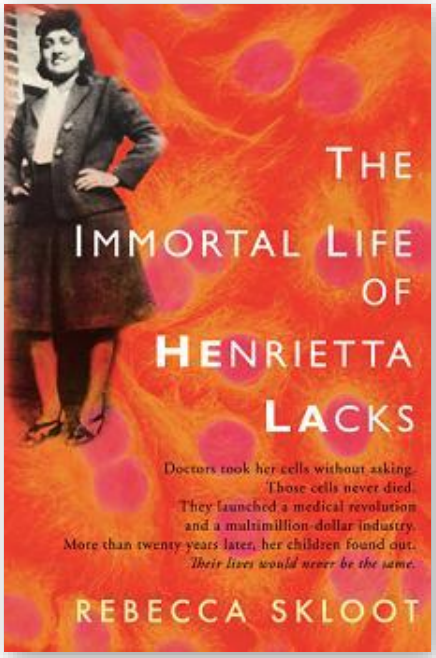
Maintaining Scientific Integrity Through Cell Line Authentication

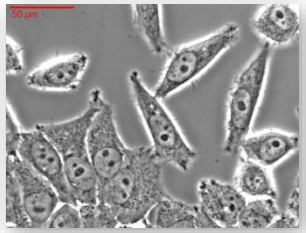
Gabriela Saldanha
Genomics Business Unit
Life Sciences Research

Presentation Outline

- Background on cell line authentication (CLA)
- Basics of STRs and CLA protocols

The legend of HeLa





HeLa cells

- HeLa- first immortal cancer line to be established in 1951 by George Gey
- Cervical cancer derived from patient Henrietta Lacks
- Widely used in research

  According to Rebecca Skloot, "More than 60,000 scientific articles had been published about research done on HeLa, and that number was increasing steadily at a rate of more than 300 papers each month."

- Rapid growing; can contaminate and overtake other cell types

Stanley Gartler describes HeLa contamination of cell lines

1968: Gartler publishes first evidence of HeLa contamination based on G6PD isoenzymes

NAME	DESCRIPTION	ORIGIN	G6PD VARIANT
HeLa	Cervical adenocarcinoma, human	African	Type A (fast)
KB	Oral epidermoid carcinoma, human	Caucasian	Type A (fast)
HEp-2	Larynx epidermoid carcinoma, human	Caucasian	Type A (fast)
Chang liver	Liver, human	Caucasian	Type A (fast)
Int-407	Embryonic intestine, human	Caucasian	Type A (fast)

Nature Reviews, 2010

Conclusion: 90% (18/20) human cell lines are 'HeLa'

Poor culture condition leads to contamination

- Poor tissue culture environment
- No disposable, plastic culture dishes
- No commercial media
- Cells grown on bench-top
- Bunsen burners and steam used for sterilization
- Technicians wore surgical masks, coats, gloves, booties, hair covers

How does false cell line identity occur?

- Cross-contamination during original cell line development
- Cross-contamination in the lab from which the cell line was acquired
- Cross-contamination in your lab
- Mislabeling

1/3

OF RESEARCHES

obtain their cell lines from another lab

Good cell culture practice

BJC

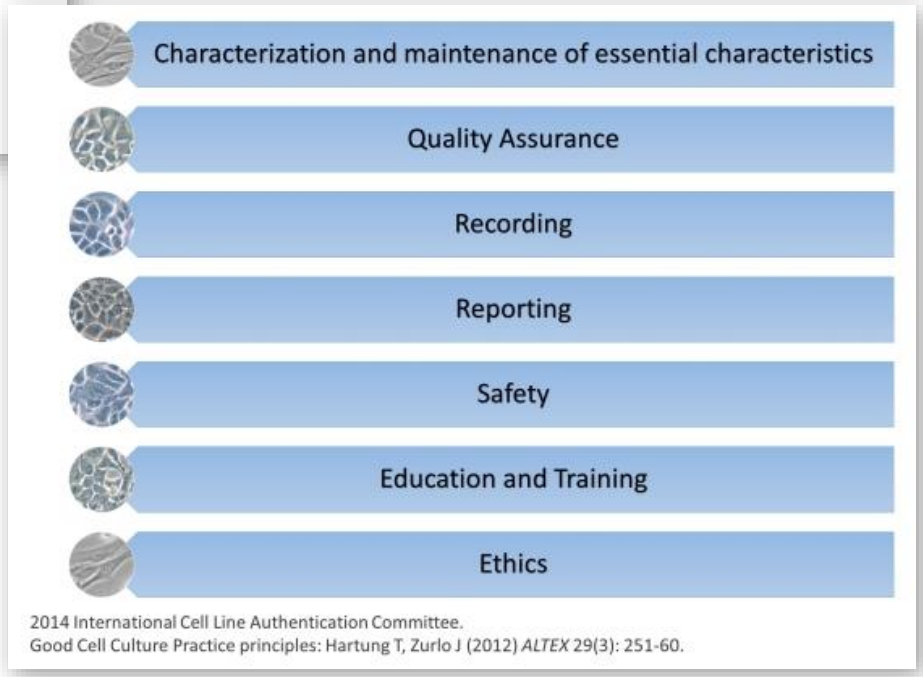
GUIDELINE

British Journal of Cancer (2014) 111, 1021–1046 | doi: 10.1038/bjc.2014.166

Keywords: cell culture; mycoplasma contamination; Human Tissue Act; cell line; cell line misidentification; cryostorage; Human Tissue Authority; STR profiling; human tissue; Human Fertilisation and Embryology Act

Guidelines for the use of cell lines in biomedical research

R J Geraghty*[,1], A Capes-Davis[2], J M Davis[3], J Downward[4], R I Freshney[5], I Knezevic[6], R Lovell-Badge[7], J R W Masters[8], J Meredith[9], G N Stacey[10], P Thraves[11] and M Vias[1]



2014 International Cell Line Authentication Committee.
Good Cell Culture Practice principles: Hartung T, Zurlo J (2012) *ALTEX* 29(3): 251-60.

Misidentification of cell lines persists…

Year	Title of Article	Reference
2004	LCC15-MB cells are MDA-MB-435: a review of misidentified breast and prostate cell lines	Clinical & Experimental Metastasis 21 (6): 535, 2004
2007	MDA-MB-435: the questionable use of a melanoma cell line as a model for human breast cancer is ongoing	Cancer Biology & Therapy 6(9): 2007
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	Journal: Cancer Letters, 245 (1-2): 350, 2007
2008	Persistent use of “false” cell lines	Journal: International Journal of Cancer 122 (1): 1, 2008
2010	Verification and unmasking of widely used human esophageal adenocarcinoma cell lines	NCI J Natl Cancer Inst 102(4): 271, 2010
2011	MCF-7/ADR cells (re-designated NCI/ADR-RES) are not derived from MCF-7 breast cancer cells: a loss for breast cancer multidrug-resistant research	Journal: Medical Oncology 28 (1): 135, 2011
2013	Misidentification of putative medullary thyroid cancer cell lines RO-H85-1 and RO-D81-1	J Clin Endocrinol Metab, 98(3):954, 2013
2013	Beware imposters: MA-1, a novel MALT lymphoma cell line, is misidentified and corresponds to Pfeiffer, a diffuse large B-cell lymphoma cell line	Genes, Chromosomes and Cancer (10): 986,2013

Impact of misidentified cell lines on applied research

Experimental results based on contaminated cell lines …

- Clinical trail recruiting EAC patients
- 100 scientific publications
- At least 3 NIH cancer research grants
- 11 US patents

Consequences of using misidentified cell lines

- Loss of cell lines
- Loss of time and money
- Misinformation in the public domain
- Discordant or irreproducible results
- Tarnished reputation

"If we're not using what we think we are using, we're not testing our hypotheses. We're just gumming up the literature. I'm not sure what we're doing, but that's not science."

Jeffrey Boatright, Emory University, The Big Clean-Up. The Scientist Magazine. September 2015

New ANSI/ATCC Standards Development Organization Guidelines

ASN-0002 consensus standard describing the use of **short tandem repeat (STR) analysis** for cell line authentication

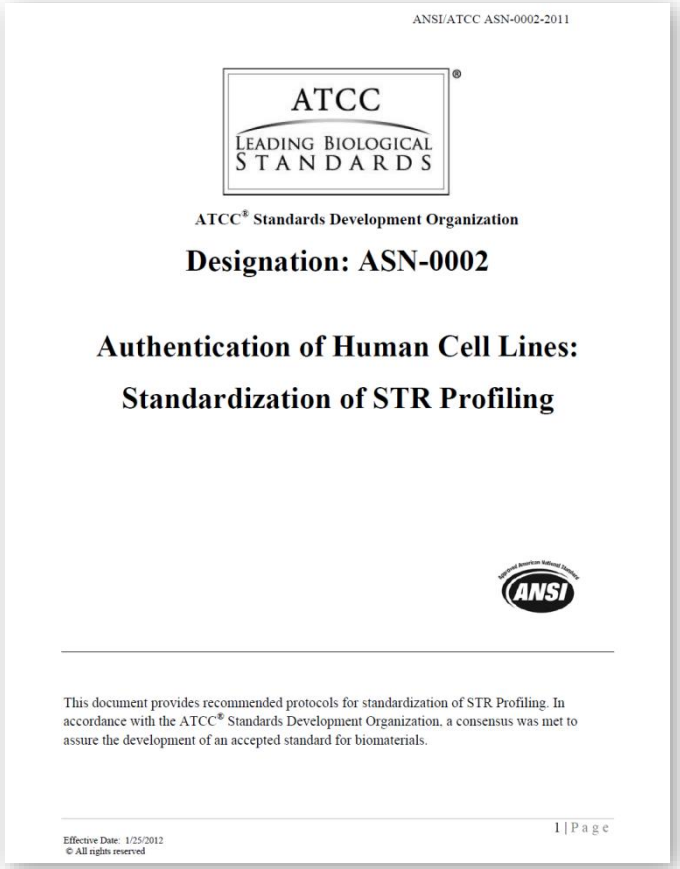
With STR profiling authentication may include:

(i) verifying the cells are of human origin

(ii) evaluating the consistency of profiles between isolates/passages

(iii) comparing STR profile to a database

(iv) detecting contaminating human DNA


ANSI/ATCC ASN-0002-2011

ATCC
LEADING BIOLOGICAL STANDARDS

ATCC® Standards Development Organization

Designation: ASN-0002

Authentication of Human Cell Lines: Standardization of STR Profiling

ANSI

This document provides recommended protocols for standardization of STR Profiling. In accordance with the ATCC® Standards Development Organization, a consensus was met to assure the development of an accepted standard for biomaterials.

1 | Page

Effective Date: 1/25/2012

Journals taking action

nature International weekly journal of science

Home | News & Comment | Research | Careers & Jobs | Current Issue | Archive | Audio & Video

Archive > Volume 520 > Issue 7547 > Editorial > Article

NATURE | EDITORIAL

Announcement: Time to tackle cells' mistaken identity

15 April 2015

PDF | Rights & Permissions

The differences between a cow and a monkey are clear. It is easy to tell a moth from a mosquito. So why are there still scientific studies that mix them up? The answer is simple: hundreds of cell lines stored and used by modern laboratories have been wrongly identified. Some pig cells are labelled as coming from a chicken; cell lines advertised as human have been shown to contain material from hamsters, rats, mice and monkeys.

Which is worse: that such crude mix-ups exist, or that, every day, researchers use cell lines that somebody, somewhere has already found to be mislabelled, misidentified or contaminated? To solve the first problem is a huge challenge. To address the second is a more manageable task, and one that researchers, journals, universities and funders must take seriously.

Nature and the Nature research journals are strengthening their policies to improve the situation. From next month, we will ask authors to check that they are not working on cells known to have been misidentified or cross-contaminated, and will ask them to provide more details about the source and testing of their cell lines.

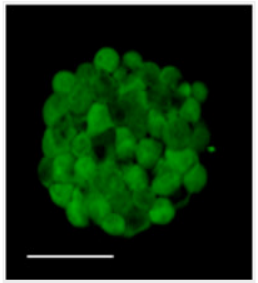
NATURE METHODS | METHAGORA

A retraction resulting from cell line contamination

09 Sep 2013 | 2:09 AM | Posted by Daniel Evanko | Category: General Interest, Journal Policy, Nature Methods papers

After nine years in print, *Nature Methods* today published its first retraction; one that could have been prevented by cell line authentication. What does this mean for journal-mandated cell line testing?

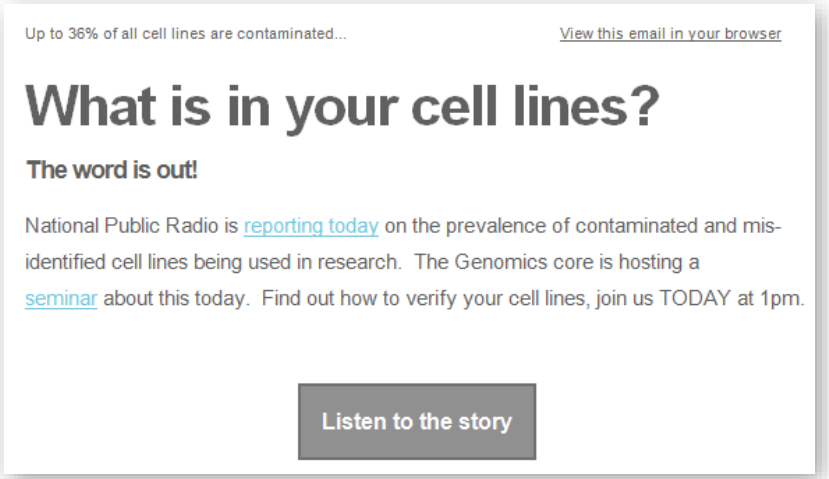
In a *Nature Methods* paper published in 2010, Ivan Radovanovic and colleagues described a method to isolate cancer-initiating cells in human glioma without the need for molecular markers. Based on morphology and on a green autofluorescence, the authors reported they could use FACS to sort cancer-initiating cells from gliomasphere cultures (which had been derived from primary tumors). They also detected autofluorescence in cells from fresh glioma specimens, but at a much lower level.

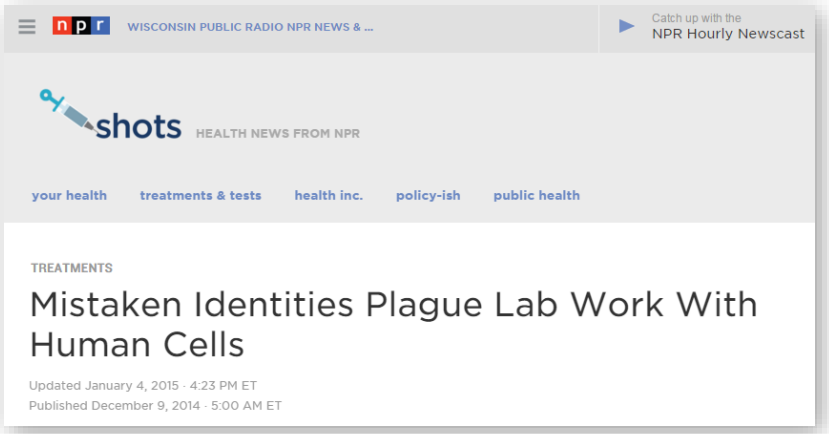


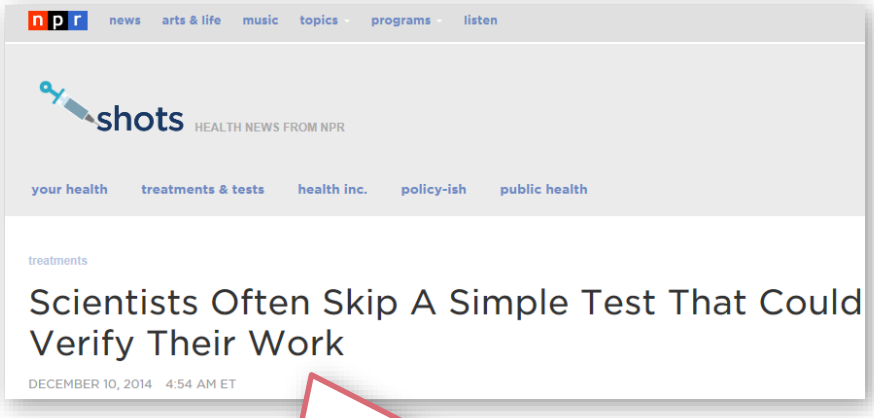
Two-photon fluorescence image of live primary gliomasphere from retracted manuscript.

Cells from the autofluorescent fraction could self renew clonogenically in vitro and were tumorigenic when transplanted into mouse brains, the authors reported, and in both cases performed better than non-autofluorescent cells from the rest of the culture or tissue. The origin of this autofluorescent signal was not understood at the time. The authors speculated it may have been related to the unique metabolism of the cancer-initiating cells.

Media outlets are paying attention







"There's a simple test that scientists could use to make sure the cells they're studying in the lab are what they think they are. But most of the time, academic scientists don't bother.

The culture of cell culture practices and authentication—Results from a 2015 survey

Concerns continue to grow regarding irreproducible basic biological and preclinical research.

Accurate documentation of cell line tissue of origin (i.e., identity), sex, and species are critical to ensure the credibility, reproducibility, and translation of data and results from cell culture based experiments.

The financial implications of misidentified or contaminated cell lines can be profound; as much as $700 million dollars per year in research that could be at risk.

GBSI
Global Biological Standards Institute®

…Results continued

18-36%

OF CELL LINES

are misidentified

52% "never" perform authentication or quality control

74% Never conduct STR DNA profiling

61% Cost

53% Time

35% Delay in research

24% Reported "I don't see the necessity; I am careful"

22% laboratory manager or principal investigator is "unaware of or ignores the issue"

Cell line authentication to improve reproducibility in cancer research

NIH revised guidelines for funding – Enhancing Reproducibility through Rigor and Transparence (effective Jan 25, 2016)

- Authentication of key biological and/or chemical resources

"*NIH plans to enhance reproducibility in multiple ways, by promoting greater scientific rigor and transparency in funding applications and publications, encouraging robust peer review, providing adequate training, removing perverse incentives, and emphasizing overlooked areas such as the consideration of both sexes in research* ***and the use of authenticated cell lines***."

Cell line authentication to improve reproducibility in cancer research

PROSTATE CANCER FOUNDATION
Accelerating the world's most promising research

Prostate Cancer Foundation
Cell Line Authentication Initiative

Perform genetic and pathogenic tests for cell lines for future funding

Duke Translational Medicine Institute

Duke University
Free Human Cell Line Authentication Assays
Free of charge, a human cell line authentication service to Duke University School of Medicine investigators.

THE UNIVERSITY OF TEXAS
MD Anderson Cancer Center
Making Cancer History®

UTMDACC INSTITUTIONAL POLICY # ACA1044

CELL LINE AUTHENTICATION POLICY

PURPOSE

The contamination or misidentification of cell cultures can seriously compromise research performed with such cell lines. The purpose of this policy is to ensure that information generated using cultured cells at The University of MD Anderson Cancer Center (MD Anderson) is obtained from samples that have been authenticated by DNA validation techniques, to rule out misidentification and inter- or intra-species contamination. **This will prevent the necessity for retraction of papers in which cell lines are found not to be of the reported lineage. It has been indicated by NIH that grant applications that fail to use acceptable experimental practices "would not fare well in the review process."**

POLICY STATEMENT

Validation by DNA analysis will be performed on cell lines used to produce scientific information at MD Anderson. This validation process **will enhance the quality of research at MD Anderson and maintain the scientific integrity and reputation of the institution.**

SCOPE

This policy **applies to all workforce members utilizing cell lines for research**.

Champions who are dedicated to advancing the #Authenticate goals

Global Biological Standards Institute®

ATCC®

CellBank Australia

CHILDREN'S MEDICAL RESEARCH INSTITUTE

CHORDOMA FOUNDATION

FasterCures
A CENTER OF THE MILKEN INSTITUTE

ICLAC

susan g. komen.

Melanoma Research Alliance

DDC MEDICAL

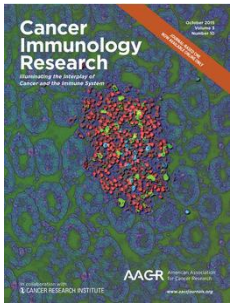
RCRF
THE RARE CANCER RESEARCH FOUNDATION

PROSTATE CANCER FOUNDATION

IDEXX BioResearch

Which journals ask for cell line authentication?*

- AACR journals:
  - Cancer Discovery
  - Cancer Research
  - Clinical Cancer Research
  - Cancer Epidemiology, Biomarkers & Prevention
  - Molecular Cancer Research
  - Molecular Cancer Therapeutics
  - Cancer Prevention Research
- Carcinogenesis
- Cell Biochemistry and Biophysics
- Cell Biology International
- Endocrine Society journals:
  - Endocrinology
  - Endocrine Reviews
  - Journal of Clinical Endocrinology & Metabolism
  - Molecular Endocrinology
  - Hormones and Cancer
- International Journal of Cancer
- In Vitro Cellular & Developmental Biology - Animal



- Journal of Molecular Biology
- Journal of the National Cancer Institute
- Molecular Vision
- Nature Publishing Group:
  - Nature
  - Nature Reviews Molecular Cell Biology
  - Nature Genetics
  - Nature Reviews Immunology
  - Nature Reviews Cancer
  - Nature Reviews Neuroscience
  - Nature Biotechnology
  - Nature Methods
- Neuro-Oncology
- Placenta
- PLoS ONE
- Society for Endocrinology journals:
  - Journal of Endocrinology
  - Journal of Molecular Endocrinology
  - Endocrine-Related Cancer

*Requirements vary with each journal

How researchers can prevent the use of misidentified cell lines?

ONLY 1%

OF AUTHORS

authenticate their cell lines

- Become aware of the magnitude of the problem
- Authenticate all cell lines you are working with!
- Eliminate laboratory practices that result in misidentification
- Notify recipients of misidentified cell lines
- Publish retractions/explanations of “false” publications

RESULTS MATTER

- Confidence in the published data
- Meet the requirements of peer reviewed journals and grant submission
- Aid in troubleshooting undetermined or unexpected results

Summary

- ✓ Cell authentication is more than just identity
- ✓ It consists of a number of orthogonal endpoints (identity, morphology, phenotype, ploidy, purity)
- ✓ Look for increasing demand for authentication of cells used in research and biological activity assays

Credit: Chris Nickels for NPR

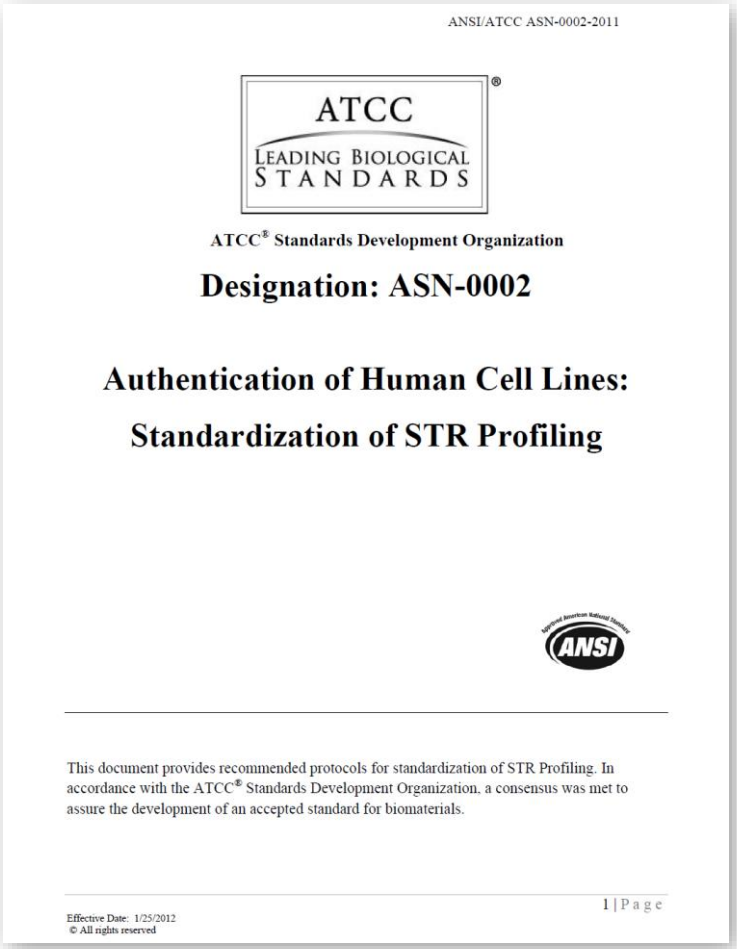
STR Profiling: Authentication Standard for Human Cell Lines

Promega Corporation

New ANSI/ATCC Standards Development Organization Guidelines

- ASN-0002 consensus standard describing the use of **short tandem repeat (STR) analysis** for cell line authentication
  - Also recommend assessing cell behavior, such as doubling time and morphology
- Plans to create a database containing STR genotypes of validated cell lines
  - Hosted by the National Center for Biotechnology Information (NCBI)
- Public database available on ATCC website

ANSI: American National Standards Institute



ANSI/ATCC ASN-0002-2011

ATCC
LEADING BIOLOGICAL STANDARDS

ATCC® Standards Development Organization

Designation: ASN-0002

Authentication of Human Cell Lines:

Standardization of STR Profiling

ANSI

This document provides recommended protocols for standardization of STR Profiling. In accordance with the ATCC® Standards Development Organization, a consensus was met to assure the development of an accepted standard for biomaterials.

1 | Page

Effective Date: 1/25/2012

ANSI authentication points

With STR profiling authentication may include:

(i) verifying the cells are of human origin

(ii) evaluating the consistency of profiles between isolates/passages

(iii) comparing STR profile to a database

(iv) detecting contaminating human DNA*

*Note that commercial STR kits are designed to detect human DNA, so other species' contaminating DNA will not be detected.

STR analysis workflow
Multiplex PCR through data analysis

DNA Sample

Multiplex PCR

- Amplification of select STR loci
- Simultaneous fluorescent labeling

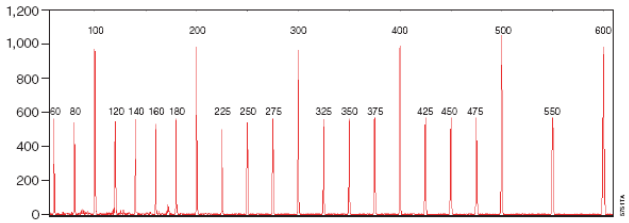
Capillary Electrophoresis (CE)

- Add Internal Lane Standard
- CE to size separate
- Fluorescent detection
- Also run Allelic Ladder in parallel

Data Analysis

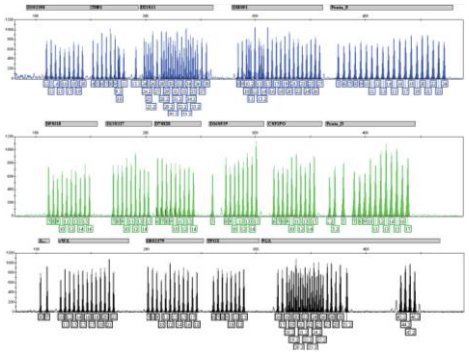
- Calculate sizes based on ILS
- Compare fragment sizes to allelic ladder to determine STR alleles present in sample
- Compare to databases

Internal Lane Standard



Labeled size standards in a different color than STR fragments

Allelic Ladder



Labeled fragments of all possible alleles for each STR locus

STR (Short Tandem Repeat) genotyping

STR Examples

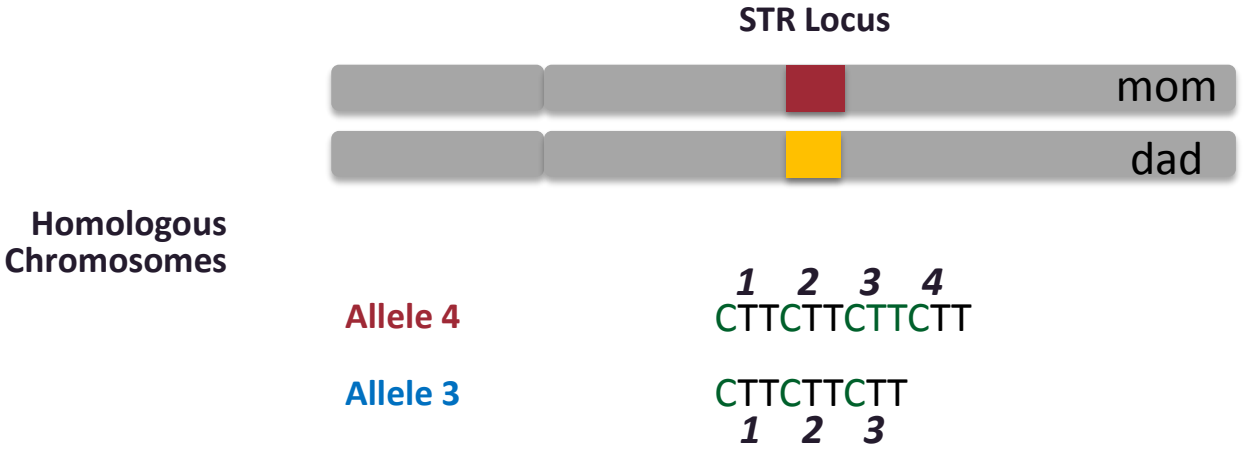
Trinucleotide: ---CTTCTTCTTCTTCTT---

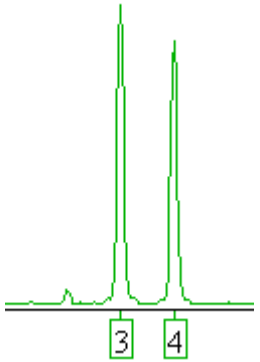
Tetranucleotide: ----AATGAATGAATG----

Pentanucleotide: -----AAAGAAAAGAAAAGAAAAGA-----

STR = Short Tandem Repeats: Multiple copies of a short (2-6bp), identical DNA sequence arranged in direct succession within a chromosome.

STR Variation in an Individual

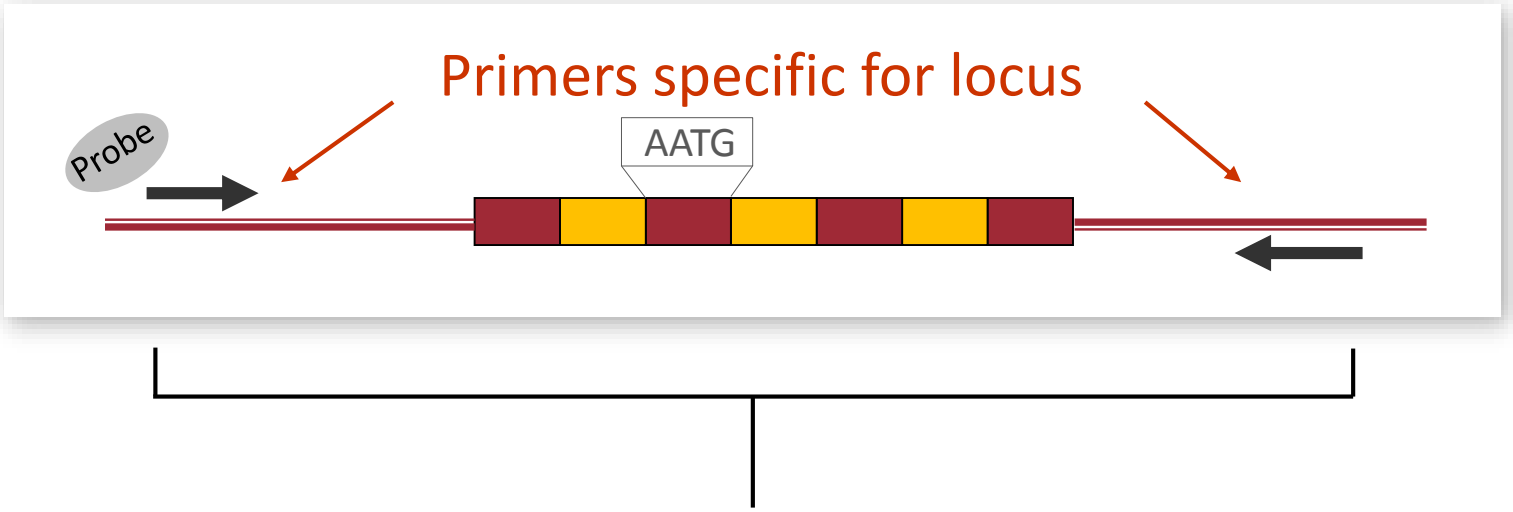




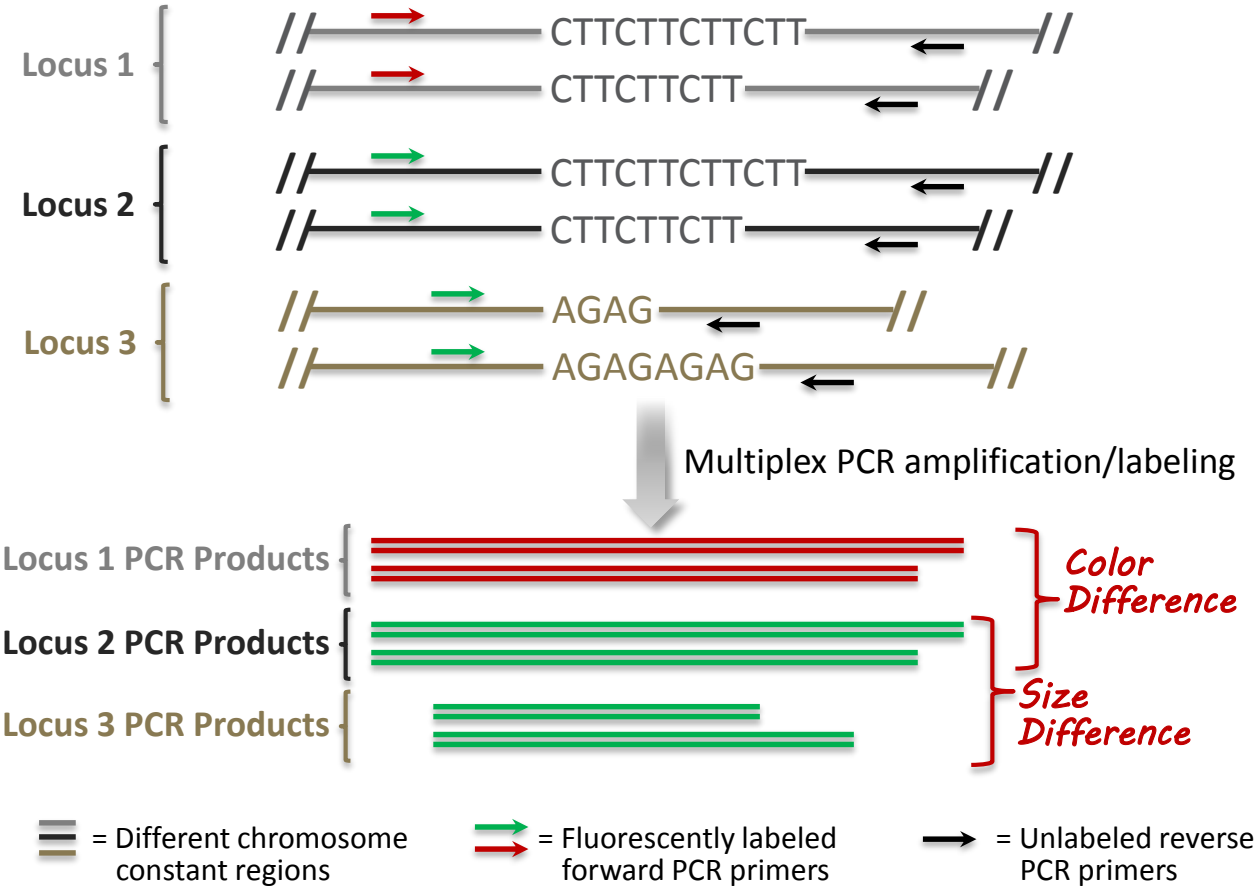
Amplifying STRs

DNA is extracted from source material

Extracted DNA is the template for PCR



Multiplex STR analysis combines size and color differences to analyze multiple loci



Separating and detecting amplified STRs

Capillary Electrophoresis

Separates amplification products based on size using Applied Biosystems' Genetic Analyzers

3130

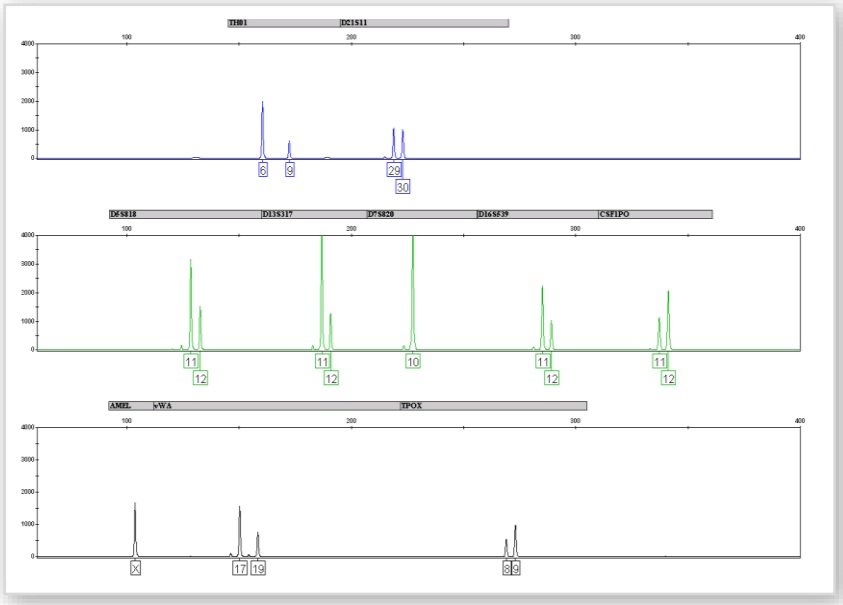
3500

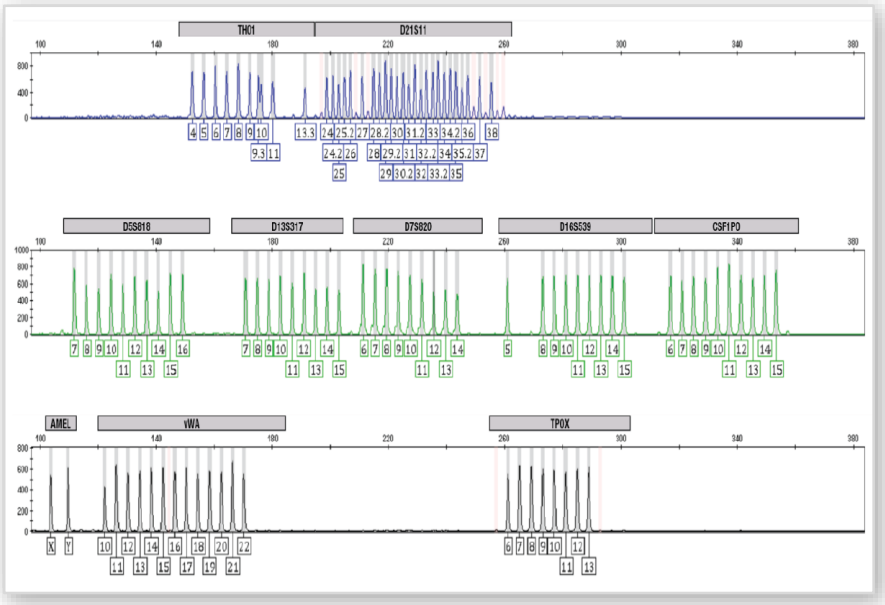
3730

Assigning allele types to peaks

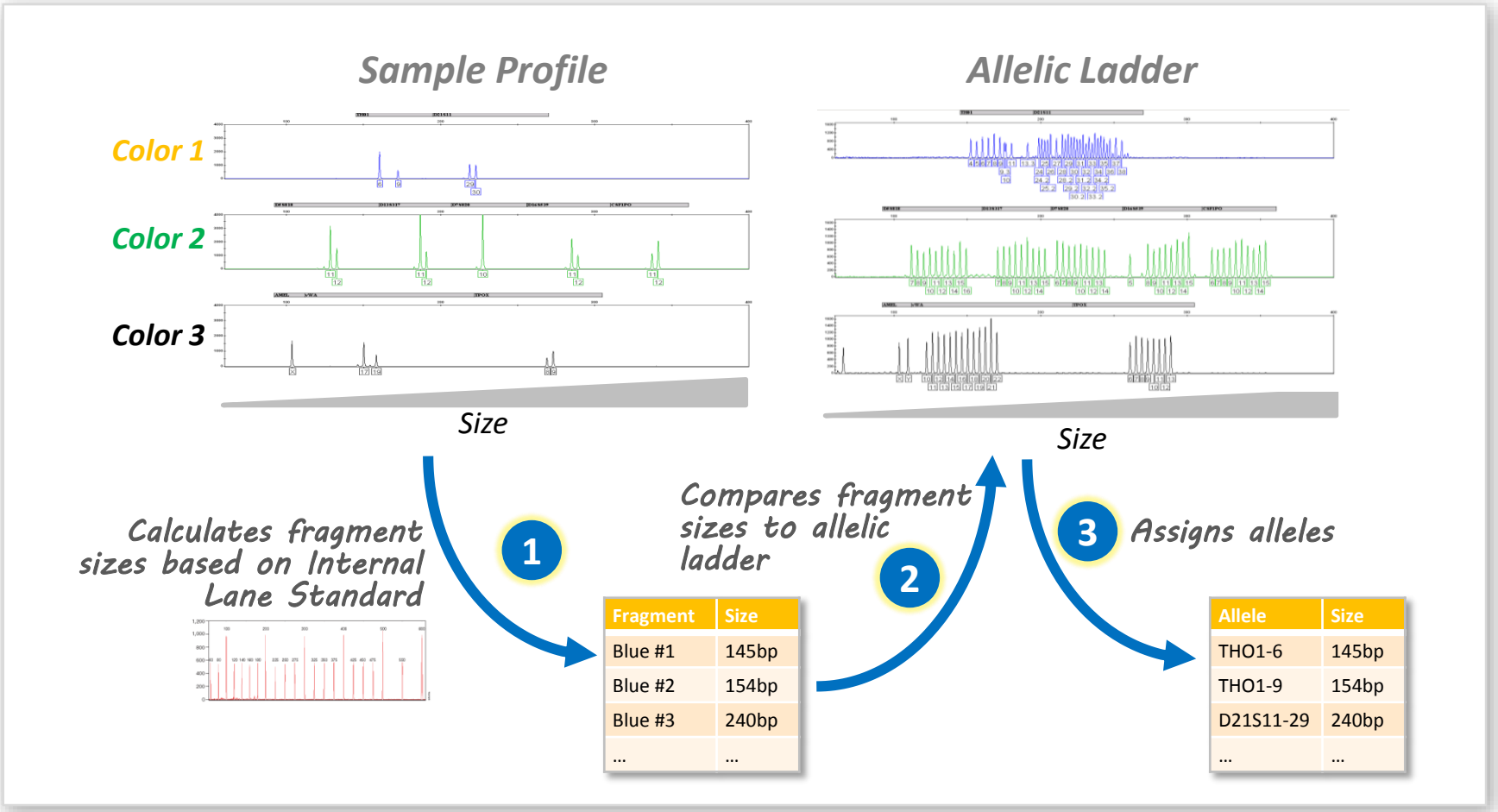
Samples are compared to an allelic ladder to determine actual alleles

- Nomenclature is based on number of repeats
- Output is a simple text-based table

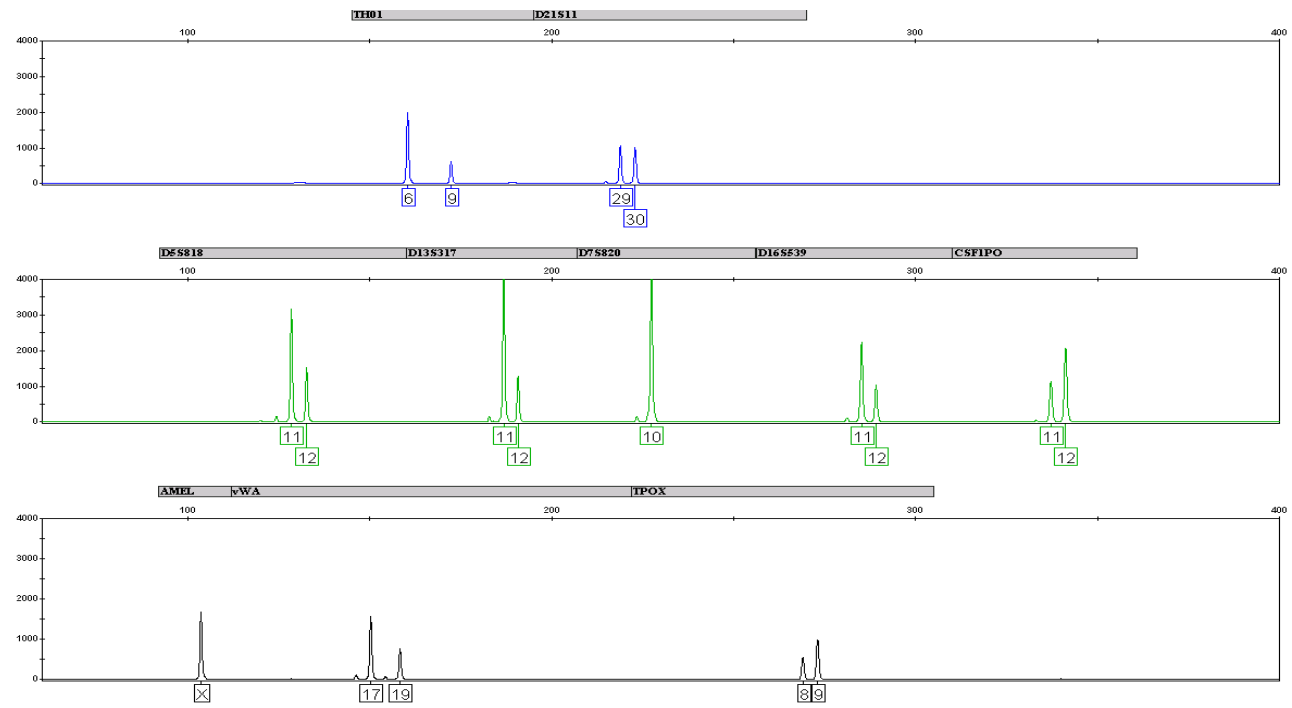




Data analysis software identifies STR alleles



Direct amplification profile from FTA® Paper
HT-29 cell line with GenePrint® 10 System

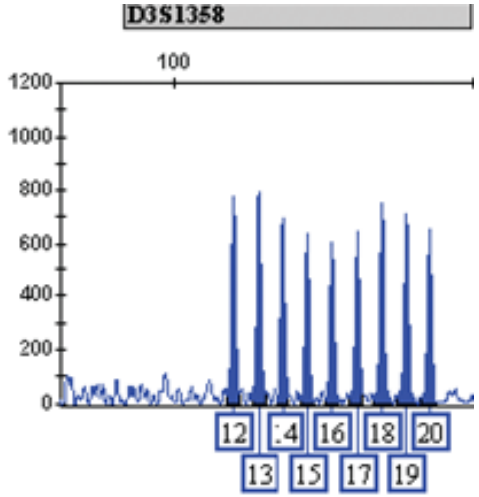


Simple text-based table

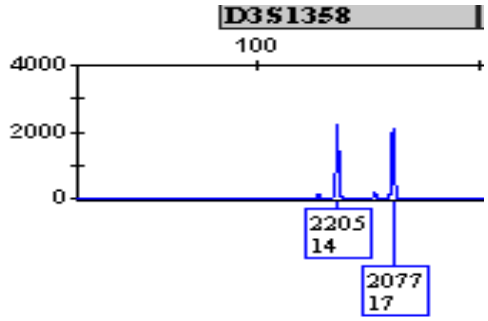
STR Locus	Alleles
TH01	6, 9
D21S11	29, 30
D5S818	11, 12
D13S317	11, 12
D7S820	10, 10
D16S539	11, 12
CSF1PO	11, 12
Amelogenin	X, X
vWA	17, 19
TPOX	8, 9

Potential versus actual alleles at 1 locus

Alleles commonly encountered in the human population at this location

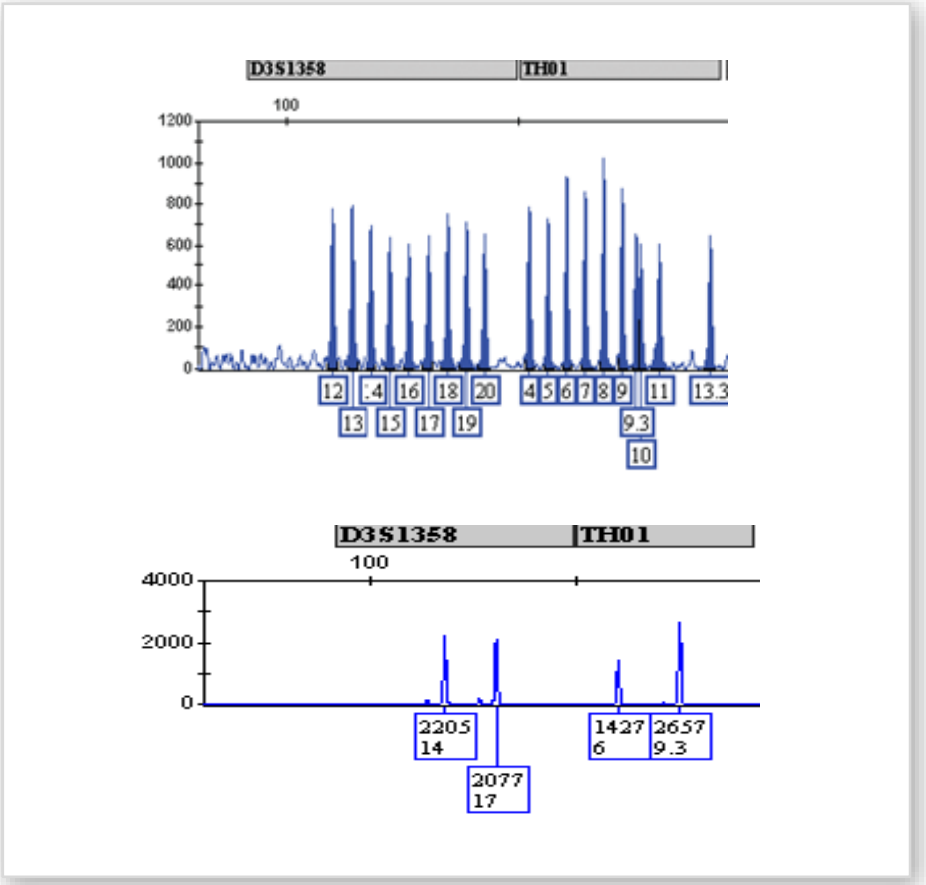


Profile generated from an individual



1 of 9 from mom
+ 1 of 9 from dad
1 of 18 possibilities

…using 2 loci



1 of 9 from mom

+ 1 of 9 from dad

1 of 18 possibilities

at first location (D3S1358)

x

1 of 10 from mom

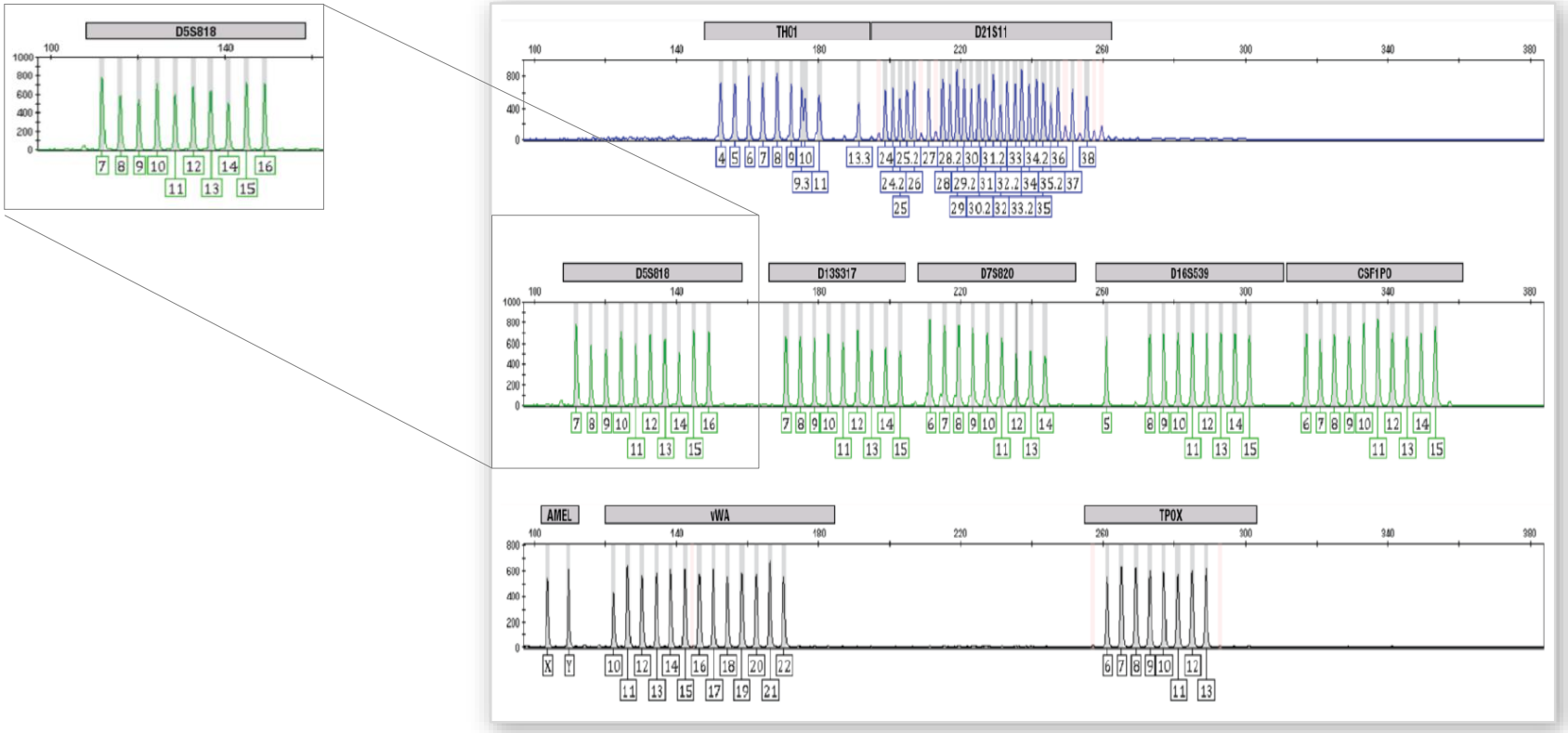
+ 1 of 10 from dad

1 of 20 possibilities

at second location (TH01)

=1:18 x 1:20 = 1:360

…using 10 loci



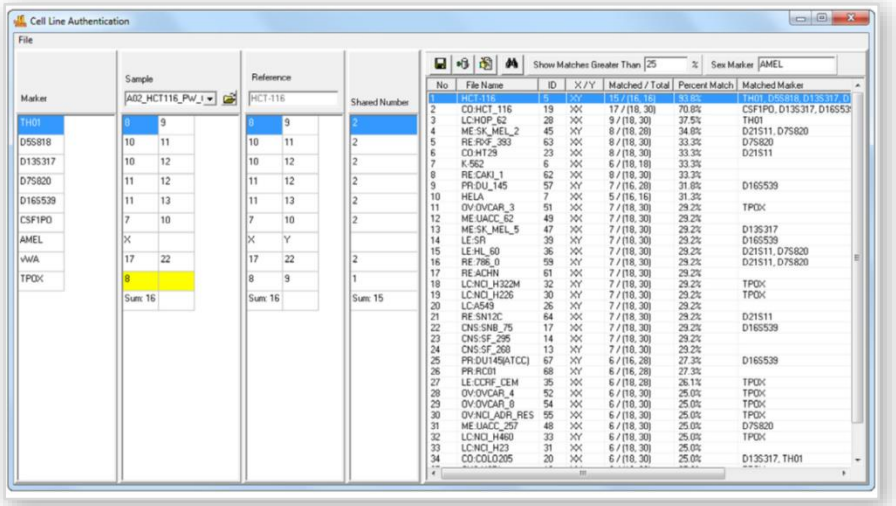
Matching probability: 1 in 18 x 1 in 20 x 1 in 50…

1 in 3 x 10^9 = 1 in 3,000,000,000 (*GenePrint*® 10)

Data analysis software options

GeneMarker® HID
Human Identity Software

By SoftGenetics



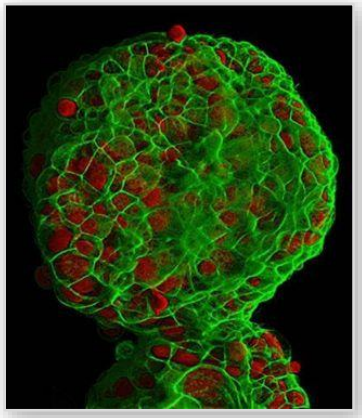
GeneMapper® Software
GeneMapper® ID-X Software

By Life Technologies

There are several versions. Some instrument files are not compatible with some versions. Additional files may be needed to perform full genotyping analysis depending upon the version used.

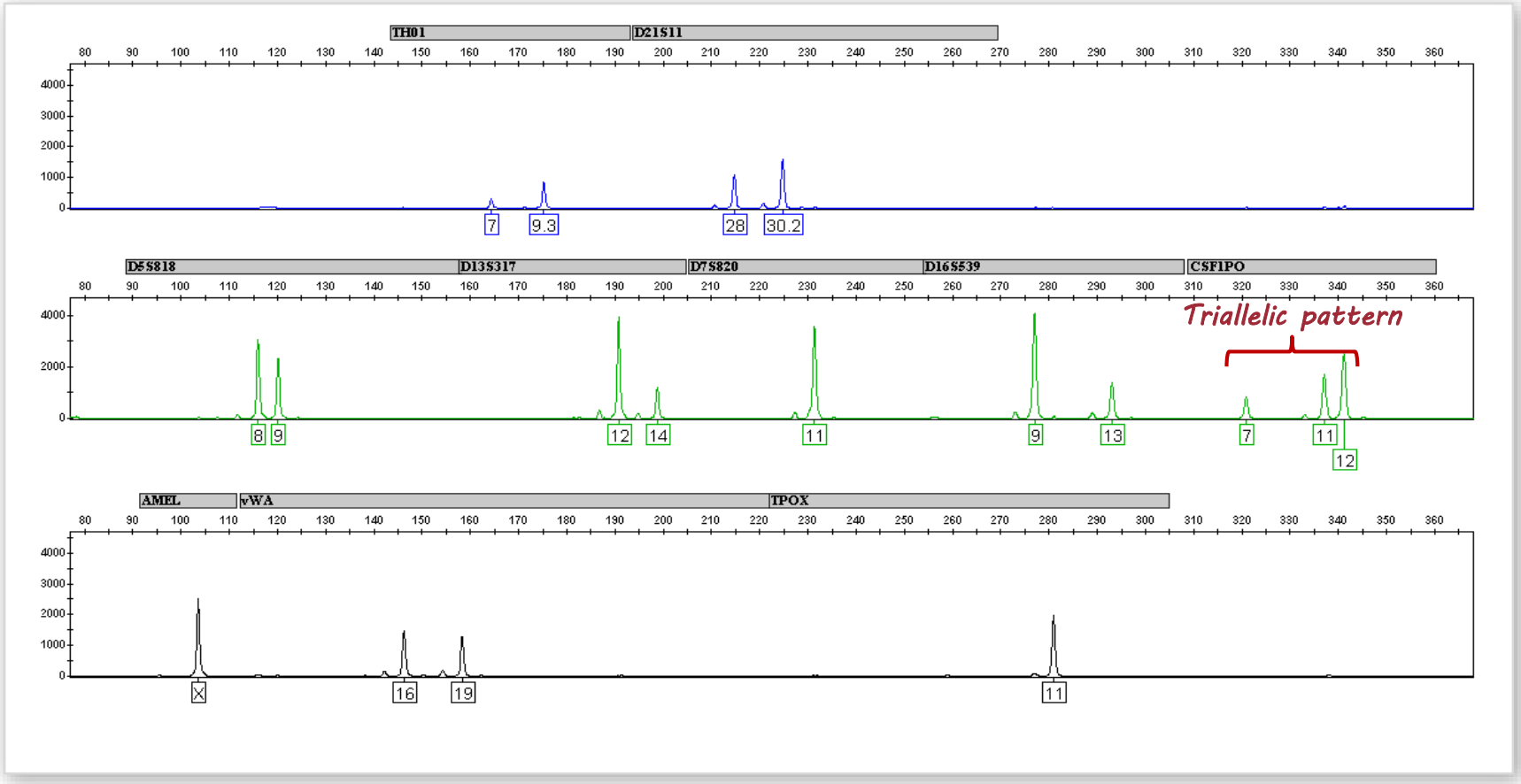
Strengths and limitations

- STR systems can determine…
  - Relatedness of a cell line to a reference standard
  - Cell line cross-contamination
  - Gender (sex)
- STR systems cannot distinguish…
  - Different cell lines created from the same individual (donor)
  - Cell lines created from identical twins
- Promega STR Systems are designed to be human-specific
  - Cannot be used to genotype non-human species

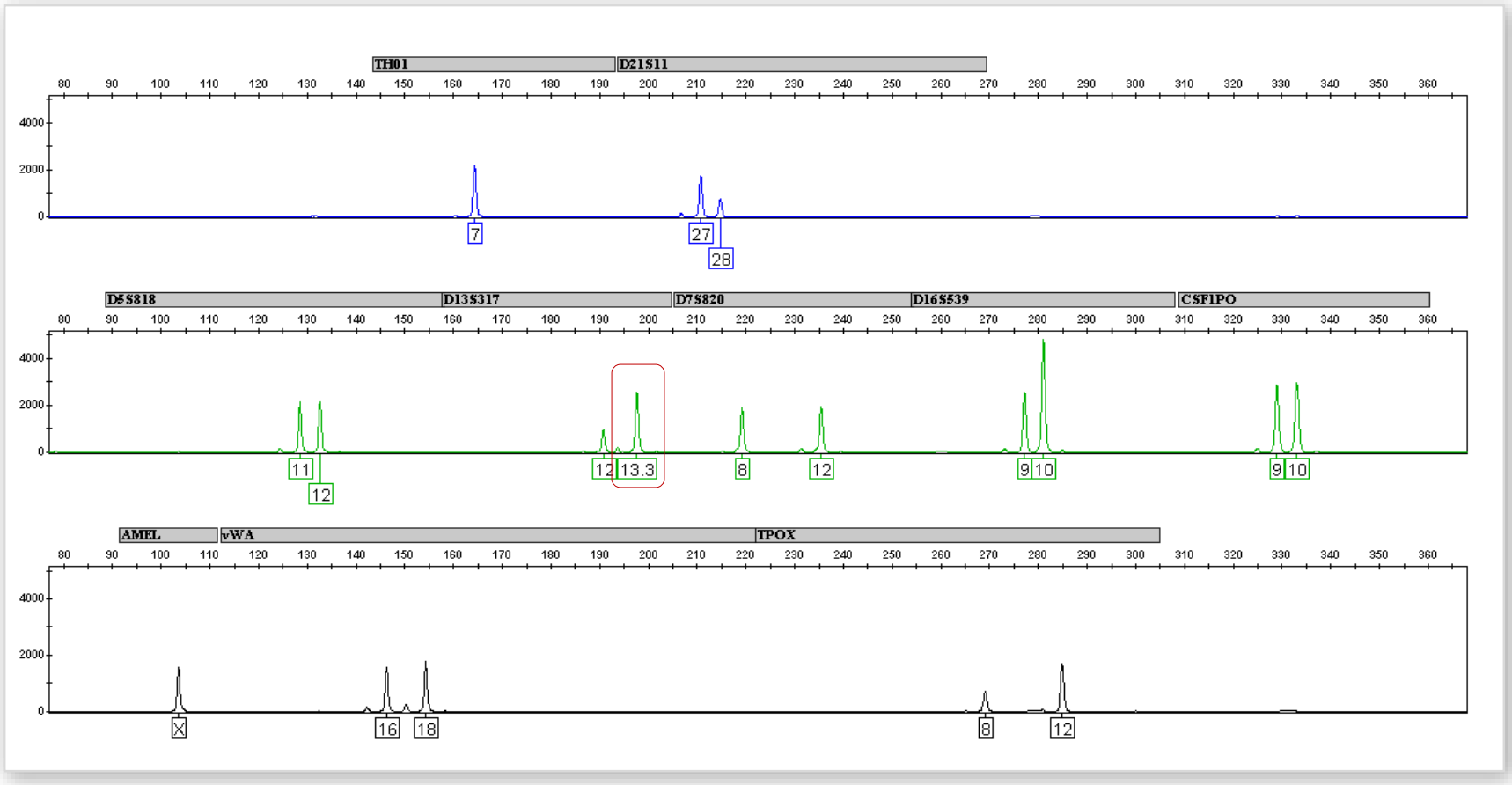


3D cell culture of breast cancer cells (MCF-7 cell line)

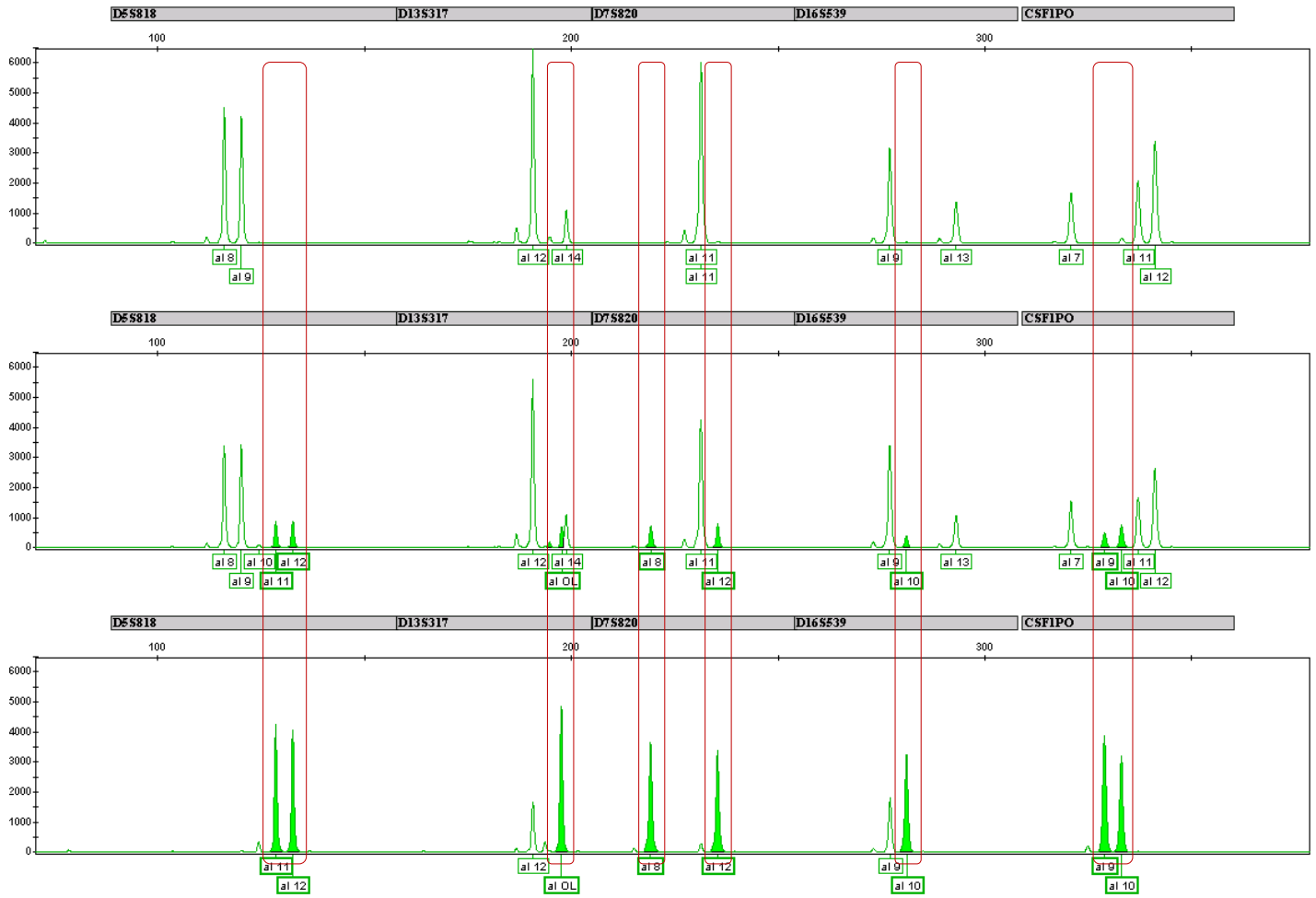
HEK293 STR profile demonstrates triallelic pattern often found in immortalized cell lines



STR profile illustrates HeLa characteristic D13 locus 13.3 allele



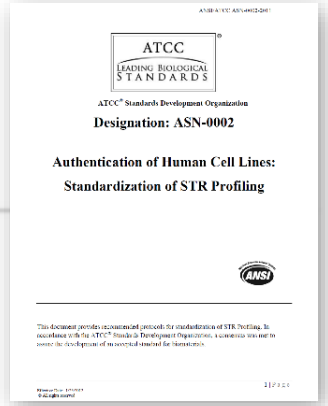
Detection of contaminating HeLa cells in HEK293 culture using GenePrint® 10 System



Sample comparison

Test Results for Submitted Sample				ATCC Profile or Reference Profile		
Loci	Query Profile			Database (or Reference) Profile: A549 (CCL-185)		
D3S1358	16					
THO1	8	9.3		8	9.3	
D21S11	29					
D18S51	14	17				
Penta_E	7	11				
D5S818	11			11		
D13S317	11			11		
D7S820	8	11		8	11	
D16S539	11	12		11	12	
CSF1PO	10	12		10	12	
Penta_D	9					
Amelogenin	X	Y		X	Y	
vWA	14			14		
D8S1179	13	14				
TPOX	8	11		8	11	
FGA	23					
D19S433	13					
D2S1338	24					
Number shared alleles between reference and test profile:						15
Total number of alleles in the database/reference profile:						15
Percent match between the submitted sample and the database/reference profile: =(Number shared)/(Total number in reference)*100						100

Matching algorithm


ATCC
LEADING BIOLOGICAL STANDARDS

ATCC® Standards Development Organization

Designation: ASN-0002

Authentication of Human Cell Lines:
Standardization of STR Profiling

1. Combine <u>total</u> number of alleles observed in the Test Sample and Reference (TOTAL ALLELES)
2. Count the number of alleles shared by the Test Sample and the Reference Sample (SHARED ALLELES)
3. Use the Match Algorithm to calculate a percent match result for the two samples:

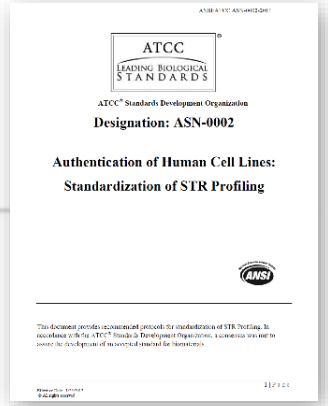
$$\text{Percent Match} = \frac{\text{SHARED ALLELES} \times 2}{\text{TOTAL ALLELES}}$$

Related samples generally yield a result in the 80-100% match range

- A percent match <80% should be investigated

International Cell Line Authentication Committee
http://standards.atcc.org/kwspub/home/the_international_cell_line_authentication_committee-iclac_/

Match criteria


ATCC
LEADING BIOLOGICAL STANDARDS
ATCC® Standards Development Organization
Designation: ASN-0002
Authentication of Human Cell Lines:
Standardization of STR Profiling

Authentic cell line: a match at ≥80% of alleles across the eight (8) core STR loci.

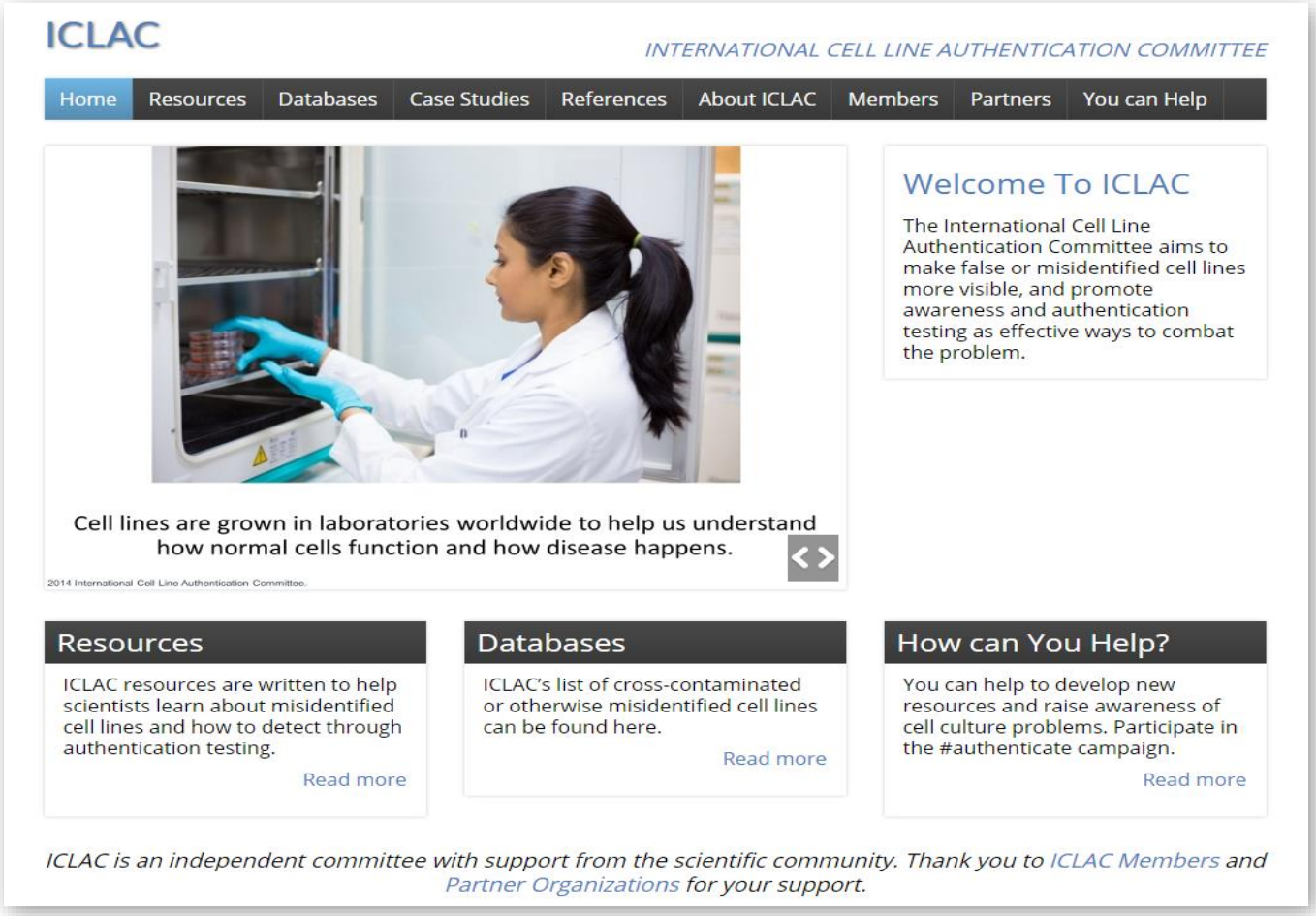
Match: when two STR profiles show identical alleles. This is described as a percentage as shown above. Cell line samples matching at ≥80% of alleles across the eight (8) core loci are said to be related.

Unrelated STR profiles: when STR profiles match at <55% of alleles. STR profiles with alleles matching at 55-80% may be related and require further investigation.

Cell line authentication is growing in importance

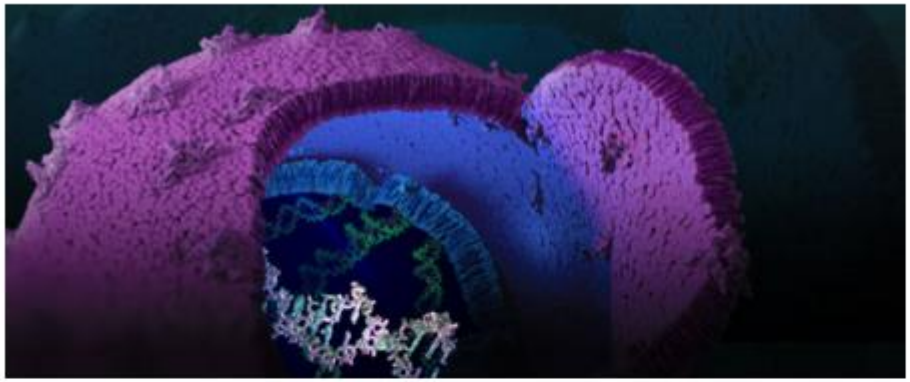
- ✓ **The number of misidentified/contaminated cells continues to grow**
  - Invalidates published and unpublished data leading to wasted time, effort and money
- ✓ **Granting agencies and journals are strongly recommending cell authentication**
  - ANS-0002 standards for cell authentication published by ANSI/ATCC
  - Check journal *Instructions to Authors* for requirements/recommendations
- ✓ **STR profiling is the gold standard for cell authentication**
  - Genotyping method that amplifies highly variable short tandem repeat loci to create a DNA fingerprint of a sample
  - Also important to observe phenotype of cells to detect possible changes
- ✓ **GenePrint® 10 System is optimized for human cell authentication**
  - Contains all loci recommended in ANS-0002 standards and available in public databases
  - Contains one additional loci for greater discriminatory power
  - Easy-to-use, complete kit compatible with multiple PCR & CE instruments

International Cell Line Authentication Committee



http://iclac.org/

Cell Line Authentication Testing



Cell Line Authentication Matters | How Authentication Works | Data Interpretation | Service Providers | Journals Requiring Authentication

Human cultured cell lines are used in a number of biomedical research and clinical applications, including cancer research, drug discovery, genetics and biobanking. However, misidentified human and animal cell lines have continued to be used even today despite multiple and repeated warnings, articles and letters by prominent scientists in the field calling for authentication.

The need is great for researchers to authenticate their human cell lines. The call for authentication has been made for over five decades now, yet still, many researches do not perform this vital quality assessment. Recently there have been more spotlights shining on this deficiency from National Public Radio (NPR) series last December, comments from Dr. Francis Collins of the NIH, Nature announcing changes to abstract submission and most recently, newly revised NIH grant application instructions. As a result, more institutions are coming around to offering, even requiring, human cell line authentication through their DNA core labs or service providers.

Cell line authentication is achieved by genetic profiling using polymorphic short tandem repeat (STR) loci. STR loci consist of repetitive sequence elements 3–7 base pairs in length. These highly discriminative markers can be utilized using rapid and inexpensive multiplex-PCR based method for the identification and detection of contaminating human cells.

EXTERNAL RESOURCES

International Cell Line Authentication Committee (ICLAC)
- Guide to Human Cell Line Authentication
- Database of Misidentified Cell Lines
- Match Criteria Worksheet
- Cell Line Checklist for Manuscripts and Grant Applications

Authentication of Human Cell Lines by STR DNA Profiling Analysis

Maintaining High Standards in Cell Culture

DATABASES

ATCC STR (login required)

DSMZ STR (login required)

NCBI BioSample (open for sample submission)

RELATED RESOURCES

Short Tandem Repeat Analysis in the Research Laboratory

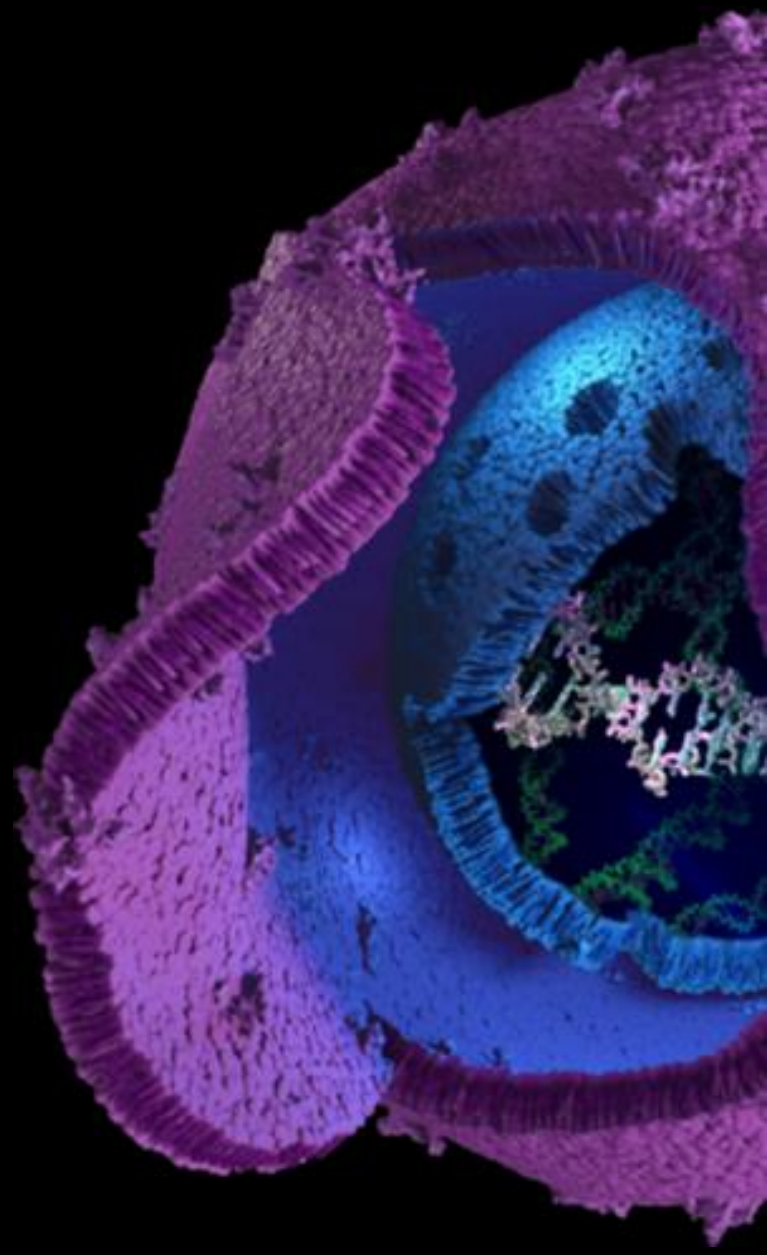
Doing Good Science: Authenticating Cell Line Identity

WEBINAR

Human Cell Line Misidentification

www.promega.com/cla

Maintaining scientific integrity through cell line authentication



Thank you