Simple, Fast and Accurate Human Cell Line Authentication by Short Tandem Repeat (STR) Analysis on GeXP™ Genetic Analysis System

Jane Luo,[1] Manfred Souquet,[2] Stephan A. Hahn,[3] Britta Redeker,[3] Handy Yowanto[1]
1 SCIEX, Brea, CA
2 SCIEX, Germany GmbH
3 Ruhr University Bochum, Germany

Introduction

Cell lines are routinely used in pharmaceutical industries and biomedical research laboratories as model systems for drug testing and studies on mechanisms of diseases. However, misidentification and cross contamination are persistent issues associated with the use of cell lines in the past 50 years.[1-4] The International Cell Line Authentication Committee (ICLAC) was formed in 2012 to make cell line misidentification more visible and to promote awareness and authentication testing as effective ways to address this issue. An STR test involving the Amelogenin gene and eight of the most informative polymorphic markers in the human genome was recommended to generate unique STR identity profile for each cell line.[5] Cell line authentication has become a prerequisite for NIH grant applications[6] as well as publication on many scientific journals if cell lines are used for scientific research projects.

In this technical note, we describe a process for cell line authentication by STR using the GenomeLab GeXP Genetic Analysis System that offers fast turnaround time (3 hrs) and high identification accuracy. Upon isolation of nucleic acid from each cell line sample, the STR loci were amplified using polymerase chain reaction (PCR), followed by fragment analysis on the GeXP Genetic Analysis System. Results are compared against the American Type Culture Collection (ATCC) public STR database to confirm the identity of the cell line.

Key Features

- GeXP Genetic Analysis System provides simple, fast and accurate cell line authentication by STR profiling
- Detects cell line cross contamination down to 2.5%
- Compatible with WellRED, Cyanine and BMN-6 dyes
- GeXP Genetic Analysis System can be used for human cell line authentication and cell line monitoring in biomedical research, pharmaceutical development, and QC testing laboratories

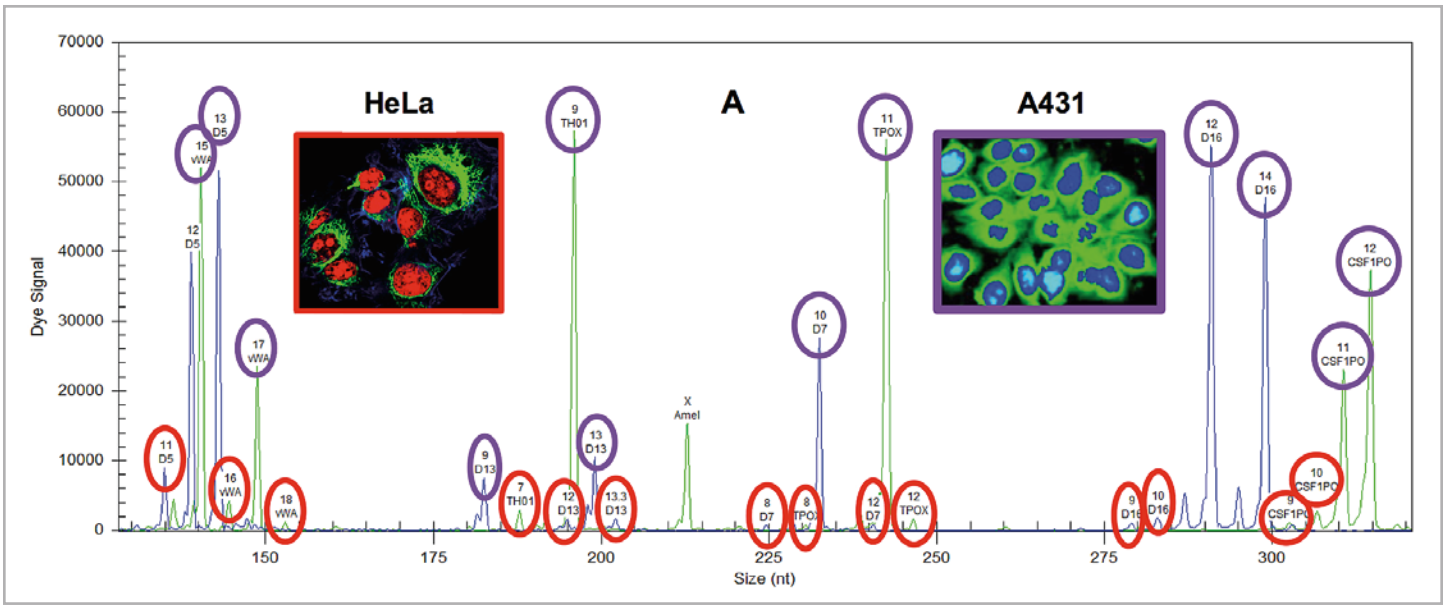

Figure 1A. STR profile of mixed cell lines. D4 (blue trace) and D3 (green trace) dyes were used to label the DNA fragments in this experiment. Genomic DNA from HeLa cells and from A431 cells were mixed at 2.5 to 97.5 ratio before STR analysis. Alleles in red circles are unique to HeLa, the ones in purple circles are unique to A431, the rest are common in both cell lines.

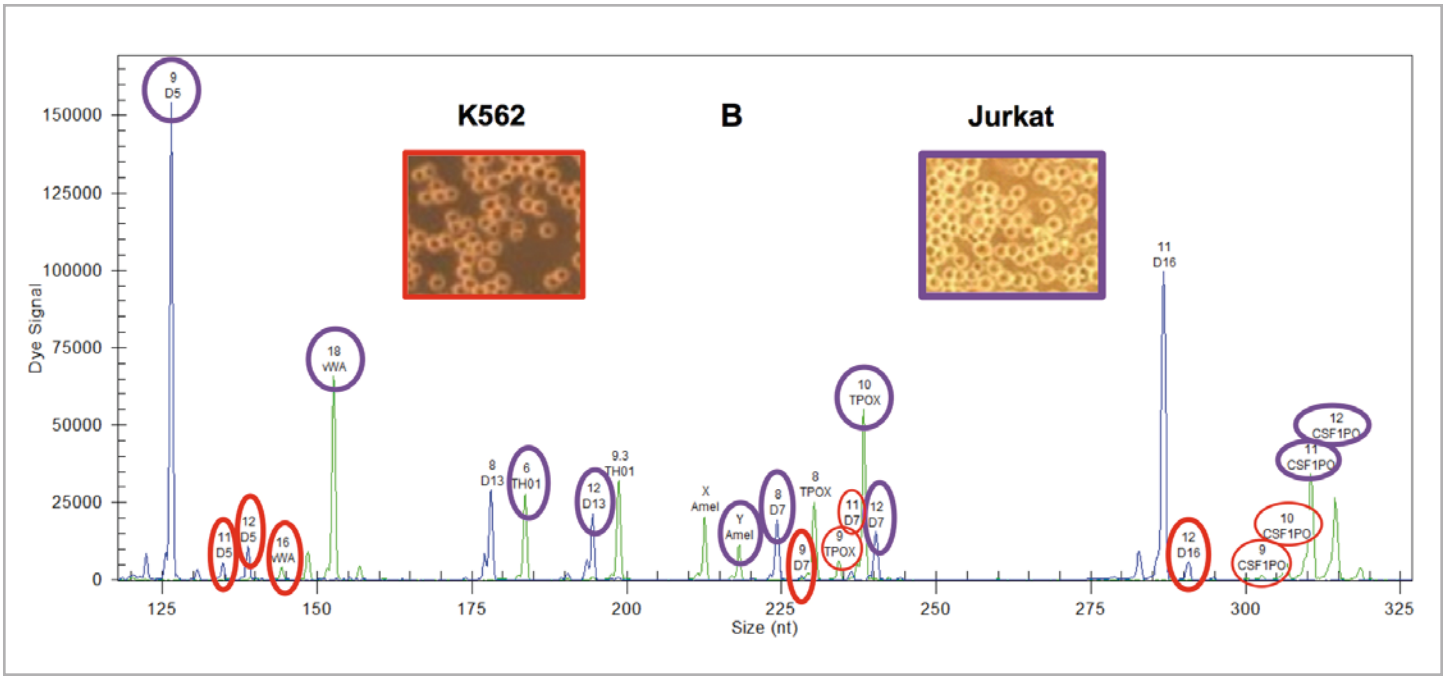

Figure 1B. STR profile of mixed cell lines. D4 (blue trace) and D3 (green trace) dyes were used to label the DNA fragments in this experiment. Genomic DNA from K562 cells and from Jurkat cells were mixed at 2.5 to 97.5 ratio before STR analysis. Alleles in red circles are unique to K562, the ones in purple circles are unique to Jurkat, the rest are common in both cell lines.

Materials and Equipment:

Materials

DRNAse Free surface decontaminant (PN D6002) was from Argos Technologies, Inc (Elgin, IL). Sanihol 70 surface decontaminant (PN 8616) was from Decon Labs Inc (King of Prussia, PA). Bioclean brand disposable filter tips for 1000 ul, 200 ul and 20 µl volumes (PN: GP-L1000F, GP-L200F and GP-L10F) were from Mettler Toledo (Oakland, CA). RNase, DNase free microcentrifuge tubes at the volume of 1.7 ml (PN 20170-355) and 0.65 ml (PN 20170-293) were from VWR (Radnor, PA). Thin-walled PCR tubes (TBS0201) and caps (TCS0803) were from BioRad (Hercules, CA). PCR plates (609801) were from SCIEX (Framingham, MA). DNase, RNase free water (not DEPC-treated, PN AM9930) was from Thermo Fisher Scientific (Carlsbad, CA); WellRED dye and Cy5 Dye labeled primers and unlabeled primers were from MilliporeSigma (The Woodlands, TX). BMN-6 labeled primers were from www.Biomers.net, Germany. Invitrogen Platinum Taq DNA Polymerase (PN 10966-026) and 10 mM dNTP solution (PN 18427-088) were obtained from Thermo Fisher Scientific (Carlsbad, CA); Aluminum foil lid (PN 538619) was from Beckman Life Sciences; DNA array (608087), Separation gel (PN 391438), DNA Size Standard 400 Kit (PN 608098), Sample Loading Solution (SLS, PN 608082), Mineral oil (PN 608114) and Separation Buffer (608012) were from SCIEX (Framingham, MA); K562 genomic DNA (PN E493A) was from Promega; HeLa genomic DNA (PN N4006S) was from New England Biolabs (Ipswich, MA); A431 (PN D1255801) and Jurkat (D1255815) genomic DNA were from BioChain Institute Inc. (Newark, CA).

Equipment

Bench top AirClean 600 PCR Workstation with UV light (PN 4263-2532) was from USA Scientific (Ocala, FL); PHENIX QuickSpin Micro Centrifuges were from Phenix Research Products (Candler, NC); Thermal Cyclers, PTC-200 and S1000 were from BioRad (Hercules, CA). A GenomeLab GeXP Genetic Analysis System (A22033) was from SCIEX (Framingham, MA).

Methods

Prevention of DNA carry-over contamination: Pre-PCR work and Post-PCR were always carried out in separate laboratories. Separate sets of pipettman, tips, plasticware, lab coats, gloves, ice-buckets, and paper towels were used for Pre-PCR and Post-PCR labs. In Pre-PCR room, benchtops were cleaned

with DRNAse Free surface decontaminant. Pipettmans, tube racks and outer surface of plastic containers were sprayed with Sanihol, and wiped dry with clean paper towel. Preparation of PCR reactions was carried out in a bench top AirClean 600 PCR Workstation with UV light within the Pre-PCR room. At the end of each experiment, pipettmans, tube racks, tip boxes and containers for tubes were irradiated with UV light for at least 15 min. Disposable microcentrifuge tubes and filter tips were used for both Pre-PCR and Post-PCR works.

PCR amplification of STR loci in cell line genomic DNA:
Genomic location and sequences for STR primers used in this technical note were as described in Ref 1. PCR reactions were prepared in a PCR Workstation within a Pre-PCR room. Each individual reaction contained the components at amounts shown in Table 1. The concentration of each individual primer was in the range of 50 nM to 150 nM. Usually, a master mix contained every component except the genomic DNA was prepared, and aliquoted into 0.2 ml thin-walled PCR tubes at 23 µl per tube. Then, 2 µl of genomic DNA at 20 to 100 ng/µl was added per tube. After mixing the components by pipetting up and down for a few times, the tubes were capped tightly, moved to Post-PCR room and placed in a PCR thermal cycler (Figure 1). Amplification of STR loci were carried out using the following thermal cycling conditions: First, the Platinum Taq polymerase was activated at 95° C for 5 min. Next, 10 cycles of 94° C for 30 seconds, 60° C for 20 seconds and 70° C for 45 seconds were performed. Then, another 20 cycles of 90° C for 30 seconds, 60° C for 20 seconds and 70° C for 45 seconds were carried out, followed by a 30 min incubation at 60° C for 30 min. At the end, the PCR reactions were hold at 4° C.

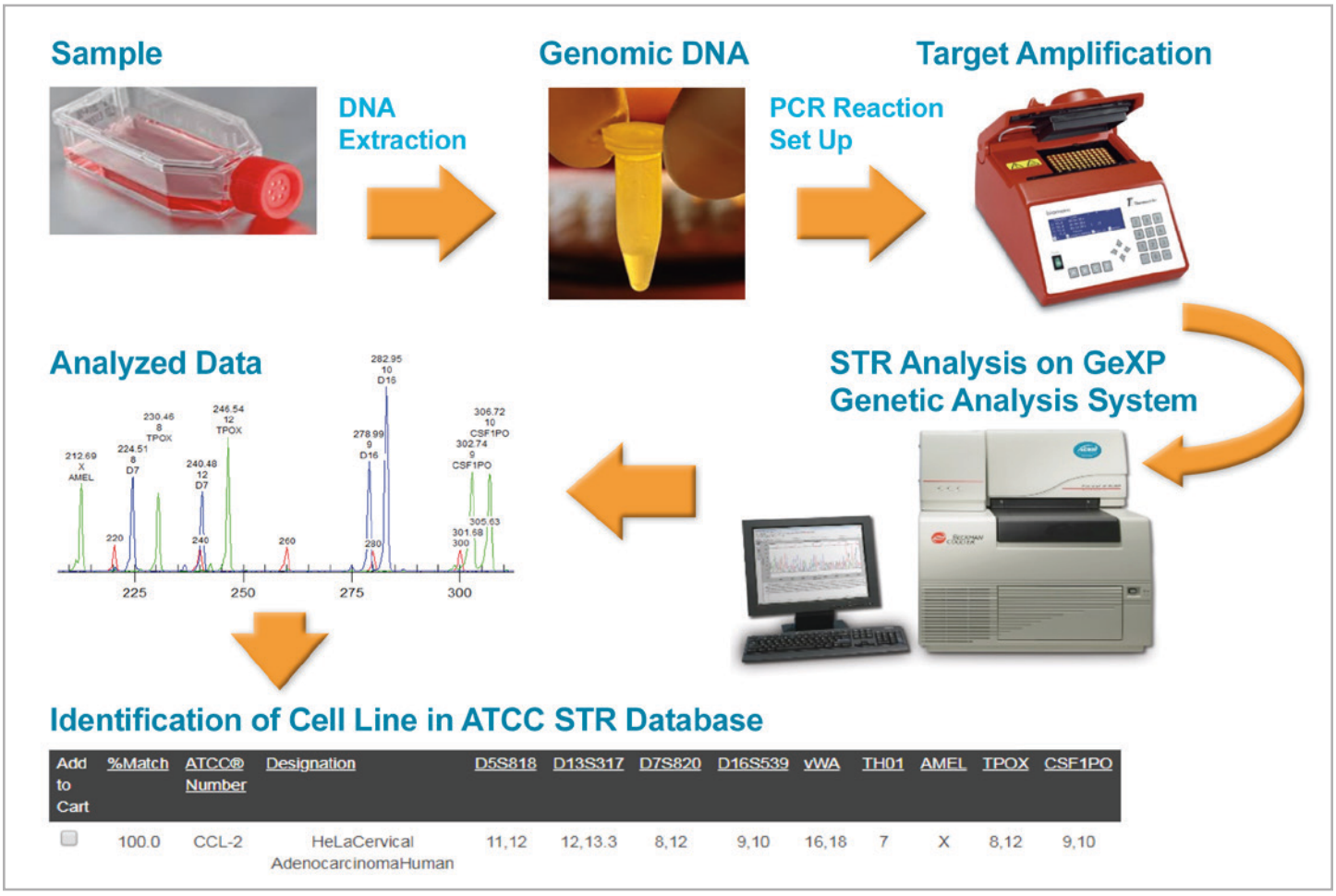

Figure 2. Work Flow for cell line authentication using GenomeLab GeXP Genetic Analysis System.

Component	Volume (µl)
DNase, RNase Free Water	17.8
10 x PCR buffer without $MgCl_2$	2.5
50 mM $MgCl_2$	1
10 mM dNTP	0.5
25 x Primer Stock Solution	1
Platinum Taq DNA Polymerase	0.2
Genomic DNA (20 to 100 ng per µl)	2
Total Volume	**25**

Table 1. Components in each PCR reaction.

Separation of dye labeled PCR products by capillary electrophoresis (CE): After thermal cycling was complete, each sample was diluted 20-fold with DNase RNase free water. One microliter of this diluted sample was added to a mixture of 30 µl Sample Loading Solution (SLS) and 0.3 µl DNA Size Standard 400 (SS-400) on a GeXP sample plate (PN 609801). The sample plate was covered with a piece of aluminum foil lid and centrifuged at 1000 rpm for 2 minutes to remove any air bubbles. After one drop of mineral oil was overlaid on top of the samples per well, the sample plate was loaded onto the GenomeLab GeXP Genetic Analysis System (Figure 1). Capillaries were pre-heated to 50° C before separation was started. During this time, manifold was purged 3 times with 0.4 ml of separation gel, capillaries were filled 2 times with fresh gel, wetting tray was refreshed. Separation was performed using the separation method Frag-3 as shown in Table 2. Further sample dilutions may be needed if the signal gets saturated.

Capillary Temperature	50° C
Denaturation Temperature and Time	90° C for 120 seconds
Pause Duration	1 min
Injection Voltage and Time	2.0 kV for 30 seconds
Separation Voltage and Time	6.0 kV for 35 min

Table 2. Frag-3 Run Method.

Data Analysis

Raw data was automatically analyzed using Human Cell Line STR Analysis Parameters containing locus tags for the 8 STR loci and Amelogenin gene. Locus tags were created based on STR Fact Sheets in the NIST STR database (https://strbase.nist.gov[7]) and apparent sizes of STR fragments determined by the GeXP software when genomic DNA from known cell lines were analyzed. After analysis, DNA fragments will be labeled with locus tags. A genotype summary report will be created by the GeXP software. Automated export option is available for further analysis with third party tools.

Search in ATCC Public STR Database

After the genotype summary report is created for a cell line sample, the identity of the cell line was confirmed by searching the ATCC public STR database. A simple registration was required to access the database. Once logged in, the database was searched by entering the allele IDs for each STR locus. For homozygotes, one allele ID was entered. For heterozygotes, two allele IDs were entered with a comma in between. After the “Matches >= 80%” button was clicked, search results were listed in a table with percentage of matches, corresponding ATCC cell line number, name of the cell line and genotype of each locus.

Results

Capability of detecting cell line cross contamination on GeXP: Genomic DNA from K562 cells and from HeLa cells were mixed at different ratio to simulate genomic DNA isolated from cross contaminated cell lines. Figure 1A shows results obtained when genomic DNA from HeLa cells were mixed with genomic DNA from A431 cells in a 2.5 to 97.5 ratio. The presence of alleles to both cell lines were detected. Similarly, when genomic DNA from K562 cells were mixed with genomic DNA from Jurkat cells in a 2.5 to 97.5 ratio, the STR profile (Figure 1B) also shows the presence of alleles from both cell lines.

Accurate STR profiling of four different cell lines on GeXP: PCR amplification of 8 STR loci and Amelogenin gene was carried out with genomic DNA isolated from 4 cell lines (A431, K562, Jurkat and Hela) as described in Methods. After separation of dye labeled PCR amplicons by CE, raw data was automatically analyzed by the GeXP software using Human Cell Line STR Analysis Parameters. Locus tags and allele IDs were assigned to DNA fragments based on their apparent sizes (Figure 3A). Information on allele ID for each of the 8 STR loci and for Amelogenin gene was entered into the ATCC STR database. Search results shown in Figures 3B to 3E indicated that each of the 4 cell lines were correctly identified based on its STR profile.

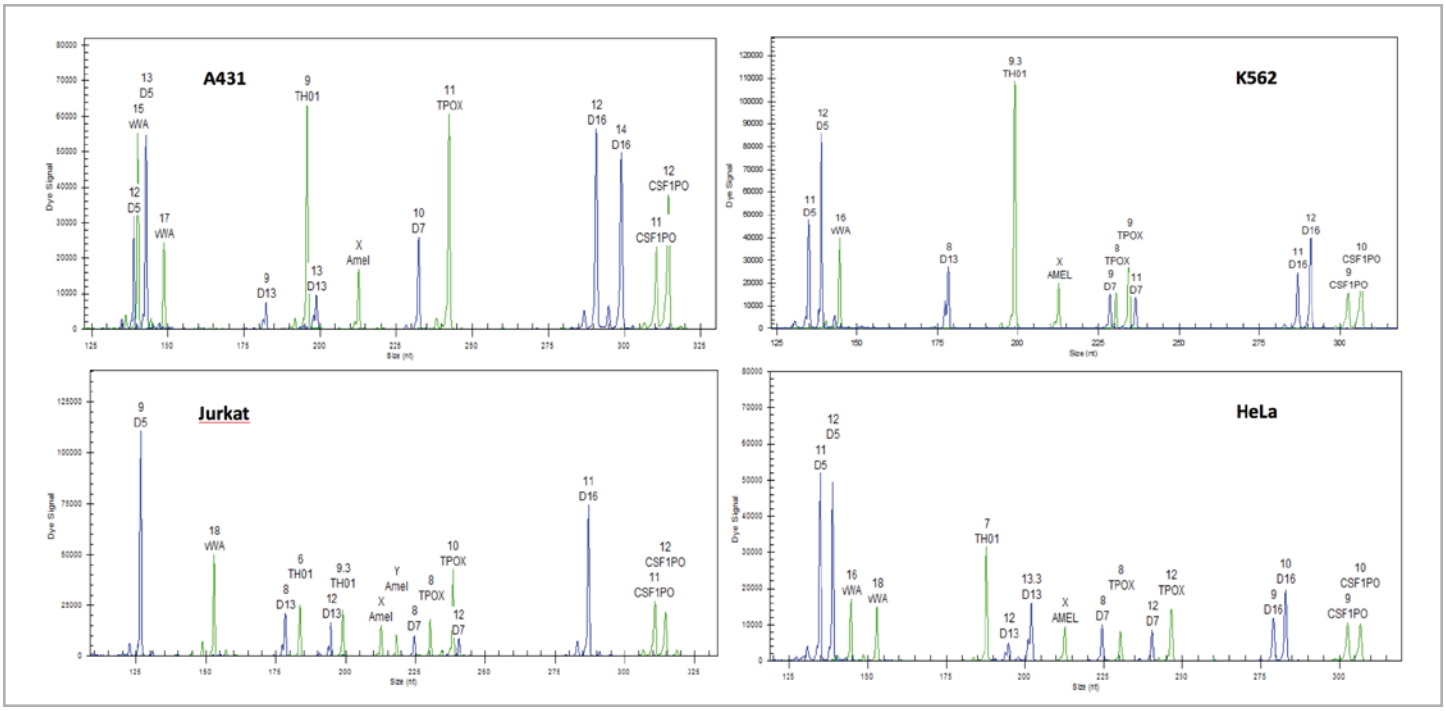

Figure 3A. STR profiles for four different cell lines: A431, K562, Jurkat and HeLa. Each Electropherogram shows results obtained with genomic DNA from one cell line. In this experiment, DNA fragments were labeled with Beckman WellRED dyes D4 (blue trace) and D3 (green trace).

%Match	ATCC® Number	Designation	D5S818	D13S317	D7S820	D16S539	vWA	TH01	AMEL	TPOX	CSF1PO
100.0	CRL-1555	A-431Epidermoid CarcinomaHuman	12,13	9,13	10	12,14	15,17	9	X	11	11,12
100.0	CRL-2592	A431NSEpidermoid CarcinomaHuman	12,13	9,13	10	12,14	15,17	9	X	11	11,12

Figure 3B. A-431 cell line was correctly identified by ATCC STR database based on allele ID information entered.

%Match	ATCC® Number	Designation	D5S818	D13S317	D7S820	D16S539	vWA	TH01	AMEL	TPOX	CSF1PO
100.0	CCL-243	K-562Leukemia (CML)Human	11,12	8	9,11	11,12	16	9.3	X	8,9	9,10

Figure 3C. K562 cell line was correctly identified by ATCC STR database based on allele ID information entered.

%Match	ATCC® Number	Designation	D5S818	D13S317	D7S820	D16S539	vWA	TH01	AMEL	TPOX	CSF1PO
100.0	TIB-152	Jurkat, Clone E6-1Acute T Cell LeukemiaHuman	9	8,12	8,12	11	18	6,9.3	X,Y	8,10	11,12
100.0	CRL-2898	Neo JurkatAcute T Cell LeukemiaHuman	9	8,12	8,12	11	18	6,9.3	X,Y	8,10	11,12
100.0	CRL-2900	BCL2 (S70A) JurkatAcute T Cell LeukemiaHuman (Homo sapiens)	9	8,12	8,12	11	18	6,9.3	X,Y	8,10	11,12
100.0	CRL-2901	BCL2 (S87A) JurkatAcute T Cell LeukemiaHuman (Homo sapiens)	9	8,12	8,12	11	18	6,9.3	X,Y	8,10	11,12

Figure 3D. Jurkat cell line was correctly identified by ATCC STR database based on allele ID information entered.

Add to Cart	%Match	ATCC® Number	Designation	D5S818	D13S317	D7S820	D16S539	vWA	TH01	AMEL	TPOX	CSF1PO
☐	100.0	CCL-2	HeLaCervical AdenocarcinomaHuman	11,12	12,13.3	8,12	9,10	16,18	7	X	8,12	9,10

Figure 3E. HeLa cell line was correctly identified by ATCC STR database based on allele ID information entered.

Using Cy5 and BMN-6 dyes as alternatives to Beckman WellRED dyes: Cy5 and BMN-6 dyes are two commonly used alternative dyes for labeling primers in DNA fragment analysis by capillary electrophoresis. STR analysis was done with genomic DNA from HeLa and K562 cells using DNA primers labeled with Cy5 and BMN-6 dyes. As shown in Figure 4, STR profiles obtained with Cy5 and BMN-6 dyes for these two cell lines were identical to those obtained using Beckman WellRED dyes D4 and D3. There were slight changes on migration time and apparent sizes of some fragments due to differences on dye structure. These changes can be easily taken care of by adjusting the apparent sizes in locus tags accordingly.

Discussions

With the combined international efforts coordinated by the ICLAC, cell line authentication has become an essential task to scientists in biomedical research and pharmaceutical industries. Guidelines for STR based cell line authentication have been published by ATCC.[8] Several interactive, searchable STR databases are now available for public use. Among them, the ATCC and DSMZ cell line STR databases are the most user-friendly. The BioSample database at NCBI website provides a searchable list of Human Cell Line STR Profiles. Another important tool provided by ATCC, DSMZ and BioSample is the list of Known Misidentified Cell Lines. It is now a recommended practice that scientists authenticate their cell lines and check them against the list of known misidentified cell lines before starting their projects.

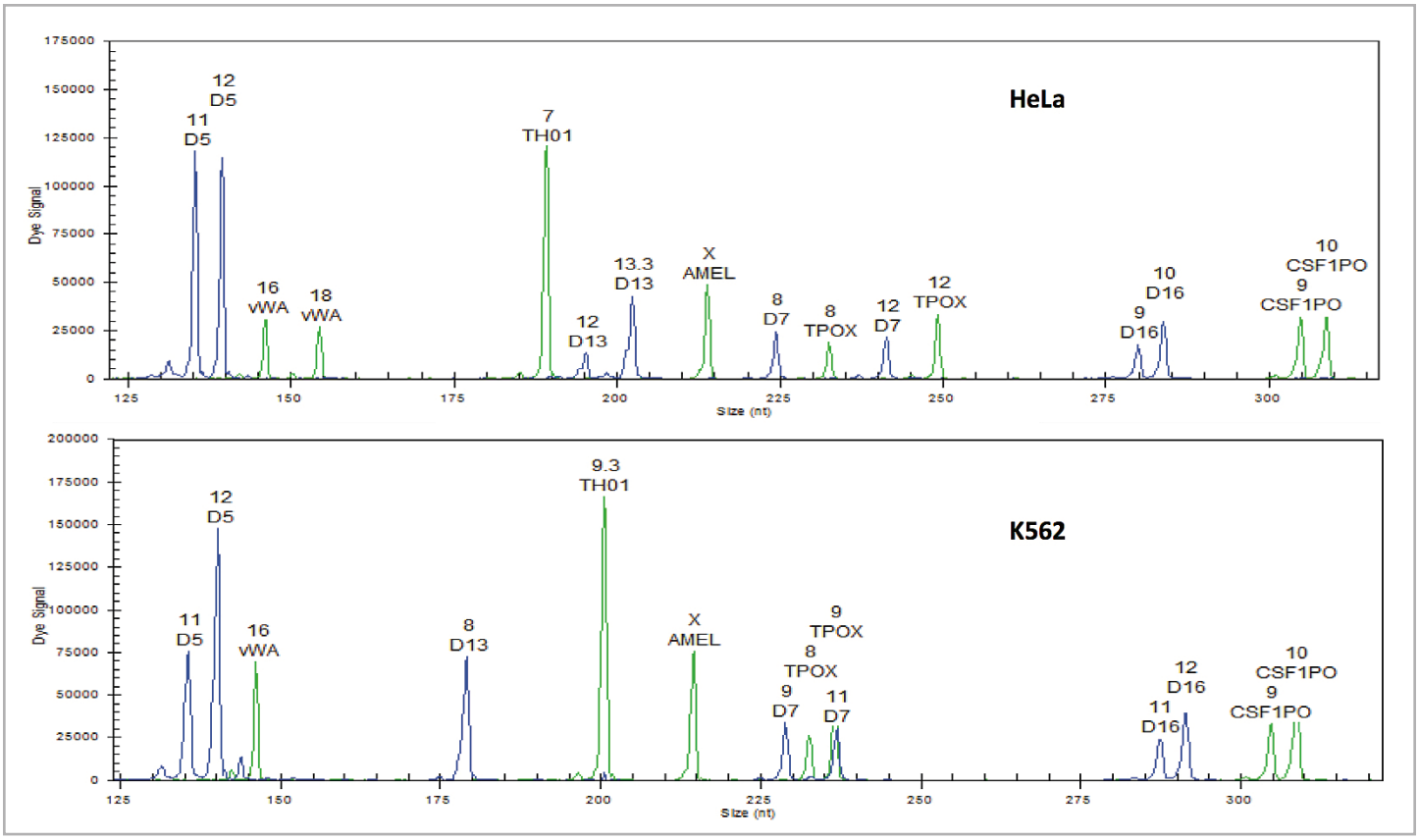

Figure 4. STR profiles for HeLa and K562 cells obtained using Cy5 (Blue trace) and BMN-6 (green trace) dyes.

Once a project starts, it is often necessary to perform frequent cell line authentication tests throughout the duration of the project. For example, in drug discovery industry, human cell lines are often used to select drug resistant clones or stable transfectants. Since cell lines can be genetically heterogeneous, the selection process may enrich cross contaminated cells or a population of a subclone that no longer reflects the initial cell line. Frequent cell line tests will detect these problems promptly. Other cases that could benefit from frequent cell line authenticity test include studies on the contribution of Y chromosome in certain diseases. Some male cell lines would lose the Y chromosome during passages in cell culture.[9] Frequent authentication test for these cell lines is important to ensure results obtained are meaningful and relevant.

The GenomeLab GeXP Genetic Analysis System is an ideal system for STR authentication of cell lines. As shown in Figure 2, the work flow of STR analysis on GeXP Genetic Analysis System is simple and fast. It takes only 3 hours from setting up PCR reactions with genomic DNA samples to completion of automated STR analysis. In addition, the GeXP Genetic Analysis System is sensitive in detecting cell line cross contamination. Figure 1 illustrated the capability of detecting 2.5% contamination of HeLa cells in A431 cells as well as 2.5% contamination of K562 cells in Jurkat cells. Both HeLa cells and A431 cells are adherent cells. Both K562 cells and Jurkat cells are non-adherent cells. It is difficult to recognize the contaminating cells by morphology when they are present at a very low percentage. Therefore, STR on GeXP Genetic Analysis System provides a great tool for monitoring cell line and for detecting cell line cross contamination early. Furthermore, users can use Beckman WellRED dyes, Cyanine dyes and BMN-6 dyes for doing the STR analysis on GeXP Genetic Analysis System.

Conclusions

1 GeXP Genetic Analysis System provides simple, fast and accurate cell line authentication by STR profiling.

2 It detects cell line cross contamination down to 2.5%.

3 It is compatible with WellRED dyes, Cyanine dyes and BMN-6 dyes.

4 GeXP Genetic Analysis System can be used for human cell line authentication and cell line monitoring in biomedical research, pharmaceutical development, and QC testing laboratories.

Who is SCIEX? SCIEX company's global leadership and world- class service and support in the capillary electrophoresis and liquid chromatography-mass spectrometry industry have made it a trusted partner to thousands of the scientists and lab analysts worldwide who are focused on basic research, drug discovery and development, food and environmental testing, forensics and clinical research.

Contact Us: sciex.com/contact-us

References

1. W. G. Dirks, S. Fähnrich, I. A. J. Estella and H. G Drexler. Short Tandem Repeat DNA typing provides an international reference standard for authentication of human cell lines (2005) ALTEX 22(2), 107-109.

2. M. Lacroix. Persistent use of "false" cell lines (2008) *Int J Cancer*. 122:1-4.

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

4. L. Vaughan, W. Glänzel, C. Korch and A. Capes-Davis. Widespread use of misidentified cell line KB (HeLa): Incorrect attribution and its impact revealed through mining the scientific literature (2017) *Cancer Res.* 77(11): 1-5.

5. Cell Line Authentication Test Recommendations (2010) ATCC Technical Bulletin, #8.

6. Enhancing Reproducibility through Rigor and Transparency (2015) NIH Notice Number: NOT-OD-15-103. Release Date: 6/9/2015.

7. C. M. Ruitberg, D. J. Reeder, J. M. Butler. Short tandem repeat DNA database for the human identity testing community (2001) *Nucleic Acids Research* 29(1), 320-322.

8. ATCC Standards Development Organization (ATCC SDO) The American National Standard (ANS) ANSI/ATCC ASN-0002-2011. Authentication of Human Cell Lines: Standardization of STR Profiling.

9. K. Shah, C. E. McCormack, N. A. Bradbury. Do you know the sex of your cells? (2014) *Am J Physiol Cell Physiol.* 306(1):C3-18.

Document number: RUO-MKT-02-8785-A 12/2018

Headquarters
500 Old Connecticut Path, Framingham, MA 01701, USA
Phone 508-383-7800
sciex.com

International Sales
For our office locations please call the division headquarters or refer to our website at
sciex.com/offices