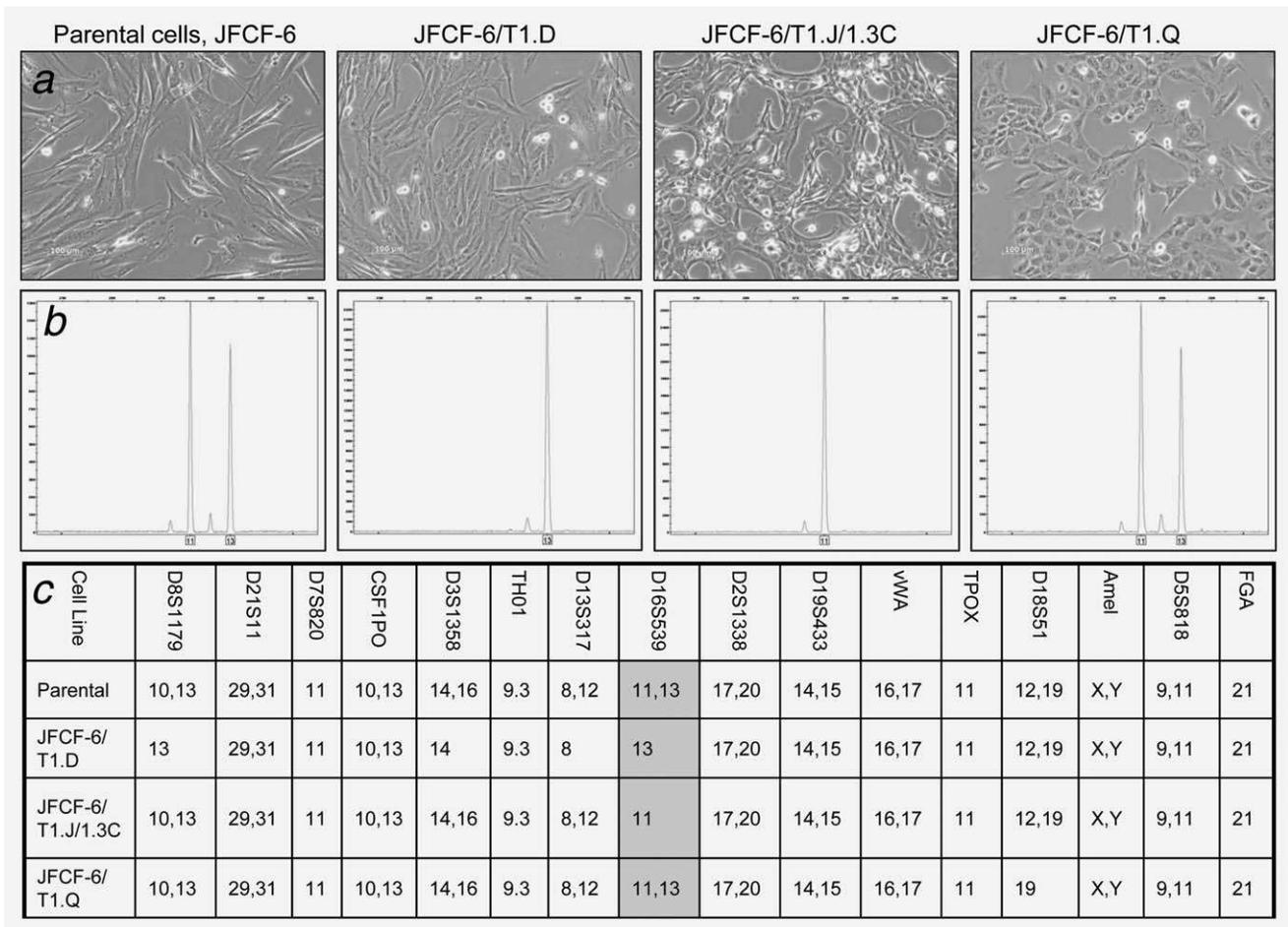
banks to publish such fingerprints online. It is advisable to always compare the test sample to a known sample within the same experiment, ideally using DNA from the blood or tissue of the original donor.

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<sup>26</sup> and accepted within the legal system for human identity testing.<sup>39</sup> 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.<sup>45,46</sup> This is even more evident in tumor-derived cell lines, where evolution or genetic drift continues to occur with passage.<sup>47</sup> 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.<sup>26,44</sup> 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,



**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.<sup>43</sup> 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 AmpFISTR Identifier 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  $\mu$ m). 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.<sup>48,49</sup> SNP analysis can also be used to identify individual samples<sup>50</sup> and has been used for cell line authentication,<sup>51</sup> 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.<sup>52</sup>

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.*<sup>53</sup> Sequence divergences within this fragment allow species discrimination across almost all animal phyla.<sup>42</sup> Although debate is ongoing as to whether DNA barcoding is sufficient for assignment of species in taxonomic terms,<sup>54</sup> 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.<sup>55</sup> DNA barcoding has been tested for species identification of cell lines<sup>27</sup> 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 practi-

cally 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.

## References

- Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell* 2000;100:57–70.
- 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.
- 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.
- 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.
- Reddel RR. The role of senescence and immortalization in carcinogenesis. *Carcinogenesis* 2000;21:477–84.
- Boehm JS, Hahn WC. Immortalized cells as experimental models to study cancer. *Cytotechnology* 2004;45:47–59.
- Masters JR. Human cancer cell lines: fact and fantasy. *Nat Rev Mol Cell Biol* 2000;1:233–6.
- 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.
- 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.
- 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.
- 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.
- Shimada Y. Researchers should have respect for the originator of the cell lines. *Clin Cancer Res* 2005;11:4634.
- Drexler HG, Uphoff CC, Dirks WG, MacLeod RA. Mix-ups and mycoplasma: the enemies within. *Leuk Res* 2002;26:329–33.
- 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.
- Stacey GN. Cell contamination leads to inaccurate data: we must take action now. *Nature* 2000;403:356.
- Masters JR. False cell lines: the problem and a solution. *Cytotechnology* 2002;39:69–74.
- Nardone RM. Eradication of cross-contaminated cell lines: a call for action. *Cell Biol Toxicol* 2007;23:367–72.
- 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.
- Nelson-Rees WA, Flandermeier RR. HeLa cultures defined. *Science* 1976;191:96–8.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- Cooper JK, Sykes G, King S, Cottrill K, Ivanova NV, Hanner R, Ikononi P. Species identification in cell culture: a two-pronged molecular approach. *In Vitro Cell Dev Biol Anim* 2007;43:344–51.
- MacLeod RA, Drexler HG. Public repositories: users reluctant to give materials. *Nature* 2006;439:912.

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, Pugazhenth 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, Beroukhir 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.