Gene 763 (2020) 145048

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Short communication

Advantages of a 21-loci short tandem repeat method for detection of cross-contamination in human cell lines

Meijia Gu[a], Jingxuan Liu[c], Meimei Yang[d], Mingmin Zhang[a], Jian Yang[a], Suling Duan[a], Xuxu Ding[a], Jie Liu[e], Chuguang Chen[e], Yundian Zeng[a], Chao Shen[b,d,*]

[a] Key Laboratory of Combinatorial Biosynthesis and Drug Discovery, Ministry of Education, School of Pharmaceutical Sciences, Wuhan University, Wuhan, China
[b] College of Life Sciences, Wuhan University, Wuhan, China
[c] School of Resources and Environmental Engineering, Wuhan University of Technology, Wuhan, China
[d] China Center for Type Culture Collection, Wuhan University, Wuhan, China
[e] Beijing Microread Genetics Co., Ltd., Beijing, China

ARTICLE INFO

Keywords:
Cross-contamination
Cell line authentication
STR
21 loci

ABSTRACT

Cross-contamination of cell lines is a highly relevant and pervasive problem. The analysis of short tandem repeats (STR) is a simple and commercially available technique to authenticate cell lines for more than two decades. At present, STR multiple amplification kits have been developed up to 21 loci while the current STR databases only provide 9-loci STR profiles. Here, we compared the advantages of 21-loci STR methodology using the same algorithm as 9-loci method. The 21-loci method reduced the uncertainty ratio for authentications by 97.5% relative to the 9-loci method and exclude effectively false positive. We show that the additional 12 loci helped to greatly reduce sample-site marker specificity arising from genetic isolation and the occurrence of null alleles, suggesting that inclusion of additional loci in these databases will ultimately improve the efficiency and accuracy of authentication of cell lines. Taken together, we demonstrate the utility of a 21-loci method in human cells, providing a novel marker panel for use as a valuable alternative to 9-loci analyses to minimize cell line authentication errors and reduce costs due to erroneous experiments.

1. Introduction

Cell lines are widely used in biological research and the production of biological reagents and treatments (Popovič et al., 1982). Thousands of cell lines have been employed as *in vitro* models in scientific research and are widely used as models to investigate the molecular mechanisms of disease, especially cancer (Boonstra et al., 2010; Gu et al., 2010; Gu and Shen, 2018). However, caution is needed when using continuously cultured cell lines due to the possibility of contamination, as another cell line or microorganism can easily be introduced during this time period (Garnett et al., 2012; Geraghty et al., 2014). The misidentification and cross-contamination of cell lines has been a widespread problem for more than 60 years and is considered to be one of the most compelling quality-control issues today (Hammond et al., 1994; Drexler et al., 2003; Liscovitch and Ravid, 2007).

Many major journals, including *Nature, Cell, Nature Cell Biology, International Journal of Cancer* and so on, now recommend that cell lines should be verified for authenticity before publication; however, the problem of cross-contamination remains unresolved and continues to negatively influence biomedical research and the development of new drugs. Methods for cell line authentication include karyotyping, immunological techniques, isoenzyme analysis, cellular markers, and DNA profiling (Gartler, 1968; Montes de Oca et al., 1969; O'Brien et al., 1980; Wright et al., 1981; Gilbert et al., 1990; Nims et al., 1998). Among them, short tandem repeat (STR) profiling of DNA microsatellites is considered the gold standard, providing high-fidelity identity matching between a cell line and its original donor (Gilbert et al., 1990). The highly polymorphic nature and lifelong stability of STR sequences in humans, in addition to the availability of low-cost techniques, make STR profiling ideal for authentication. This method allows identification software to assign a number to each allele at that locus, and this type of profiling is used by the American Type Culture

Abbreviations: STR, short tandem repeats; ATCC, the American Type Culture Collection; CCTCC, the China Center for Type Culture Collection; JCRB, the Japanese Collection of Research Bioresources Cell Bank; DSMZ, and the German Collection of Microorganisms and Cell Cultures; IJC, International Journal of Cancer; DNA, deoxyribonucleic acid

* Corresponding author at: College of Life Sciences, Wuhan University, Wuhan, Hubei 430072, China.
E-mail address: shenchao@whu.edu.cn (C. Shen).

https://doi.org/10.1016/j.gene.2020.145048

Available online 14 August 2020

Collection (ATCC), the China Center for Type Culture Collection (CCTCC), the Japanese Collection of Research Bioresources Cell Bank (JCRB), and the German Collection of Microorganisms and Cell Cultures (DSMZ). The Standard Development Organization at the ATCC is currently generating an international standard for human cell line identification based on STR profiling (ATCC SDO Workgroup ASN-0002). Strict criteria and databases for STR profiles derived from cancer cell lines have been developed, enabling comparison of STR sample data with authenticated cell lines. These STR databases use a 9-loci method, including eight core STR loci (D5S818, D13S317, D7S820, D16S539, vWA, Th01, TPOX, and CSF1PO) plus the amelogenin marker, for the identification of human cell lines (Institute, 2011) . Cell line matching is based on an algorithm that compares the number of shared alleles between two cell line samples and expresses this value as a percentage. To date, STR multiple amplification kits have been developed up to 21-loci, with 9-loci, 16-loci and 21-loci kits being the most widely used. Nevertheless, the current STR databases only provide 9-loci STR profiles, making the compared results of extra loci useless.

Due to the genetic volatility of cancer cell lines, including microsatellite instability, loss of heterozygosity, aneuploidy and the increase in genetic instability that happens when cells are cultured continuously. STR-based authentication of these cell lines can be more difficult than that for normal cell lines (Capes-Davis et al., 2010; Yu et al., 2015). It was reported that combining 16-loci STR profiling, more than 20.5% of tumor cell lines were revealed as having been incorrectly identified, including intra-species (14.5%), inter-species (4.4%) cross-contamination and contaminating cell lines (1.7%)(Bian et al., 2017). For example, the cell lines HCCC-9810 and Calu-6 exhibited an 88.9% match in the 9-loci ATCC STR database, suggesting a common origin; however, 21-loci analysis revealed a 48.2% match, indicating differing origins (CCTCC). The results of the 21-loci STR analysis were also confirmed by single nucleotide polymorphism (SNP) analysis (Huang et al., 2017). Moreover, *International Journal of Cancer* (*IJC*) reported recently two THP-1 cell lines from two major repositories, presented numerous genomic, transcriptomic and proteomic discrepancies that have pervasive effects, namely on genes instrumental in leukemogenesis. This indicates that the two THP-1 cell lines are not the same entity and have undergone biologically important genetic drift which can be underestimated by analyses of a 9-loci STR(Noronha et al., 2020).

Thus, we wanted to expand our performance comparison between the 9-loci and 21-loci authentication methodologies to a larger number of cell lines and human subjects. With the growing popularity and accessibility of STR technology, a considerable number of studies involve the use of STR data, and these studies could also be used to check the authenticity of the cell lines. Because cell line authentication databases only provide STR profiles of 9 loci, this study sought to compare the sensitivity and discrimination power of 9-loci STR profiling with the newly developed 21-loci STR profiling using 197 human cell lines (S1) and 299 unrelated human subject volunteers (S2).

2. Materials and methods

2.1. Human cell lines and samples

The human cell line group consisted of 197 cell lines, and the unrelated human subject volunteer group consisted of 299 samples. The 197 human cell lines are obtained from different universities, companies and hospitals in China. Besides, students from local universities cover all the volunteers as unrelated human group.

2.2. STR typing

The methods of STR profiling were described previously (Huang et al., 2017). Briefly, total genomic DNA from human cell lines and subjects was extracted using the TIANamp genomic DNA purification kit (Tiangen Biotech, Beijing, China). PCR of genomic DNA was performed using an STR multiple amplification kit-Microreader TM21ID System (Suzhou Microread Genetics, Beijing, China). Mutiplex PCR reactions were performed on a T-Gradient thermal cycler (Biometra, Gottingen, Germany). The TM21D system uses primers to co-amplify 20 STR loci (including 13 combined core STR loci of DNA index system, namely CSF1PO, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, FGA, TH01, TPOX, vWA, and 7 other loci, namely Penta D, Penta E, D19S433, D16S1043, D2S441, D12S391, D2S1338) and amelogenin. The reaction mixture contained 2 ng genomic DNA, 2.5 mM dNTPs, 10 × PCR buffer, 0.2 μM of each primer, as well as 2.5 U of Taq DNA polymerase. The PCR protocol was as follows: denaturing at 94 °C for 5 min, denaturing at 94 °C for 30 s, annealing at 60 °C for 60 s, extension at 70 °C for 60 s in 30 cycles and final extension at 60 °C for 30 min. One μl aliquot of amplicon was mixed with 7.5 μl HiDiformamide (Applied Biosystems, CA, USA) and 0.5 μl ROX-500 internal-lane size standards. The mixture was denatured at 95 °C for 5 min, then placed immediately on ice for 5 min, followed by adding 1 μl allelic ladder (Suzhou Microread Genetics, Beijing, China) and loading on a POP-7 polymer gel (Applied Biosystems, CA, USA) for electrophoresis. The samples were electrophoresed with the GeneScan program on the ABI 3130 Genetic Analyzer (Applied Biosystem, CA, USA). The amplicon sizes were determined using GeneMapper 3.2 (Applied Biosystems, CA, USA). Alleles were designated by comparison to the allelic ladder. Each sample was repeated at least three times to confirm the results.

2.3. Statistical analyses

All cell line and human subject STR data were uploaded and analyzed by the China Center for Type Culture Collection (CCTCC) Multifunction STR Profile Comparison System. STR results for cell lines were subjected to pairwise comparison. The comparison was performed according to the following formula for a combination: C (n, r) = n!/r! (n-r)!, where n is the number of things that are choosing from, r is the number of items, ! is a factorial of a number. Here, we used 197!/2! (197–2)! = 19306, indicating that 19,306 pairs of results were obtained for the comparison of the 9- and 21-loci methods for each pairwise comparison of the 197 cell lines. Using the same algorithm, 299!/2! (299–2)! = 44551, indicating that 44,551 pairs of results were obtained for the 299 human sample pairwise comparisons.

3. Results

Availability and Implementation: Freely available on the web at http://47.92.66.97:8091/cellcontrast/

3.1. Comparative analysis of 9-loci and 21-loci STR methods

To compare performance between the 9- and 21-loci methods, we obtained 19,306 comparison results from the cell line group and 44,551 comparison results from the human subject group (Fig. 1). According to the algorithm of ATCC, The percent match between the submitted sample and the database profile equals to the number of shared alleles between query sample and database file divided by the total number of alleles in the database file. Linear regressions were determined for the cell line group (Fig. 1A) and the human subject group (Fig. 1B). We observed an apparent consistency and common trend between the two sample groups. For the cell lines, the R^2 was 0.683, while this value was 0.368 for the human subject comparisons.

3.2. The 21-loci STR method provides greater accuracy

For the cell line group results, the 9-loci method yielded 8.65% uncertain authentications, while uncertainty decreased to 0.22% using the 21-loci method (Table 1). The 21-loci method therefore reduced the uncertainty ratio compared to the 9-loci method by 97.46% for cell line

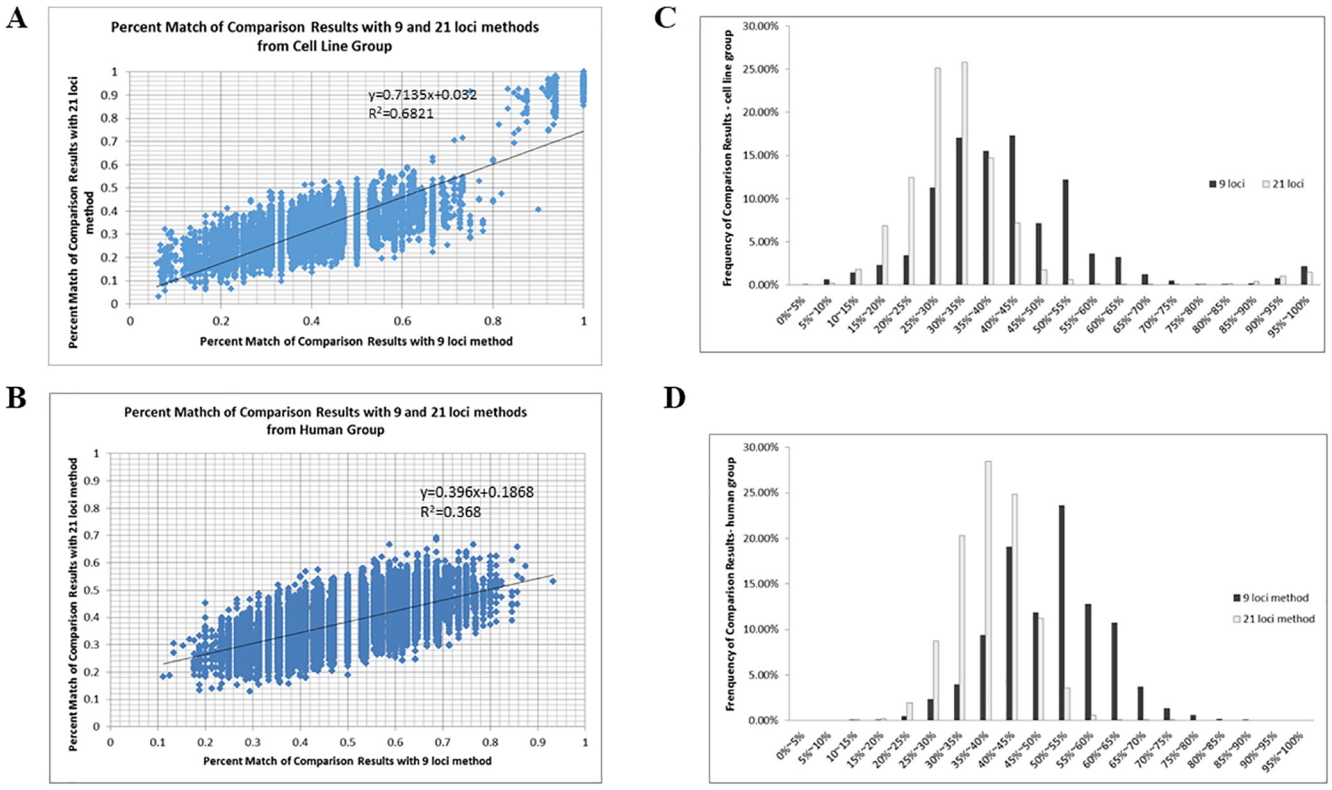

Fig. 1. Comparative analysis of 9-loci and 21-loci STR methods. Comparison results with 9 and 21 loci methods (A) and (B). (A) 19,306 pairs of comparison results from the 9- and 21-loci methods from the cell lines. Linear regression: $y = 0.7135x + 0.032$, $R^2 = 0.6821$. (B) 44,551 pairs of comparison results from the 9- and 21-loci methods from the human subjects group. Linear regression: $y = 0.396x + 0.1868$, $R^2 = 0.368$ (X means percent match of comparison results with 9 loci method, y means percent match of comparison results with 21 loci method, R^2 means relevance between the two different methods. There is an apparent consistency and common trend between the two sample groups). Frequency distribution of comparison results from the 9- and 21-loci methods (C) and (D). (C) Frequency distribution of 19,306 comparison results from the 9- and 21-loci methods from the cell line group. (D) Frequency distribution of 44,551 comparison results from the 9- and 21-loci methods from the human subjects group.

authentication. Excluding the gender loci, the additional 12 loci analyzed in the 21-loci STR method increased the fidelity of the STR analysis compared to the 9-loci analysis.

Results of uncertain identity are always an important variable that affects cell authentications. Thus, the 21-loci method could greatly reduce uncertainty, improving the accuracy of the whole analysis. With respect to the human group, the 9-loci method authenticated 0.18% (80/44551) of all data; however, because the human subject group consisted of 299 individual people without consanguinity, all of the comparison results should have yielded uncertain or unrelated results (< 80% match rate), with a majority being unrelated authentications (< 55% match rate) (Table 1). Thus, these authenticated matches from the 9-loci method are likely erroneous. Testing the same data with the 21-loci method yielded no related authentications. For the human subject group, there were 29.11% uncertain authentications with the 9-loci method and only 0.67% uncertain authentications with the 21-loci method, suggesting that the 21-loci method reduced the uncertainty ratio by 97.70%. Together, these data from both cell lines and human subject samples indicate that the 21-loci method STR analysis reduced the uncertainty ratio by 97.50%, greatly improving accuracy.

3.3. The 21-loci method offers more reliability

As shown in Table 1, for the cell lines group, the 9-loci method resulted in 611 pairs (3.16% frequency) of related comparison results, while the 21-loci method resulted in 598 pairs (3.10% frequency). Thus we made a detailed comparative analysis of these 13 (611–598 = 13) different cell lines, as shown in Table 2. Although the 21-loci method resulted in 13 fewer related pairs with 0.06% difference in frequency, the 21-loci method was able to ensure the conservative of related authentication (Table 2). According to ATCC, the Chang liver cell line was originally thought to be derived from normal liver tissue; however, this cell line was found to be contaminated by HeLa cells based on isoenzyme analysis, HeLa marker chromosomes, and DNA fingerprinting

Table 1
Count and frequency of unrelated, uncertain and related authentications of cell line group and human group with 9 loci and 21 loci method, separately.

Cell line group		9 loci method		21 loci method	
		Count	Frequency	Count	Frequency
	Unrelated authentications (0 ~ 55%)	17,025	88.19%	18,665	96.68%
	Uncertain authentications (55%~80%)	1670	8.65%	43	0.22%
	Related authentications (80%~100%)	611	3.16%	598	3.10%
Human group	Unrelated authentications (0 ~ 55%)	31,504	70.71%	44,251	99.33%
	Uncertain authentications (55%~80%)	12,967	29.11%	300	0.67%
	Related authentications (80%~100%)	80	0.18%	0	0%

Table 2
The comparison result of the 13 different cell lines of related comparison results with 9 loci method and 21 loci method.

Cell A	Cell B	9 loci (%)	21 loci (%)
Chang Liver	HeLa S3-1	93.33	77.33
WSS-1–2	293/CHE-Fc	92.31	75.36
293A	293/CHE-Fc	92.31	74.63
HeLa S3-15	Chang Liver	87.50	78.38
293/CHE-Fc	WSS-1–1	85.71	75.36
293/CHE-Fc	293/CHE-FC	84.62	73.53
293 T	293/CHE-Fc	84.62	69.57
HeLa S3	Chang Liver	81.25	77.33
Hep G2-03	Hep G2-01	80.00	62.34
Hep G2-03	Hep G2-02	80.00	61.54
CAL 27	SiHa-1	81.82	47.62
Caco-2	HCC827	80.00	44.07
Caco-2–2	HCC827	80.00	44.83

(CCTCC). In our results, the 9-loci method ratio was 87.5%, indicating a match, while the 21-loci method ratio was only 78.38%, indicating uncertainty (Table 2). All 9-loci comparison ratios between HeLa cells and Chang liver cells were greater than 80%, suggesting a common donor ancestry; however, the 21-loci method yielded lower comparison ratios. This discrepancy may be caused by the information that the additional 12 loci provided, and such information could help greatly when distinguishing different cell line subtypes from a common origin. Additionally, WSS-1 cells, which were derived from human embryonic kidney, were also verified as HEK293 cells (Table 2).

As indicated in Table 3, all of the last three pairings of cell lines (CAL 27 and SiHa-1, Caco-2 and HCC827, Caco-2–2 and HCC827) originate from tissues different than their paired counterpart. The comparison results of these three false pairs of cell line samples yielded a high degree of similarity using the 9-loci ATCC STR algorithms and database (81.82%, 80%, and 80%, respectively). While these percent matches indicated that all three pairs of cell lines were from the same origin, SiHa-1 cells had been established from cancerous tissues of the cervix, while CAL 27 originated from a middle tongue lesion. The SiHa-1 cell line shared results with AMEL, D7S820, VWA, and TPOX and differed from the CAL 27 cell line at 5 other loci based on comparisons with the ATCC 9-loci STR database. Comparison of these two cell lines using the 21-loci method revealed 4 shared loci, while the remaining 17 loci, D5S818, D18S51, D6S1043, Penta E, D12S391, and FGA were different. When the 9-loci algorithm was replaced with 21-loci, the percent match of these two cell lines was 47.62%, indicating that they were from different origins (Table 3). Obviously, these results further confirm that the additional 12 loci offer greater genetic representation and information for cell lines. While the similarities between cell line pairs Caco-2 and HCC827 and Caco-2–2 and HCC827 were both 80% with the 9-loci method, they were only 44.07% and 44.83%, respectively, using the 21-loci method The results of these three pairs of cell line samples suggest that the 21-loci method offers more reliability.

4. Discussion

CCTCC started using the 9-loci STR profiling method for cell line authentication in 2009 but currently uses the 21-loci method (Huang et al., 2017) in response to awareness of cell line contamination and human variation, resulting in higher heterogeneity of cell lines used in research. Our study expands the previously used 9-loci method of STR profiling with the addition of 12 more markers. As demonstrated in this study, the 9-loci method lacked the necessary accuracy, resulting in false authentications, and we predict even more false authentications as new primary cell lines are established. Our data suggest that the 21-loci method could reduce false authentications, and such improvement is critical for biomedical research that employs cell lines as experimental models. The 21-loci method of STR profiling may therefore be

Table 3
The comparison results of the 3 false pairs of cell line samples with 9 and 21 loci method.

cell line	AMEL*	D5S818*	D13S317*	D7S820*	D16S539*	VWA*	TH01*	TPOX*	CSF1PO*	D19S433	D21S11	D18S51	D6S1043	D3S1358	PENTA_D	D2S441	D8S1179	PENTA_E	D12S391	D2S1338	FGA
CAL 27	X	11,12	10,11	10	11,12	14,17	6,9.3	8	10,12	14,15.2	28,29	13	12	16	9,10	10,11.3	13,15	7	18.3,20	23,24	25
SiHa-1	X	9	11	10	12	14,17	6,9	8	12	14.2	29,31	15	18	16,17	9,12	10	13,16	10,12	19,22	24	21
9 loci (%)	81.82%																				
21 loci (%)	51.72%																				
Caco-2	X	12,13	11,13,14	11,12	12,13	16,18	6	9,11	11	15	30,32	12	12	14,17	9,11	10,15	12,14	7	17,23	17,25	19
HCC827	X	12	9	11,12	12	18	6	8	11	14	31	13	12	17	14	11,11.3	12	20	17	17,24	24
9 loci (%)	80.00																				
21 loci (%)	44.07																				
Caco-2–2	X	12,13	11,13	11,12	12,13	16,18	6	9,11	11	15	30,32	12	12	14,17	9,11	10,15	12,14	7	17,23	17,15	19
HCC827	X	12	9	11,12	12	18	6	8	11	14	31	13	12	17	14	11,11.3	12	20	17	17,24	24
9 loci (%)	80.00																				
21 loci (%)	44.83																				

The difference of STR data between CAL 27 and SiHa-1, Caco-2 and HCC827, Caco-2–2 and HCC827 were showed red letters.

employed to provide a more reliable and accurate result when data from 9-loci analyses is suspicious.

According to ATCC standards, Cell lines with ≥ 80% match are considered to be related; i.e., derived from a common ancestry. Cell lines with between a 55% to 80% match require further profiling for authentication of relatedness, while cell lines with a match rate of < 55% are classified as misidentified cell lines. In both Fig. 1C and 1D, the majority of the data for both methodologies fell under regions deemed unrelated and/or uncertain by ATCC standards. Within these unrelated/uncertain regions, the data showed a relatively high degree of consistency regardless of sample group; however, differences between the 9- and 21-loci analyses increased gradually as the percent match comparison results increased from 20 to 25%. Multiple peaks in the 9-loci method data implied that the changing trend was unstable for different regions of the percent match comparison results. Overall, these findings suggest that the 9- and 21-loci methodologies are somewhat redundant. By the shift of the two peaks, the frequency distribution and comparison of the two different methods could be determined simultaneously. Furthermore, the peak values for the 21-loci method are 25.82% and 28.47%, much higher than the 17.31% and 23.63% of the 9-loci method. The position of the peaks shifted leftward toward a greater match compared to those for the 9-loci method, suggesting that the 21-loci method is more informative, leading to the determination of accuracy and reliability.

While the similarities between cell line pairs Caco-2 and HCC827 and Caco-2–2 and HCC827 were both 80% with the 9-loci method, they were only 44.07% and 44.83%, respectively, using the 21-loci method, leading to different conclusions regarding cell identities with the two methods. The results of these three pairs of cell line samples suggest that the current “gold standard” of 80% matching could be reduced when using the 21-loci method due to its higher reliability.

To conclude, the 21-loci method greatly improves the accuracy of STR profiling as reducing the uncertainty ratio by 97.46% for cell line authentication and 97.70% for the human subject group, compared to the 9-loci method. Moreover, in our results, there is 0.18% false positive determination using 9-loci method, but 21-loci method effectively excluded these false positives. All these confirms that the additional 12 loci represent important gene information that is missing from the 9-loci method, indicating that the 21-loci STR method is much more reliable for cell line and human subject authentication. Additionally, the 21-loci method could improve the reliability of biological products and scientific payoffs, reducing unnecessary waste. We speculate that cell line verification using the 21-loci method will strengthen the reproducibility and comparability of cell lines in different laboratories. Therefore, we strongly encourage all of scientific researchers to perform STR analysis using 21-loci method on their cells to confirm their identity.

Funding

This work was supported by the Fundamental Research Funds for the Central Universities (2042018kf0241) and the National Science and Technology Infrastructure Grants (NSTI-CR15 and NSTI-CR16) to Dr. Chao Shen.

CRediT authorship contribution statement

Meijia Gu: Methodology, Writing - review & editing. **Jingxuan Liu:** Conceptualization, Methodology. **Meimei Yang:** Software. **Mingmin Zhang:** Data curation, Conceptualization. **Jian Yang:** Writing - original draft. **Suling Duan:** Writing - original draft. **Xuxu Ding:** Software, Validation. **Jie Liu:** Software. **Chuguang Chen:** Data curation. **Yundian Zeng:** Software. **Chao Shen:** Visualization, Investigation, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We acknowledge the support of Microread Genetics Co., Ltd, Beijing.

References

Bian, X., Yang, Z., Feng, H., Sun, H., Liu, Y., 2017. A Combination of Species Identification and STR Profiling Identifies Cross-contaminated Cells from 482 Human Tumor Cell Lines. Sci Rep 7 (1). https://doi.org/10.1038/s41598-017-09660-w.

J.J. Boonstra R. van Marion D.G. Beer L. Lin P. Chaves C. Ribeiro A.D. Pereira L. Roque S. J. Darnton N.K. Altorki D.S. Schrump D.S. Klimstra L.H. Tang J.R. Eshleman H. Alvarez Y. Shimada H. van Dekken H.W. Tilanus W.N.M. Dinjens Verification and Unmasking of Widely Used Human Esophageal Adenocarcinoma Cell Lines 102 4 2010 271 274.

Capes-Davis, A., Theodosopoulos, G., Atkin, I., Drexler, H.G., Kohara, A., MacLeod, R.A.F., Masters, J.R., Nakamura, Y., Reid, Y.A., Reddel, R.R., Freshney, R.I., 2010. Check your cultures! A list of cross-contaminated or misidentified cell lines. Int. J. Cancer 127 (1), 1–8.

Drexler, H.G., Dirks, W.G., Matsuo, Y., MacLeod, R.A.F., 2003. False leukemia–lymphoma cell lines: an update on over 500 cell lines. Leukemia 17 (2), 416–426.

Garnett, M.J., Edelman, E.J., Heidorn, S.J., Greenman, C.D., Dastur, A., Lau, K.W., Greninger, P., Thompson, I.R., Luo, X.i., Soares, J., Liu, Q., Iorio, F., Surdez, D., Chen, L.i., Milano, R.J., Bignell, G.R., Tam, A.T., Davies, H., Stevenson, J.A., Barthorpe, S., Lutz, S.R., Kogera, F., Lawrence, K., McLaren-Douglas, A., Mitropoulos, X., Mironenko, T., Thi, H., Richardson, L., Zhou, W., Jewitt, F., Zhang, T., O'Brien, P., Boisvert, J.L., Price, S., Hur, W., Yang, W., Deng, X., Butler, A., Choi, H.G., Chang, J.W., Baselga, J., Stamenkovic, I., Engelman, J.A., Sharma, S.V., Delattre, O., Saez-Rodriguez, J., Gray, N.S., Settleman, J., Futreal, P.A., Haber, D.A., Stratton, M.R., Ramaswamy, S., McDermott, U., Benes, C.H., 2012. Systematic identification of genomic markers of drug sensitivity in cancer cells. Nature 483 (7391), 570–575.

GARTLER, STANLEY.M., 1968. Apparent HeLa Cell Contamination of Human Heteroploid Cell Lines. Nature 217 (5130), 750–751.

Geraghty, R.J., Capes-Davis, A., Davis, J.M., Downward, J., Freshney, R.I., Knezevic, I., Lovell-Badge, R., Masters, J.R.W., Meredith, J., Stacey, G.N., Thraves, P., Vias, M., 2014. Guidelines for the use of cell lines in biomedical research. Br J Cancer 111 (6), 1021–1046.

Gilbert, D.A., Reid, Y.A., Gail, M.H., Pee, D., White, C., Hay, R.J., O'Brien, S.J., 1990. Application of DNA fingerprints for cell-line individualization. Am J Hum Genet 47, 499–514.

Gu, M., Li, H., Shen, C., Wu, L., Liu, W., Miao, L., Zheng, C., 2010. Cloning and characterization of a new BRCA1 variant: A role for BRCT domains in apoptosis. Cancer Letters 295 (2), 205–213.

Gu, M., Shen, C., 2018. Novel cancer cell lines derived from primary breast tumors in Chinese patients. Am J Transl Res 10, 3956–3968.

Hammond, H.A., Jin, L., Zhong, Y., Caskey, C.T., Chakraborty, R., 1994. Evaluation of 13 short tandem repeat loci for use in personal identification applications. Am J Hum Genet 55, 175–189.

Y. Huang Y. Liu C. Zheng C. Shen P. Talamas-Rohana Investigation of Cross-Contamination and Misidentification of 278 Widely Used Tumor Cell Lines PLoS ONE 12 1 e0170384 10.1371/journal.pone.0170384 10.1371/journal.pone.0170384.g001 10.1371/journal.pone.0170384.g002 10.1371/journal.pone.0170384.g003 10.1371/journal.pone.0170384.t001 10.1371/journal.pone.0170384.t002 10.1371/journal.pone.0170384.t003 10.1371/journal.pone.0170384.t004 10.1371/journal.pone.0170384.t005 10.1371/journal.pone.0170384.t006 10.1371/journal.pone.0170384.s001 10.1371/journal.pone.0170384.s002 10.1371/journal.pone.0170384.s003 10.1371/journal.pone.0170384.s004 10.1371/journal.pone.0170384.s005 https://dx.plos.org/10.1371/journal.pone.0170384 http://dx.plos.org/10.1371/journal.pone.0170384.g001 http://dx.plos.org/10.1371/journal.pone.0170384.g002 http://dx.plos.org/10.1371/journal.pone.0170384.g003 http://dx.plos.org/10.1371/journal.pone.0170384.t001 http://dx.plos.org/10.1371/journal.pone.0170384.t002 http://dx.plos.org/10.1371/journal.pone.0170384.t003 http://dx.plos.org/10.1371/journal.pone.0170384.t004 http://dx.plos.org/10.1371/journal.pone.0170384.t005 http://dx.plos.org/10.1371/journal.pone.0170384.t006 http://dx.plos.org/10.1371/journal.pone.0170384.s001 http://dx.plos.org/10.1371/journal.pone.0170384.s002 http://dx.plos.org/10.1371/journal.pone.0170384.s003 http://dx.plos.org/10.1371/journal.pone.0170384.s004 http://dx.plos.org/10.1371/journal.pone.0170384.s005.

Institute., A.N.S., 2011. Authentication of human cell lines: standardization of STR profiling. ANSI/ATCC ASN-0002-2011.

Liscovitch, M., Ravid, D., 2007. 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 Letters 245 (1-2), 350–352.

De Oca, F.M., Macy, M.L., Shannon, J.E., 1969. Isoenzyme Characterization of Animal Cell Cultures. Experimental Biology and Medicine 132 (2), 462–469.

Nims, R.W., Shoemaker, A.P., Bauernschub, M.A., Rec, L.J., Harbell, J.W., 1998. Sensitivity of isoenzyme analysis for the detection of interspecies cell line cross-contamination. In Vitro Cell.Dev.Biol.-Animal 34 (1), 35–39.
Noronha, N., Ehx, G., Meunier, M.-C., Laverdure, J.-P., Thériault, C., Perreault, C., 2020. Major multilevel molecular divergence between THP-1 cells from different biorepositories. Int. J. Cancer 147 (7), 2000–2006.
O'Brien, S.J., Shannon, J.E., Gail, M.H., 1980. A molecular approach to the identification and individualization of human and animal cells in culture: Isozyme and allozyme genetic signatures. In Vitro 16 (2), 119–135.
Popovič, M., Kalyanaraman, V.S., Reitz, M.S., Sarngadharan, M.G., 1982. Identification of the RPMI 8226 retrovirus and its dissemination as a significant contaminant of some widely used human and marmoset cell lines. Int. J. Cancer 30 (1), 93–99.
Wright, W.C., Daniels, W.P., Fogh, J., 1981. Distinction of seventy-one cultured human tumor cell lines by polymorphic enzyme analysis. J Natl Cancer Inst 66, 239–247.
Yu, M., Selvaraj, S.K., Liang-Chu, M.M.Y., Aghajani, S., Busse, M., Yuan, J., Lee, G., Peale, F., Klijn, C., Bourgon, R., Kaminker, J.S., Neve, R.M., 2015. A resource for cell line authentication, annotation and quality control. Nature 520 (7547), 307–311.