Investigation of cross-contamination among human cell lines used in China

Mingzhen Wang[a], Meimei Yang[b], Yuehong Liu[a], Yaqing Huang[a], Fang Ye[a], Congyi Zheng[a,b,*], Chao Shen [a,b,*]

[a] College of Life Sciences, Wuhan University, Wuhan, China

[b] China Center for Type Culture Collection, Wuhan University, Wuhan, China

*Correspondence should be addressed to: College of Life Sciences, Wuhan University, Wuhan, Hubei 430072 P. R. China

Chao Shen

E-mail: shenchao@whu.edu.cn

Tel: 86-27-68752093

Fax: 86-27-68754833

or

Congyi Zheng

E-mail: cctcc202@whu.edu.cn

Tel: 86-27-68754001

Fax: 86-27-68754833

Abstract

Increasing attention has been paid to the hazard of cell line cross-contamination. More journals are requiring that cell lines be authenticated prior to manuscript submission. China Center for Type Culture Collection (CCTCC) has authenticated cell lines for publications in the *International Journal of Cancer* since 2009. Between 2010 and 2016, we authenticated 485 cell lines from 66 Chinese universities and research institutes. Almost 50% of cell lines were misidentified, most were completely replaced by other cell lines (mainly HeLa and its sublines). The number of cell lines we received for authentication increased dramatically after 2013, but the proportion of misidentified cell lines remained elevated (over 50%). We strongly recommend that researchers take authentication of cell lines seriously and routinely perform authentication of the cell lines used in their laboratories.

Key Words: Cell line, Cross-contamination, STR

Novelty and Impact Statements:

China Center for Type Culture Collection has provided cell line authentication for journal publications since 2009. We authenticated 485 cell lines from 66 Chinese universities and research institutes between 2010 and 2016. Almost 50% of cell lines were misidentified. We strongly recommend that researchers take cell line authentication seriously and routinely authenticate their

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as an 'Accepted Article', doi: 10.1002/ijc.30923

own cell lines.

Introduction

Cell lines are commonly used in life sciences and clinical medicine. Cell line contamination can lead to biased or incorrect interpretation of cell-based experiments. Since 2007, increasing attention has been paid to the hazard of cell line cross-contamination[1-3]. However, in China, many researchers have overlooked cell line quality, leading to serious cell line contamination and misidentification[4]. In recent years, many leading journals have added a requirement to authenticate cell lines prior to manuscript submission to avoid the resulting academic controversy[5]. While there are many ways to authenticate cell lines, DNA typing based on short tandem repeats (STRs) is still the gold standard for cell authentication[6-8]. STRs are a class of DNA sequences formed by tandem repeats of a core unit, which is a series of two to seven base pairs. STRs are widely used to detect cross-contamination in human cell lines because they have a wide distribution, are well researched, have a high degree of polymorphism, and are easily PCR amplified and typed.

In 2009, we received a letter from the editorial board of International Journal of Cancer, inviting the China Center for Type Culture Collection (CCTCC) to authenticate cell lines for manuscripts submitted by Chinese researchers, and we subsequently accepted this invitation. From 2010 to 2016, the CCTCC received authentication applications from 66 Chinese universities and research institutes, and authenticated 485 cell lines. The results of these cell line authentications are summarized below.

Materials and methods

Cell lines

All cell lines were sent to the CCTCC solely for authentication prior to manuscript submission. A total of 485 human cell lines were submitted by 66 universities and research institutions in China. Submitted samples included cells from existing and newly established cell lines. Among samples of newly established cell lines, none came from the lab that originally developed the cell line.

STR typing

The total genomic DNA of cells from submitted cell lines was extracted with a genomic DNA

purification kit (Tiangen Biotech, Beijing, China). PCR of the genomic DNA was performed using an STR multiple amplification kit-Microreader TM21ID System (Suzhou Microread Genetics, Beijing, China). The TM21D system contained 20 STR loci (CSF1PO, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, FGA, TH01, TPOX, vWA, Penta D, Penta E, D19S433, D16S1043, D2S441, D12S391 and D2S1338) and amelogenin. The amplified product was added to a 96-well plate and analyzed using capillary electrophoresis with an ABI 3730X1 instrument[9]. The STR profiles were compared to those found in the ATCC and DSMZ STR databases to authenticate the cell lines. Cell lines with a match rate of ≥80% were classified as correct cell lines, i.e., derived from a common donor ancestry. Cell lines with a match rate of <55% were classified as misidentified cell lines. Cell lines with a match rate of approximately 55%–80% required further analysis to determine the degree of relatedness to the common donor ancestry.

Results and discussion

The samples in this study were tested to authenticate cell lines for manuscript submission. All 485 submitted cell lines were divided into three categories according to the authentication results: correct cell lines, misidentified cell lines, and unknown cell lines. The misidentified cell lines were further divided into three categories: nonhuman cell lines, mixed human cell lines, and replaced cell lines. As shown in Figure 1, misidentified cell lines account for almost 50%, of the 485 cell lines authenticated, and most were completely replaced by other cell lines (cell cross-contamination). Cell lines were mainly contaminated with HeLa cells and its sublines. The proportion of mixed human cell lines was small, likely because cells could not coexist long term under the same culture conditions. Thus, one cell line would completely replace the other after several passages.

The results of our cell line authentication from 2010 to 2016 are shown in Table 1. The number of cell lines authenticated in 2013–2016 was more than twice that authenticated in 2010–2012. This result could be due to an increasing number of Chinese researchers prioritizing the quality of their cell lines or an increasing number of journals requiring cell line authentication prior to manuscript submission. While the number of cell lines submitted for authentication has increased significantly over time, the proportion of misidentified cell lines remained elevated. One

reason for misidentification may be because the experiments were mainly carried out by graduate students who might not pay enough attention to the quality of cell lines. Another reason might be the difficulty in obtaining a cell line from overseas cell banks. Thus, labs may obtain cell lines from other labs or small companies and maintain them in their lab or transfer them to other labs. If the original cell lines were misidentified, the users would never know without authentication.

Human cell lines are important materials, not only for basic research, but also as cell models for cancer research and drug development. To ensure the scientific accuracy of the research and avoid unnecessary waste[11], it is absolutely critical for researchers to choose correctly identified and uncontaminated cell lines. Therefore, cell lines should be authenticated frequently, especially before the start of the new research project[12, 13]. In China, raising awareness about the importance of cell line authentication is still urgently needed.

Among the cell lines authenticated by the CCTCC over seven years, less than 50% were correctly identified, and there are no signs of improvement at the moment. Therefore, we strongly recommend that Chinese and global researchers: 1) recognize the grave situation of cell line misidentification, 2) purchase cell lines from professional cell banks, and 3) develop good laboratory practices, including regular authentication of cell lines. These steps are the only way to reduce the risk of using misidentified cell lines in research and to achieve accurate/reliable research outcomes.

References

1. Chatterjee R. Cell biology. Cases of mistaken identity. *Science* 2007;**315**: 928-31.

2. Identity crisis. *Nature* 2009;**457**: 935-6.

3. Lorsch JR, Collins FS, Lippincott-Schwartz J. Cell Biology. Fixing problems with cell lines. *Science* 2014;**346**: 1452-3.

4. Ye F, Chen C, Qin J, Liu J, Zheng C. Genetic profiling reveals an alarming rate of cross-contamination among human cell lines used in China. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* 2015;**29**: 4268-72.

5. American Type Culture Collection Standards Development Organization Workgroup ASN. Cell line misidentification: the beginning of the end. *Nature reviews Cancer* 2010;**10**: 441-8.

6. Barallon R, Bauer SR, Butler J, Capes-Davis A, Dirks WG, Elmore E, Furtado M, Kline MC, Kohara A, Los GV, MacLeod RA, Masters JR, et al. Recommendation of short tandem repeat profiling for authenticating human cell lines, stem cells, and tissues. *In vitro cellular & developmental biology Animal* 2010;**46**: 727-32.

7. Dirks WG, Drexler HG. STR DNA typing of human cell lines: detection of intra- and interspecies

cross-contamination. *Methods in molecular biology* 2013;**946**: 27-38.

8. Masters JR, Thomson JA, Daly-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. *Proceedings of the National Academy of Sciences of the United States of America* 2001;**98**: 8012-7.

9. Huang Y, Liu Y, Zheng C, Shen C. Investigation of Cross-Contamination and Misidentification of 278 Widely Used Tumor Cell Lines. *PloS one* 2017;**12**: e0170384.

10. Nelson-Rees WA, Daniels DW, Flandermeyer RR. Cross-contamination of cells in culture. *Science* 1981;**212**: 446-52.

11. Neimark J. Line of attack. *Science* 2015;**347**: 938-40.

12. Freedman LP, Gibson MC, Ethier SP, Soule HR, Neve RM, Reid YA. Reproducibility: changing the policies and culture of cell line authentication. *Nature methods* 2015;**12**: 493-7.

13. McLaren RS, Reid Y, Storts DR. Human cell line authentication: the critical first step in any project using human cell lines. *Methods in molecular biology* 2013;**963**: 341-53.

Table 1. Analysis of cell lines collected from 2010 to 2016

	2010	2011	2012	2013	2014	2015	2016	Total
A is A	27	12	7	23	57	24	49	199
A is B	15	7	6	9	40	38	50	165
Cross-contaminant cells	4	2	3	5	18	7	13	52
Nonhuman cells	2	7	0	2	11	2	3	27
Unknown cells	5	4	2	1	9	2	19	42
Total	53	32	18	40	135	73	134	485

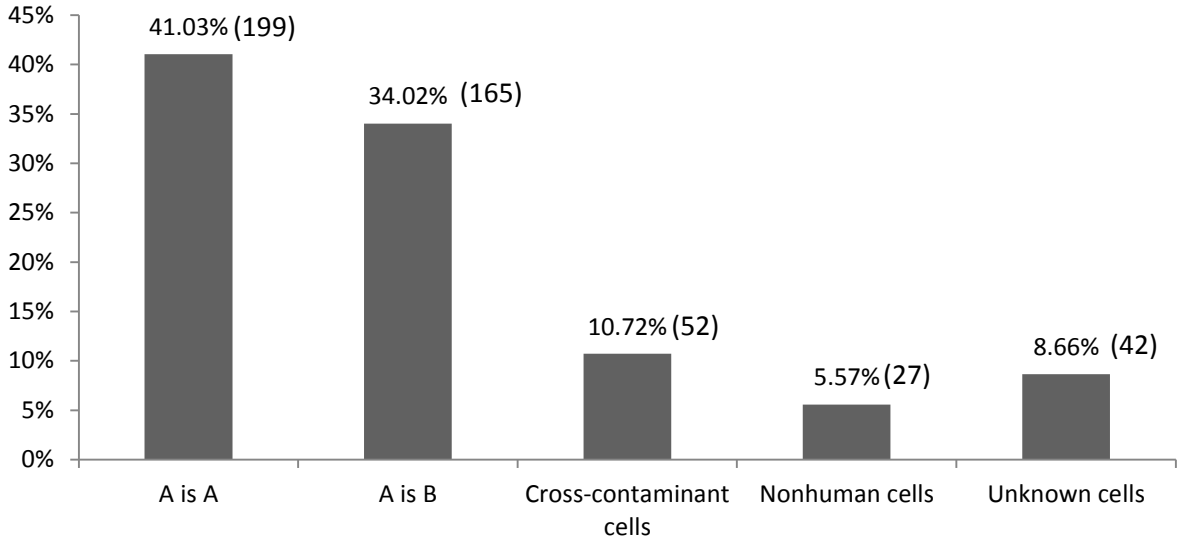

Fig 1. Percentage of total cell lines collected during 2010-2016

Note: "A is A" means the identity of the cell line was correct. "A is B" means the cell line was misidentified as another cell line. For example, samples that correspond to known misidentified cell lines (e.g., KB cells, Chang liver cells[10]) were included in the 'A is B' category. "Cross-contaminant cells" were cell lines that were mixtures of two to three human cell lines, "Nonhuman cells" were nonhuman cell lines. "Unknown cells" were cell lines that did not have matching STR profiles in the surveyed databases, implying they are unknown cell lines.