The costs of using unauthenticated, over-passaged cell lines: how much more data do we need?

Peyton Hughes¹, Damian Marshall², Yvonne Reid¹, Helen Parkes², and Cohava Gelber¹

Increasing data demonstrate that cellular cross-contamination, misidentified cell lines, and the use of cultures at high-passage levels contribute to the generation of erroneous and misleading results as well as wasted research funds. Contamination of cell lines by other lines has been recognized and documented back to the 1950s. Based on submissions to major cell repositories in the last decade, it is estimated that between 18% and 36% of cell lines may be contaminated or misidentified. More recently, problems surrounding practices of over-subculturing cells are being identified. As a result of selective pressures and genetic drift, cell lines, when kept in culture too long, exhibit reduced or altered key functions and often no longer represent reliable models of their original source material. A review of papers showing significant experimental variances between low- and high-passage cell culture numbers, as well as contaminated lines, makes a strong case for using verified, tested cell lines at low- or defined passage numbers. In the absence of cell culture guidelines, mandates from the National Institutes of Health (NIH) and other funding agencies or journal requirements, it becomes the responsibility of the scientific community to perform due diligence to ensure the integrity of cell cultures used in research.

Effects of Over-Subculturing

Scientific contributions from investigators using continuous cell lines as research tools are significant. However, the ability of continuous cell lines to exist almost indefinitely opens the possibility that cell lines are used beyond safe passage numbers (the point at which the cell culture no longer maintains key gene functions and consistent morphology). Divergent effects of long-term culturing on cell line morphology, development, and gene expression have been documented on key cell lines (1–6). Long-term subculturing places selective pressure on cell line traits leading to, for example, faster growing cells that eventually overrun slower proliferators in the population. In addition, cell lines maintained in culture over a long period of time may experience mutations that alter the original functional characteristics of the cell lines identified at earlier passage levels (6). When cell lines are obtained from colleagues, they often lack verification or documentation about the condition or passage number of the lines. This practice increases the likelihood that inferior, malperforming cultures are used, leading to results that may not be accurate or reproducible (7). Mossberg found that of 1402 surveyed scientists working in industrial settings, 69% reported obtaining one or more cell lines from another researcher or another source, and nearly half of the respondents reported never testing for identity with the lines they use (8).

Passage Number

The age of a tree can be determined by counting the rings in a cross-section of the trunk. A similarly straightforward method for determining the passage number of adherent cell lines does not exist. Stocks of adherent cell lines maintained in laboratories may differ by hundreds of passages. The impact of a cell line’s age (or the number of times it has been passaged) on any given cell line is complex and dependent on several factors including the tissue and species of origin, the culture conditions, and the application for which the cells are used (4,9–12). Furthermore, cell lines do not behave similarly with increased passage number. For example, high-passage Caco-2 cells show an increase in the expression of the green fluorescent protein (GFP) reporter gene after transfection, while high-passage MCF7 cells exhibit a decrease in GFP expression (P. Ikonomi, manuscript in preparation). While observed effects appear at different degrees of subculturing, the potential consequences of using over-subcultured cell lines remain.

The human Caco-2 cell line has been the focus of several studies regarding the influence of passage number on several cell line-specific characteristics, including transport and toxicity of endogenous and exogenous compounds (1,13–15). The Caco-2 cell line is an established model of intestinal absorptive epithelium due to the ability of the cells to form a tight monolayer and to express key enterocyte markers and drug transport mechanisms upon differentiation. Reports demonstrate cell passage level can lead to variability in the key properties that define the ability of Caco-2 cells to predict drug absorption in vivo (Table 1).

In 1996, Sun Lu et al. demonstrated passage-related differences in cell proliferation and transepithelial electrical resistance (TEER) linked to the tightness of the cell monolayer between early- (passage 35–47) and late-passage cells (passage 87–112). This indicates that as passage levels increase, a positive selection of faster-growing subpopulations of cells present in the heterogeneous parental line occurs, forming a tighter monolayer (13). This positive selection of subpopulations of cells was also shown by Briske-Anderson et al. in 1997 (1). In this report, the authors examined Caco-2 cells from passage 20 through passage 109 and found the TEER values increase up to passage 36 then decline after passage 60. They also found the proliferation rates of the cells and the activity of alkaline phosphatase increased in the later passage cells. Also in 1997, H. Yu et al. showed low- (passage 28–36) and high- (passage 93–108) passage

¹ATCC, Manassas, VA, USA and ²LGC Limited, Teddington, Middlesex, UK

doi 10.2144/000112598
Caco-2 cells differ in their compound transport characteristics. Cells at high-passage levels showed reduced carrier-mediated transport, reduced paracellular permeability, and increased transcellular permeability consistent with a reduction in the functional expression of a brush border enzyme and several transport proteins (6).

More recently, in 2005, Sambuy et al. examined the effect of cell-related factors on parental and clonally derived Caco-2 cell lines. This data supported the proposition that passage number influences brush border enzyme activities, morphology, TEER, proliferation rate, cell density, glucose transporter expression, carrier-mediated transport activities, and metabolic activities. It also highlighted the importance of culture conditions, such as seeding density, and medium composition, in influencing Caco-2 variability (5).

To investigate the impact these sources of variability have on the reproducibility of data from the Caco-2 cell model, LGC and ATCC, in a collaborative study with the National Institute for Biological Standards and Controls (NIBSC), investigated the effect of passage number, cell source, and medium composition on the transepithelial permeability, TEER, and proliferation rates of Caco-2 cells. This study observed a decrease in the TEER properties of cells passed over 50 times, an increase in the paracellular permeability of the monolayer (consistent with the decreasing TEER values), a decrease in the transcellular permeability of the cells (consistent with a reduction in p-GP), and an increase in the proliferation rate of the cells as the passage number increased. These increases occurred in a number of different cell sources [ATCC, German Collection of Microorganisms and Cell Cultures (DSMZ), and European Collection of Cell Cultures (ECACC)] over the same passage range and were heavily influenced by the type of medium used to grow the cells (www.mfprog.org.uk/publications/publications_item.asp?intPublicationID=1365).

The effect of extended passaging on cell culture characteristics is not limited to the Caco-2 cell model. In response to an increasing interest in agonists able to modulate growth and differentiation of prostate tumor cells, Esquenet et al. reported results of LNCaP prostatic adenocarcinoma cells derived from low-(passage 24–32) and high- (approximately passage 80) passage cells (3). This cell line, originally derived from a lymph node metastasis in 1977, has become an established model for the study of prostate cancer progression due to its proliferative and secretory responses to androgens. Esquenet et al. reported the high-passage cells showed divergent results compared with low-passage controls. The high-passage cells displayed an increased proliferative response when exposed to increasing concentrations of androgens, as well as a loss of the characteristic inhibition of growth in the presence of the synthetic androgen R1881. The high-passage cells also had a lower secretion rate of the differentiation marker prostate-specific antigen (3). These changes in cell line characteristics, due to increasing passage, are most likely attributable to genetic heterogeneity within the parental line, which leads to a positive selection process during prolonged cell culture (16,17).

In further studies into the effect of passage number, Wenger et al. reported results of a study comparing two sources (cell repositories and other laboratories) for each of the established cancer cell lines, MCF7 (breast cancer) and Ishikawa (endometrial tumor), which had all been subcultured for over 100 passages (7). All four cell line samples were examined by karyotype and comparative genomic hybridization (CGH). The authors found that the MCF7 cell lines showed slight karyotype differences (also observed by CGH analysis), suggesting genotype changes occurring over time due to cell passage level and maintenance in different laboratories. The Ishikawa cell lines showed no karyotype similarities other than a missing X chromosome with CGH analysis, suggesting that one or both of the lines became contaminated or that extensive passaging resulted in genetic drift. This study illustrates the changes a cell line may undergo due to multiple passaging and also emphasizes the importance of confirming the identity of cultured cells used for research, especially when obtained from a source other than a cell repository (7).

Cell line authentication is equally important with human embryonic stem (hES) cells, which are purported to have a limitless proliferation potential. A number of reports have been produced recently that highlight potential genetic frailty in hES cells, which can be exposed during the passaging procedure. Draper et al. reported the duplication of the long arm

Table 1. Summary of Reports Detailing the Effects of Cell Passage Level on Caco-2 Cells

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Cell Response After Passaging</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>TEER</td>
<td>Increase in TEER from passage 29 to passage 198.</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Increase in TEER from passage 35–47 to passage 87–112.</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>Decrease in TEER from passage 36 to 86.</td>
<td><a href="http://www.mfprog.org.uk/topics/themes.asp">www.mfprog.org.uk/topics/themes.asp</a></td>
</tr>
<tr>
<td>Cell density</td>
<td>Increase in cell density from passage 29 to passage 198.</td>
<td>6</td>
</tr>
<tr>
<td>Proliferation rates</td>
<td>Growth rates lower in passage 35–47 to passage 87–112.</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>Growth rates lower in passage 36 to passage 86.</td>
<td><a href="http://www.mfprog.org.uk/topics/themes.asp">www.mfprog.org.uk/topics/themes.asp</a></td>
</tr>
<tr>
<td></td>
<td>Growth rates higher in passage 20 compared with passage 72.</td>
<td></td>
</tr>
<tr>
<td>Paracellular permeability</td>
<td>Decrease in paracellular permeability from passage 29 to passage 198.</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Increase in paracellular permeability from passage 36 to passage 86.</td>
<td><a href="http://www.mfprog.org.uk/topics/themes.asp">www.mfprog.org.uk/topics/themes.asp</a></td>
</tr>
<tr>
<td></td>
<td>No significant difference in permeability from passage 35–47 to passage 87–112.</td>
<td>13</td>
</tr>
<tr>
<td>Transcellular permeability</td>
<td>Increase in transcellular permeability from passage 29 to passage 198.</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>No significant difference in permeability from passage 35–47 to passage 87–112.</td>
<td>13</td>
</tr>
<tr>
<td>Carrier-mediated transport</td>
<td>Decrease in transcellular permeability from passage 29 to passage 198.</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>No significant difference in permeability from passage 35–47 to passage 87–112.</td>
<td>13</td>
</tr>
<tr>
<td>Alkaline phosphatase</td>
<td>Higher expression of AP in passage 29 compared with passage 198.</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Higher expression of AP in passage 20 compared with passage 72.</td>
<td>14</td>
</tr>
</tbody>
</table>

TEER, transepithelial electrical resistance.
of chromosome 17 translocated to 6q in a subpopulation of the hES cell line H7 following 60 passages in culture (18). A similar result was reported by Inzunza et al. (19), who observed a normal chromosomal content in an hES cell line (HS237) at passage 35, but chromosomal aberrations in the X chromosome when the cell line was reanalyzed at passage 61. Mitalipova et al. (20) showed techniques used for cell passaging can also have an effect on the genetic stability of stem cells. They demonstrated that bulk passage methods after as few as 23 passages compared to manual passage methods, leads to aneuploidy detectable by karyotyping (20). Bulk passage methods include use of either enzymatic disaggregation (collagenase/trypsin) or nonenzyme-based methods using cell association buffer. Maitra et al. reported that following long-term culturing, a number of late-passage hES cell lines from different sources had at least one genetic abnormality commonly observed in human cancer cells, including DNA copy number and promoter methylation (21).

Passage number has also been shown to influence pluripotency of ES cells. Li et al. using the technique called tetraploid embryo complementation (where a mouse ES cell is aggregated with a four-cell-stage tetraploid embryo, then implanted, and allowed to develop in the uterus of a pseudopregnant female mouse), demonstrated that increased ES cell passage number negatively affected the ability of the cells to form an adult mouse (22). This technique is considered the most reliable method for determining ES cell pluripotency and demonstrates the important effects passaging can have on cells considered to have limitless proliferative potential.

Cross-Contamination and Misidentification

The problem of intraspecies and interspecies cross-contamination among cell lines has been recognized for half a century, and although reviews have been published, evidence of continued use of misidentification and cellular cross-contamination of cell cultures has not declined (23–36). Masters et al. found as much as 20% of all ostensibly cell lines in use today are not as they are purported (37), and other results support these findings (23). An additional report estimates that more than one-third of cell cultures are cross-contaminated either with cells from other species (interspecies contamination) or with unrelated cells from the same species (intraspecies contamination) (32). MacLeod et al. (24) tested 252 human cell lines deposited at the DSMZ and found similar levels of cross-contamination: namely, 18% or 45 lines. The authors reported, “These misidentified cell lines have already been used in several hundreds of potentially misleading reports, including use as inappropriate tumor models and subclones masquerading as independent replicates” (24). A sample of known past cross-contamination reports is depicted in Table 2.

One of the earliest reports suggesting a high frequency of HeLa cell cross-contamination was made by Stanley Gartler in 1966 at a cell culture conference. Gartler later published in Nature, where it was shown that 18 lines thought to be of Caucasian tumor origins were found to contain isoenzyme A of G6PD, the same isoenzyme variant present in the aggressive HeLa cell line, which was derived from African American Henrietta Lacks (38).

Between this early report and today, the problem has intensified (36). In “Human Cell Cross-Contamination Since 1983,” Masters reported results from a Medline search for the years 2000–2004 revealing 19 citations for the putative intestinal cell Int 407, 45 citations for the putative amnion cell WISH, 59 citations for the putative liver cell Chang liver, 470 citations for the putative human nasal carcinoma cell Hep-2, and 556 citations for the putative oral carcinoma KB (37). Masters reported hundreds of papers being published each year in high-impact journals using these cell lines, yet fewer than 10% of the papers reveal the model system being used is in fact HeLa cells (37).

Over 220 publications were found by Buehring et al. (23) in a PubMed database search from 1969 to 2004, in which known HeLa contaminants were used as a model for the tissue type of the original cell line (23). Furthermore, the authors collected survey data from mammalian cell biologists to determine how many were using HeLa contaminants without being aware of their true identity and how many were not using available means to ensure correct identity. The survey respondents included scientists, staff, and graduate students in 48 countries. The survey revealed of the 483 respondents, 32% used HeLa cells, 19% unwittingly were using HeLa contaminants, nearly half (46%) never tested for cell identity, 35% obtained all cell lines from another laboratory, and 63% obtained at least one cell line from another laboratory rather than from a major repository.

In 2005, Melcher et al. reported the putative normal colon epithelial cell line NCOL-1 (commonly used for colon cancer research) was not representative of a normal colon epithelial cell line. The researchers used spectral karyotyping to show this cell line is identical to LoVo, a cell line derived from a colon carcinoma (39). Yoshino et al. used short tandem repeat (STR) polymorphism analysis to examine approximately 400 cell lines in the Cell Engineering Division of the Japanese research institution RIKEN and found that 10 of the human cell lines were identical to a different cell line that had been earlier deposited in the collection and that had been misidentified by the depositor (40). Liscovitch et al. reported the misidentification of a human breast adenocarcinoma cell line, MCF7/AdR, which has been used in over 300 molecular studies of cancer cell drug resistance. This cell line, originally thought to be from the parental MCF7 cell line, was recently matched to the OVCAR-8 ovarian adenocarcinoma (41), suggesting many studies will need further verification of results against the OVCAR-8 parental cell line. Azari et al. used a technique called combined DNA index system (CODIS), which uses tetrameric STR sequences on 13 distinct chromosomes to create a fingerprint that has a random match probability of one in a trillion (42). Using this technique, the authors examined 100 cell lines deposited in the National Cell Bank of Iran and compared them to STR reference profiles obtained from ATCC and the Japanese Collection of Research Bioresources cell bank. They found that 18.8% of the cell lines examined had been cross-contaminated (42).

Discussion

The number of publications containing spurious data as a result of over-passaged, misidentified, or contaminated cell lines is unknown. Drexler et al. summarizes, “the problem of false cell lines operates as a classic positive feedback loop, whereby information about false cell lines goes unreported or buried in specialist journals escaping the attention of unwary beginners who may later become victims or even perpetrators” (29).

The existing data and reports of wasted time and funds underscore the need for establishing proper controls and standards for cell culture conditions. Scientists have choices and can adopt safeguards. Researchers can perform authentication tests or, when possible, acquire cell cultures from reputable sources, such as cell banks or culture collections. Cell lines have already been used in several hundreds of potentially misleading reports.
<table>
<thead>
<tr>
<th>Year</th>
<th>Cell Cultures Observed</th>
<th>Results</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1968</td>
<td>Eighteen independent human cell lines.</td>
<td>All 18 lines were contaminated with HeLa.</td>
<td>39</td>
</tr>
<tr>
<td>1973</td>
<td>Two cell lines claimed to be of human origin.</td>
<td>Both shown to be of mouse origin.</td>
<td>48</td>
</tr>
<tr>
<td>1974</td>
<td>Twenty independent cell lines.</td>
<td>Nine of the 20 cell cultures displayed HeLa markers.</td>
<td>49</td>
</tr>
<tr>
<td>1974</td>
<td>Two purported human carcinoma cancer lines (HTB-3 and HTB-39B) and 1 HEK cell line.</td>
<td>All three lines were HeLa cells.</td>
<td>27</td>
</tr>
<tr>
<td>1976</td>
<td>246 lines investigated specifically for evidence of cellular cross-contamination or mislabeling.</td>
<td>Overall 30% of the lines were incorrectly designated (14% for interspecies and 25% for intraspecies contamination).</td>
<td>50</td>
</tr>
<tr>
<td>1977</td>
<td>279 cell lines from 45 laboratories.</td>
<td>In total, 41 lines not as purported; 21 were of the wrong species, 15 were HeLa instead of other human cell lines, one rodent cell line was contaminated with manta cells, one purported normal diploid human cell line was actually a BT-20 breast carcinoma cell line, and three purported mixed species lines lacked one of two species (one lacked rat in an avian-rat mixture; two lacked mouse in a human-mouse mixture).</td>
<td>51</td>
</tr>
<tr>
<td>1979</td>
<td>Two established human breast carcinoma cell lines of metastatic origin.</td>
<td>Both shown to be cross-contaminated.</td>
<td>52</td>
</tr>
<tr>
<td>1981</td>
<td>Four Hodgkin’s disease human cell lines (FQ, RB, SpR, and Ry).</td>
<td>None of the cell lines were Hodgkin’s. Three of the cell lines were shown to be identical. The origin of the fourth (RY) could not be established with certainty.</td>
<td>53</td>
</tr>
<tr>
<td>1981</td>
<td>From 103 sources, lines derived from endometrium, amniotic cells, breast carcinoma, and other gynecologic cancers, larynx, and lung cancer, gastrointestinal cancer, liver and bone-marrow-derived leukemia and lymphoma lines, urologic cancer, etc.</td>
<td>About 100 documented events of contamination.</td>
<td>33</td>
</tr>
<tr>
<td>1984</td>
<td>257 cultures.</td>
<td>Overall 35% were contaminated; 36% of the human lines were cross-contaminated (25% by cells of another species and 11% by another human cell line).</td>
<td>54</td>
</tr>
<tr>
<td>1988</td>
<td>Insulin-producing cell line, clone 16, thought to be derived from human fetal endocrine pancreatic cell.</td>
<td>Documented to be of Syrian hamster origin.</td>
<td>55</td>
</tr>
<tr>
<td>1993</td>
<td>Sister cell lines SPI-801 and SPI-802 thought to be established from a patient with acute lymphoblastic leukemia (ALL).</td>
<td>Shown to be subclones of K-562, a chronic myeloid leukemia cell line.</td>
<td>56</td>
</tr>
<tr>
<td>1994</td>
<td>Macrophage-monocyte U-937.</td>
<td>Found to be contaminated with K-562.</td>
<td>57</td>
</tr>
<tr>
<td>1999</td>
<td>189 cell cultures received by the German Collection of Microorganisms and Cell Cultures (DSMZ) cell bank (1990–1999) representing 170 human hematopoietic lines.</td>
<td>Seventeen out of 117 (14.5%) from original source and 11 out of 72 (15.3%) from a secondary source were shown to be cross-contaminated with another hematopoietic cell line.</td>
<td>28</td>
</tr>
<tr>
<td>1999</td>
<td>252 human cell lines submitted to the German DSMZ cell bank.</td>
<td>45 of the cell lines were contaminated, mostly by intraspecies contamination.</td>
<td>24</td>
</tr>
<tr>
<td>2000</td>
<td>Cell line, ECV 304, a putative human endothelial line of umbilical vein origin.</td>
<td>Shown to be identical to cell line, T24/83 derived from human urinary bladder carcinoma.</td>
<td>58</td>
</tr>
<tr>
<td>2002</td>
<td>TI-1 cell line reportedly established from peripheral blood blasts from male patient.</td>
<td>Shown to be a cross-contaminant of K-562, a line derived from bone marrow of a 53-year-old female patient with chronic myelogenous leukemia (CML).</td>
<td>46</td>
</tr>
<tr>
<td>2003</td>
<td>550 human leukemia-lymphoma cell lines.</td>
<td>Unequivocal evidence of misidentification for 82 (14.9%) of the lines was found.</td>
<td>29</td>
</tr>
<tr>
<td>2005</td>
<td>Normal colon epithelial cell line NCOL-1.</td>
<td>Identified as the colon carcinoma cell line LoVo.</td>
<td>38</td>
</tr>
<tr>
<td>2006</td>
<td>Approximately 400 human cell lines deposited in the RIKEN BioResource Center.</td>
<td>Ten human lines were identical to a different cell line in the collection.</td>
<td>40</td>
</tr>
<tr>
<td>2007</td>
<td>Human breast adenocarcinoma cell line MCF-7 and the derived line MCF7/AdR.</td>
<td>Both were found to have originated from the ovarian adenocarcinoma cell line OVCAR-8.</td>
<td>41</td>
</tr>
<tr>
<td>2007</td>
<td>100 cell lines deposited at the National Cell Bank of Iran.</td>
<td>Eighteen of the lines were cross-contaminated.</td>
<td>42</td>
</tr>
</tbody>
</table>
For those scientists working on cell lines derived themselves or received from a colleague, basic authentication tests such as STR profiling, isoenzyme analysis, and contamination tests are readily available and should be routinely used (11, 12, 28, 32, 46, 47). Transferring cell lines to colleagues should be avoided, or when it does occur, accompanied with comprehensive documentation verifying the integrity of the material or tests need to be repeated (36).

The basis for any research, development, or production program involving cell cultures is the selection of an identity-verified and low-passage cell line. The use of similar and identified passage numbers throughout a project will better ensure reproducible results and comparisons between laboratories. To further ensure the use of authenticated cell lines, full cell line documentation, including the source and passage numbers used during experiments, should be submitted for scientific publications (28, 31, 34, 36). Cell lines are critical components of experiments and should be considered as standard research reagents and given the same care and quality control measures that surround the use of kits, enzymes, and other laboratory products commercially obtained.

Until a greater consciousness and consensus regarding authentication compliance within the scientific community is achieved, malperforming and/or contaminated cell lines will continue to be released into the research community and spurious scientific conclusions will continue to affect the credibility and integrity of all biomedical science.

ACKNOWLEDGMENTS

We thank Bettyanne Eilers, Aina Downing, and Scott Jenkins for their valuable assistance.

COMPETING INTERESTS STATEMENT

The authors declare no competing interests.

REFERENCES

Review

36. Stacey, G.N. 2000. Cell contamination leads to inaccurate data: we must take action now. Nature 403:356.

Received 23 April 2007; accepted 4 September 2007.

Address correspondence to Peyton Hughes, ATCC, 10801 University Blvd., Manassas, VA 20110, USA. e-mail: phughes@atcc.org

To purchase reprints of this article, contact: Reprints@BioTechniques.com

584 | BioTechniques | www.biotechniques.com