HITACHI
Inspire the Next

Cell Line Authentication with the DS3000 Compact CE Sequencer

Abstract

Cell line contamination, misidentification, and genetic drift result in invalid research.
It has been estimated that 18% to 36% of cell lines are contaminated or misidentified (1). To address this issue, a number of scientific journals request researchers to perform cell line authentication (CLA) prior to submitting a manuscript that involves the use of cell lines (2). Short tandem repeat (STR) analysis is a commonly used method for CLA. STRs are short repetitive sequences that are scattered throughout the whole genome. As each individual harbors a different number of repeats, STRs are powerful tool to identify the origin of biomaterials. To perform CLA, multiple STRs are enriched by polymerase chain reaction (PCR). Then, the number of repeats (alleles) at each locus (marker) are identified by capillary electrophoresis. Finally, the identified alleles are compared to alleles registered in biobanks, such as the American Type Culture Collection (ATCC) (3); we refer to such registered alleles as reference alleles. Considering the impact and importance of CLA, instruments for CLA must have high reliability and accuracy.
Here, we introduce a practical example of CLA using the DS3000 Compact CE Sequencer (DS3000). More than 80% of alleles observed using DS3000 combined with the GenePrint™ 10 System or GenePrint™ 24 System (Promega®) agreed with the reference alleles. This rate is higher than the generally accepted threshold to claim the authentication of the cell line (3). These results indicate that the performance of DS3000 is sufficient for CLA.

DS3000
Compact CE Sequencer

Results

Analysis of non-mixed cell lines

We used the genomic DNA extracted from five commercially available human cell lines (HeLa, Jurkat, MCF7, K562, and 2800M DNA). Reference alleles are available for all five cell lines (see Materials and Methods). STRs in the genomic DNA were amplified with the GenePrint™ 10 System or GenePrint™ 24 System according to the manufacturer's instructions. The resulting products were electrophoresed on DS3000, followed by analysis on GeneMarker HID (ver. 2.9.0, SoftGenetics) or GeneMapper™ ID-X (ver. 1.6, Applied Biosystems®).
Figure 1 (GenePrint™ 10 System) and Figure 2 (GenePrint™ 24 System) show representative examples of electropherograms. Tables 1 and 2 summarize the observed alleles and expected alleles in each cell line. The results showed that more than 80% of the observed alleles matched with the reference alleles. The percent match was calculated by Master's formula (number of observed alleles that matched with the reference / total number of observed alleles) (4).

Table 1: Markers and alleles identified with the GenePrint™ 10 System

Marker	HeLa				Jurkat				MCF7			
	Observed allele		Reference allele		Observed allele		Reference allele		Observed allele		Reference allele	
AMEL	X		X		X	Y	X	Y	X		X	
CSF1PO	9	10	9	10	11	12	11	12	10		10	
D13S317	12		12	13.3	8	12	8	12	11		11	
D16S539	9	10	9	10	11		11		11	12	11	12
D21S11	27	28			31.2	33.2			30			
D5S818	11	12	11	12	9		9		11	12	11	12
D7S820	8	12	8	12	8	12	8	12	8	9	8	9
TH01	7		7		6	9.3	6	9.3	6		6	
TPOX	8	12	8	12	8	10	8	10	9	12	9	12
vWA	16	18	16	18	18		18		14	15	14	15

Marker	K562				2800M DNA			
	Observed allele		Reference allele		Observed allele		Reference allele	
AMEL	X		X		X	Y	X	Y
CSF1PO	10		9	10	12		12	
D13S317	8		8		9	11	9	11
D16S539	11	12	11	12	9	13	9	13
D21S11					29	31.2	29	31.2
D5S818	11	12	11	12	12		12	
D7S820	9	11	9	11	8	11	8	11
TH01	9.3		9.3		6	9.3	6	9.3
TPOX	8	9	8	9	11		11	
vWA	16		16		16	19	16	19

The gray rows indicate the marker that is not registered in the ATCC database.

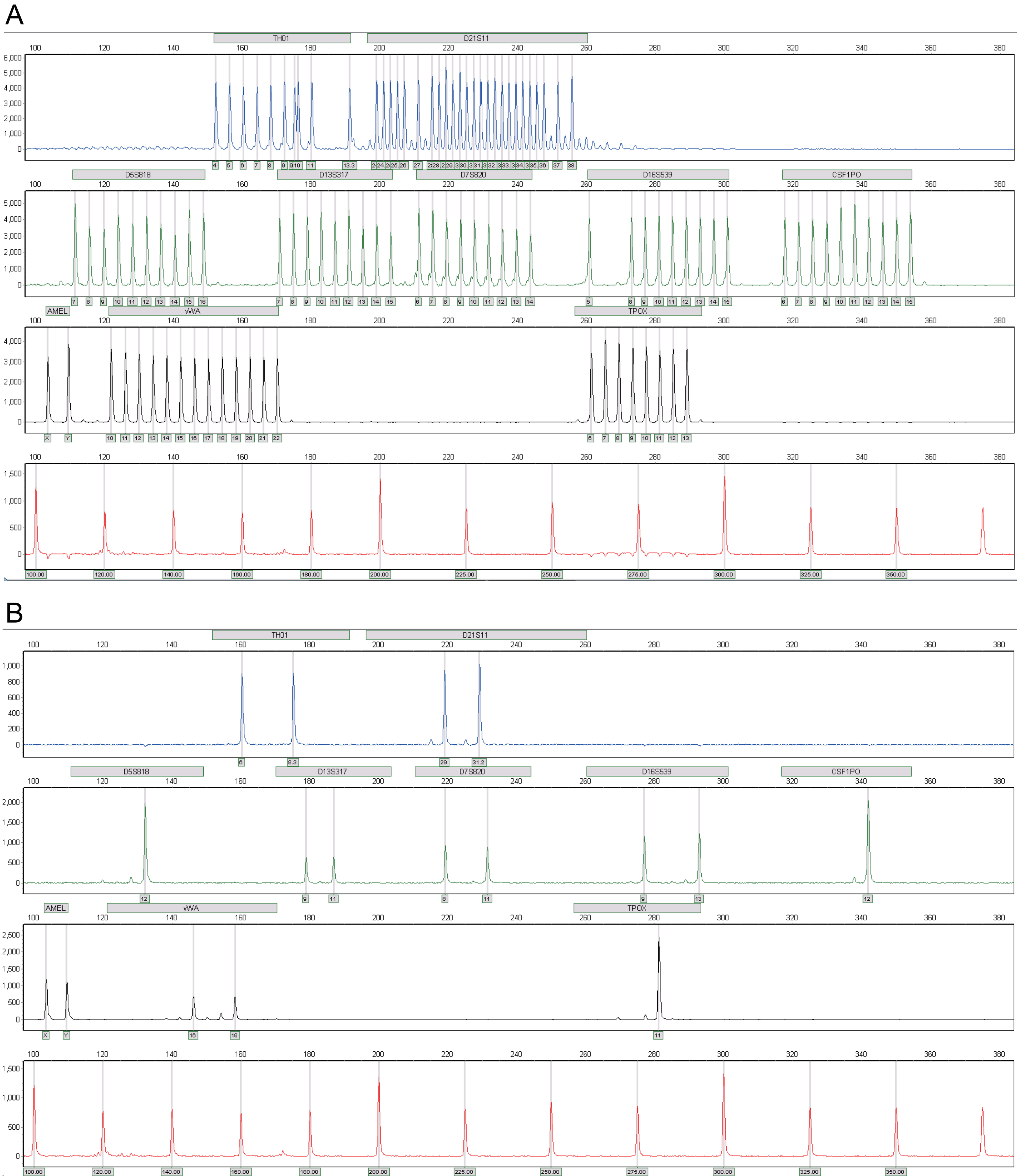

Figure 1: Electropherograms obtained with the GenePrint™ 10 System

Representative examples of the allelic ladder (A) and 2800M DNA (B) are shown.
The x-axis shows the size of the DNA fragment (bp), and the y-axis shows the signal intensity (RFU).

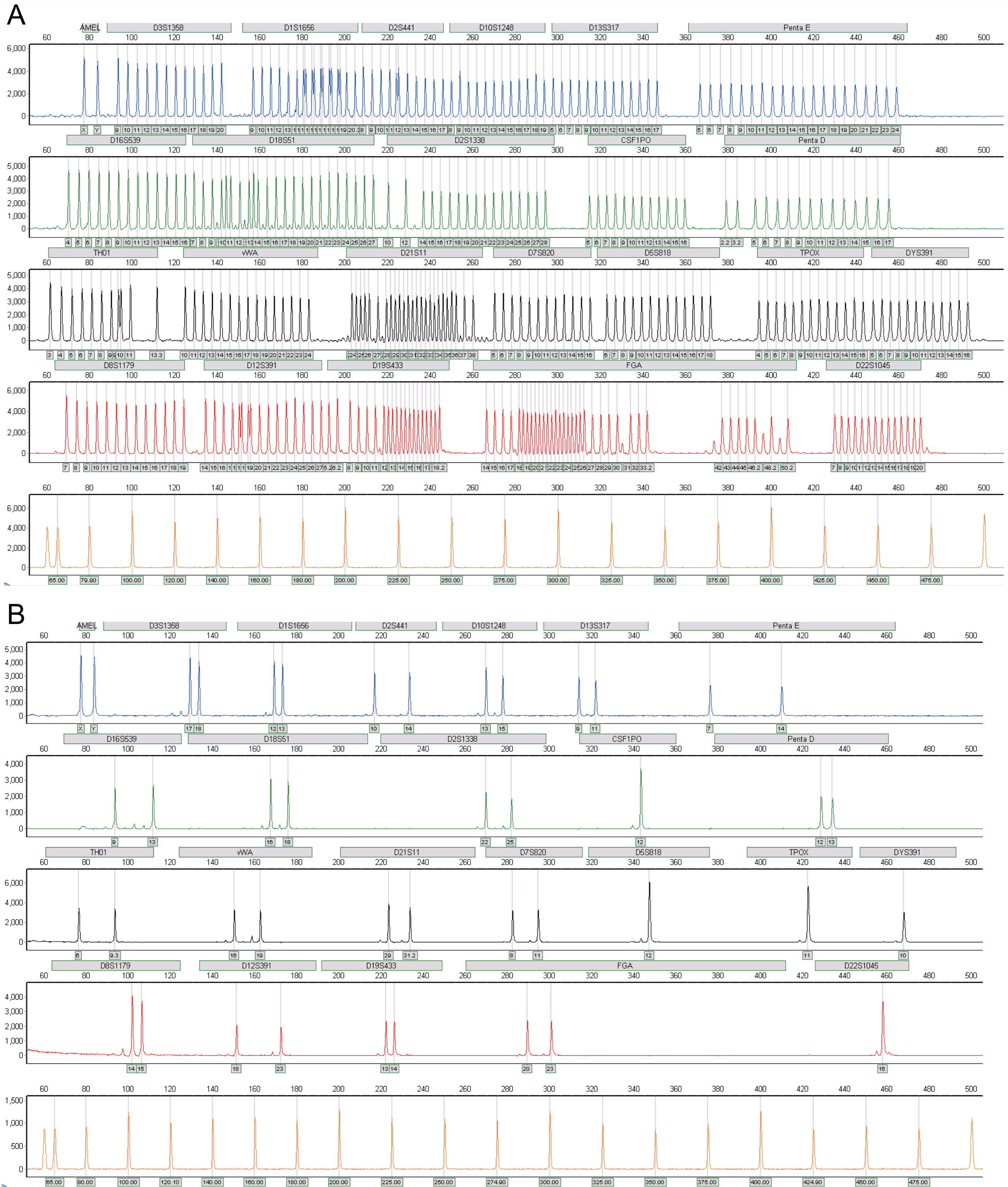

Figure 2: Electropherograms obtained with the GenePrint™ 24 System

Representative examples of the allelic ladder (A) and 2800M DNA (B) are shown.
The x-axis shows the size of the DNA fragment (bp), and the y-axis shows the signal intensity (RFU).

Analysis of mixed cell lines

We evaluated the ability of DS3000 to detect sample contamination by mixing the genomic DNA of two different cell lines. MCF7 genomic DNA was added to Jurkat genomic DNA to prepare samples containing 0%, 1%, 2%, 3%, 4%, 5%, 10%, 30%, and 100% MCF7 DNA. STRs in the mixed DNA samples were amplified by the GenePrint™ 10 System or GenePrint™ 24 System. The resulting products were subjected to analysis on DS3000. Figure 3 shows an example of the migration pattern in D13S317 with the GenePrint™ 10 System. Jurkat-specific alleles (alleles 8 and 12) and an MCF7-specific allele (allele 11) are shown. The signal intensity of the MCF7-specific allele increased in proportion to the amount of MCF7 DNA. The percent match of the mixed samples with the Jurkat reference alleles is summarized in Figure 4. The percent match was less than 80% when the amount of MCF7 was higher than 3%. Since the threshold of 80% has been established to indicate the absence of contamination in cell lines, these results indicated that the performance of DS3000 was sufficient for detecting contamination of 3% or

Table 2: Percent match between the observed and reference alleles identified with the GenePrint™ 10 System

Item	HeLa	Jurkat	MCF7	K562	2800M DNA
Number of alleles that matched with the reference	15	15	14	13	15
Total number of observed alleles	15	15	14	13	15
Percent match	100%	100%	100%	100%	100%

The percent match was calculated by Master's formula (number of observed alleles that matched with the reference / total number of observed alleles) (4).

Conclusion

In this study, we validated the performance of DS3000 for CLA. We performed STR analysis in five cell lines. In all cell lines, DS3000 combined with a STR analysis kit, GenePrint™ 10 System or GenePrint™ 24 System, successfully identified alleles with a percent match greater than 80% to reference alleles. In addition, we analyzed a mixture of MCF7 and Jurkat DNA. DS3000 could identify contamination even when the mixture contained only 3% MCF7 DNA.

These results indicate that DS3000 is fully capable of meeting the performance requirements for CLA using the cell line-derived genomic DNA and STR reagents employed in this study.

Reference

1, Hughes, P., et al. The costs of using unauthenticated, over-passaged cell lines: how much more data do we need? *BioTechniques* **43**: 575-586 (2007).

2, https://www.atcc.org/the-science/authentication/cell-line-authentication-publication-requirements

3, American Type Culture Collection, Authentication of Human Cell Lines: Standardization of STR Profiling. ANSI/ATCC ASN-0002-2011.

4, Capes-Davis, A., et al. Match criteria for human cell line authentication: Where do we draw the line? *Int. J. Cancer* **132**, 2510–2519 (2013)

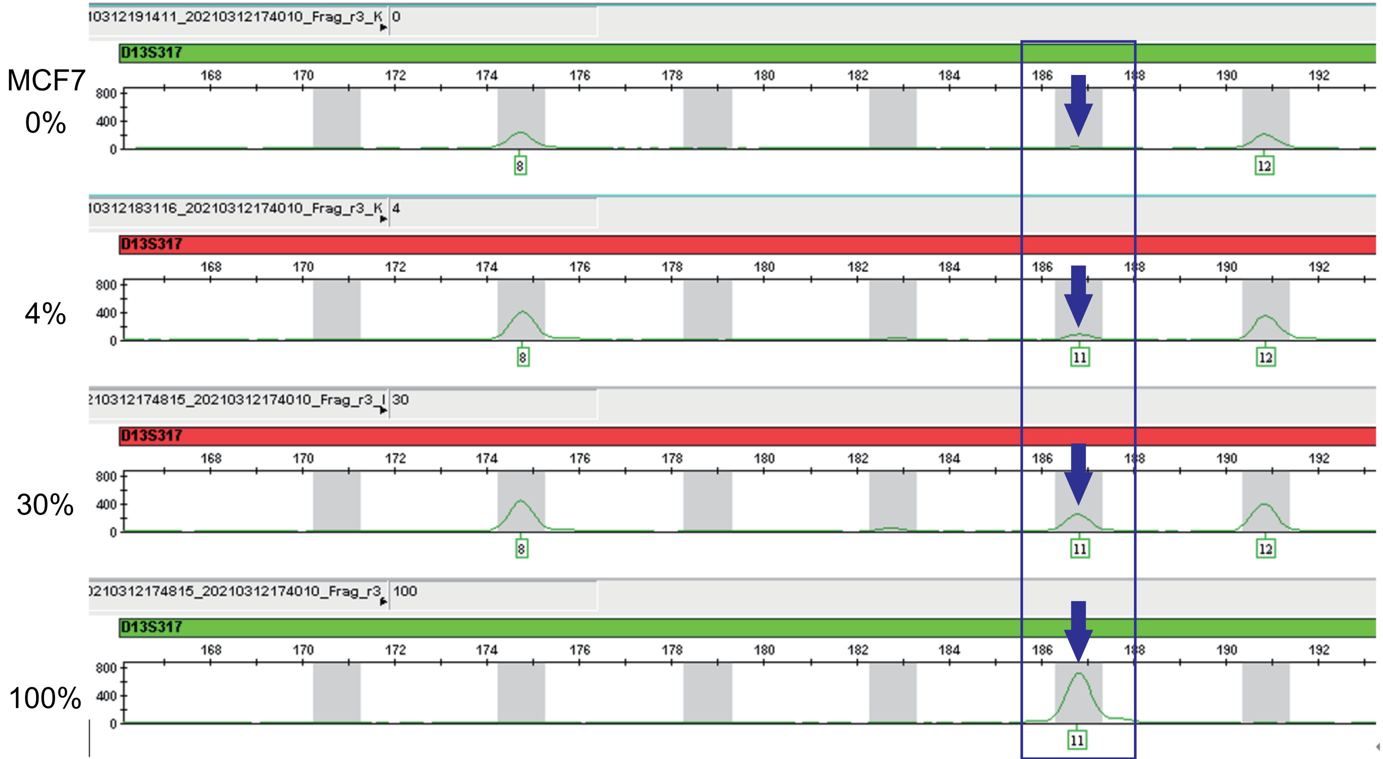

Figure 3: Electropherograms of mixed DNA samples (Jurkat and MCF7 DNA)

The signal intensity of the MCF7-specific allele (D13S317: 11) increased in relation to the amount of MCF7 DNA.
The x-axis shows the size of the DNA fragment (bp), and the y-axis shows the signal intensity (RFU).

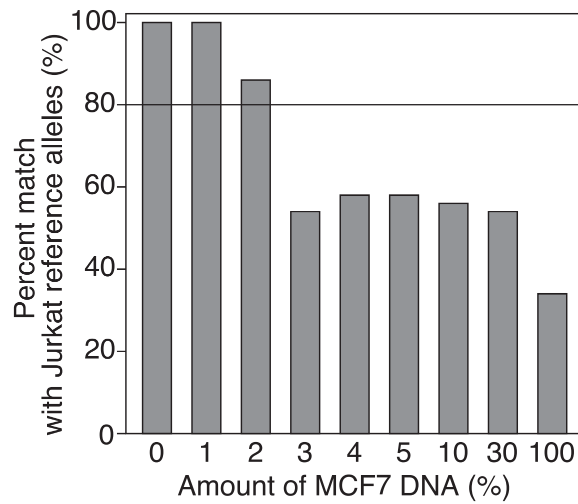

Figure 4:
Percent match of the mixed DNA samples (Jurkat and MCF7 DNA) with the Jurkat reference alleles.

The x-axis shows the amount of MCF7 DNA (%).

HITACHI
Inspire the Next

Materials and Methods

HeLa genomic DNA, Jurkat genomic DNA, and MCF7 genomic DNA were obtained from BioChain Institute Inc. K562 genomic DNA and 2800M DNA were obtained from Promega®. STRs were amplified according to the manufacturer's instructions. Either 5 ng (GenePrint™ 10 System) or 10 ng (GenePrint™ 24 System) of genomic DNA was used for STR amplification. Capillary electrophoresis was performed with the settings shown in Table 3. Size calling, allele calling, and visualization of the electropherograms were performed using GeneMarker HID (ver. 2.9.0, SoftGenetics) and GeneMapper™ ID-X (ver. 1.6, Applied Biosystems®). Reference alleles are available from the ATCC (HeLa, Jurkat, MCF7, and K562) and Promega® (2800M DNA).

Table 3: Instrumental settings for electrophoresis

STR kit	Polymer	Module	Run time	Run voltage
GenePrint® 10 System	Polymer4	Promega_4Dye_ILS600_36_P4	ca. 45 min	13 kV
GenePrint® 24 System	Polymer7	Promega_5Dye_WENILS_36_P7	ca. 35 min	13 kV

Main specifications for the DS3000

■ Main unit specifications

Item	Details
Number of capillaries	4
Capillary length	36 cm
Sample format	8-tube strip × 4
Device control	Touch panel PC
Number of Dyes	6
Application	Sequencing analysis / Fragment analysis
Size	400 (W) × 600 (D) × 600 (H) mm
Weight	45 kg
Performance guarantee temperature	15 – 30°C
Performance guarantee humidity	20 – 80% RH (no condensation)
Power input	100 – 240 ±10% VAC, 50/60 Hz
Rated power	260 VA
Supported secondary analysis software	• Mutation Surveyor (SoftGenetics, LLC, sold separately) • GeneMarker (SoftGenetics, LLC, sold separately) • GeneMarker HID (SoftGenetics, LLC, sold separately)

■ Run module specifications

Run Module	Application	Polymer type	Contiguous Read Length*1 (bp, QV20 CRL)	Average run time (minutes)
Fast_Sequence36_Polymer7	Sequencing analysis	Polymer7	≥600	≤32
Standard_Sequence36_Polymer7	Sequencing analysis	Polymer7	≥700	≤60
BDx_Fast_Sequence36_Polymer7	BDx sequencing analysis	Polymer7	≥600	≤32
BDx_Standard_Sequence36_Polymer7	BDx sequencing analysis	Polymer7	≥700	≤60

Run Module	Application	Polymer type	Average run time (minutes)	Sizing precision*2 (bp, 50 – 400 bp)
Fragment_Analysis36_Polymer7	Fragment analysis	Polymer7	≤35	NA
Fragment_Analysis36_Polymer4	Fragment analysis	Polymer4	≤44	<0.16

*1 Contiguous Read Length (bp, QV20 CRL) is measured with BigDye® Terminator v3.1. Sequencing Standard Kit (Thermo Fisher Scientific, sold separately)

*2 Sizing precision (bp, 50 - 400 bp) is measured with PowerPlex™ ESI17 Fast Allelic Ladder and WEN ILS 500 ESS. (Promega®, sold separately)

· BigDye is a registered trademark of Thermo Fisher Scientific Inc.

· Promega, PowerPlex is a registered trademark of Promega Corporation.

■ Consumables specifications

Product name	Part number	Details	Remarks
Capillary Cartridge 36 cm	613-0330	1 pcs	Storage temperature: 15 – 30°C
Buffer	613-0252	Anode Buffer × 2 cartridges Cathode Buffer × 2 cartridges	Storage temperature: 2 – 10°C
Polymer7	613-0251	4 cartridges	Storage temperature: 2 – 10°C
Polymer4	613-0250	4 cartridges	Storage temperature: 2 – 10°C
Septa for Cathode Buffer Cartridge	613-7231	10 pcs	
Retainer for Cathode Buffer Cartridge	613-7233	4 pcs	
Septa for 8 well tubes	613-7230	24 pcs	
Base and Retainer for 8 well tubes	613-7232	4 pcs	
Anode Electrode Assembly	613-7263	1 pcs	

Specifications in this catalog are subject to change with or without notice,
as Hitachi High-Tech Corporation continues to develop the latest technologies and product for its customers.

CAUTION:For correct operation, follow the instruction manual when using the instrument.

NOTICE:The system is For Research Use Only, and is not intended for any animal or human therapeutic or diagnostic use.
Hitachi High-Tech does not guarantee the performance on the document with every possible sample under every possible condition.
Copyright (C) Hitachi High-Tech Corporation 2022 All rights reserved.

Tokyo, Japan
www.hitachi-hightech.com/global/science/
Toranomon Hills Business Tower, 1-17-1 Toranomon, Minato-ku,
Tokyo 105-6409, Japan
customercenter.ev@hitachi-hightech.com

HTB-DS-E007 2022.6