Mouse STR Profiling: A new addition to ATCC’s authentication portfolio

Balsam Shawky, M.S.
Senior Biologist, ATCC

Brian Shapiro, Ph.D.
Scientific Content Specialist, ATCC

Credible Leads to Incredible™
About ATCC

- Founded in 1925, ATCC is a non-profit organization with HQ in Manassas, VA, and an R&D and Services center in Gaithersburg, MD
- World’s largest, most diverse biological materials and information resource for microbes – the “gold standard”
- Innovative R&D company featuring gene editing, microbiome, NGS, advanced models
- World leader in cell line and microbe authentication
- Partner with government, industry, and academia
- Leading global supplier of authenticated cell lines, viral and microbial standards
- Sales and distribution in 150 countries, 18 international distributors
- Talented team of 450+ employees, over one-third with advanced degrees
Agenda: Cell line authentication for human and mouse cells

Introduction: Misidentification of cell lines

Cell line authentication

- Short tandem repeat (STR) profiling technique
- Mouse STR profiling

When and why should you authenticate

Summary
Consequences of using misidentified cell lines

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

“If we’re not using what we think we’re 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 1, 2015
## Misidentification of cell lines in the literature: 1994-2017

<table>
<thead>
<tr>
<th>Year</th>
<th>Title of article</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010</td>
<td>Verification and unmasking of widely used human <strong>esophageal adenocarcinoma</strong> cell lines.</td>
<td>JNCI 102(4):271, 2010.</td>
</tr>
<tr>
<td>2014</td>
<td>SNP Array profiling of <strong>mouse cell lines</strong> identifies their strains of origin and reveals cross-contamination and widespread aneuploidy</td>
<td>BMC Genomics 15:847, 2014.</td>
</tr>
<tr>
<td>2018</td>
<td>A comprehensive analysis of e-CAS cell line reveals they are <strong>mouse</strong> macrophages</td>
<td>Sci Rep 8(1):8237, 2018</td>
</tr>
</tbody>
</table>
Impact of misidentified cell lines on applied research

<table>
<thead>
<tr>
<th>Cell Line</th>
<th>Purported</th>
<th>STR confirmed (ATCC STR Profile database)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SEG-1</td>
<td>Esophageal adenocarcinoma cell line</td>
<td>H460 (ATCC® HTB-177™)                             Lung carcinoma (large cell lung cancer)</td>
</tr>
<tr>
<td>BIC-1</td>
<td>Esophageal adenocarcinoma cell line</td>
<td>SW620 (ATCC® CCL-227™)                             Colorectal adenocarcinoma</td>
</tr>
<tr>
<td>SK-GT-5</td>
<td>Esophageal adenocarcinoma cell line</td>
<td>SK-GT-2                                              Gastric fundus carcinoma</td>
</tr>
</tbody>
</table>

Experimental results based on contaminated cell lines…

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

Economic impact of misidentified cell lines

NIH Reporter for projects using “cell line” or “cell culture” – US $3.7B

- If 25% of research is conducted using misidentified cell lines then the loss could be $925 M
- If this number could be lowered to 10% then the cost is be reduced to $370 M

Short Tandem Repeat (STR) Profiling

A549 non-small cell lung carcinoma cell line expressing p53
STR – A standard for the authentication of human cell lines

- **ASN-0002 - Authentication of Human Cell Lines: Standardization of STR Profiling**
  - The standard describes a **consistent, inexpensive, and universally applicable method** for authenticating new and established cell lines and their criteria for use
  - **Chair:** John R.W. Masters, University College of London
  - **Co-Chair:** Yvonne A. Reid, ATCC (Retired)
  - **Final action by ANSI:** January 25, 2012
  - **Published date:** February 2, 2012

STR analysis for cell line identity: characteristics

- Target sequence consists of microsatellite DNA (short repeats, 2 – 6 bp, 5 – 50 times)
- Typically use 1-2 ng DNA
- Discrete alleles allow digital record of data
- Markers distributed throughout the genome
- Highly variable within populations; highly informative; high discriminating power
- High observed heterozygosity >70% (more alleles = higher power of discrimination)
- Robust and reproducible results
- Low stutter characteristics
  - 2 bp has high stutter
  - 4 bp has low stutter
## Properties of STRs for DNA profiling (Human Loci)

<table>
<thead>
<tr>
<th>Locus name</th>
<th>Chromosome location</th>
<th>Repeat motif</th>
<th>No. repeating units</th>
</tr>
</thead>
<tbody>
<tr>
<td>D16S539</td>
<td>16q24-qtr</td>
<td>GATA</td>
<td>5-15</td>
</tr>
<tr>
<td>D7S820</td>
<td>7q11.21-22</td>
<td>GATA</td>
<td>6-15</td>
</tr>
<tr>
<td>D13S317</td>
<td>13q22-q31</td>
<td>TATC</td>
<td>5-15</td>
</tr>
<tr>
<td>D5S818</td>
<td>5p21-q31</td>
<td>AGAT</td>
<td>7-16</td>
</tr>
<tr>
<td>CSF1PO</td>
<td>5q33.3-34</td>
<td>TAGA</td>
<td>6-16</td>
</tr>
<tr>
<td>TPOX</td>
<td>2p23-pter</td>
<td>GAAT</td>
<td>6-13</td>
</tr>
<tr>
<td>vWA</td>
<td>12p23-pter</td>
<td>[TCTA] [TCTG]</td>
<td>10-24</td>
</tr>
<tr>
<td>TH01</td>
<td>11p15.5</td>
<td>TCAT</td>
<td>3-14</td>
</tr>
<tr>
<td>Amelogenin</td>
<td>Gender determination</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Power of discrimination** 1 : 1.2 x 10⁸

Retrospective studies on 500 human cell lines – minimum of 8 STR markers required to uniquely identify a human cell line
Outline of STR profiling procedure

**Outline of STR profiling procedure**

- **Multiplex PCR**
  - Amplification of STR loci
  - Simultaneous fluorescent labeling

- **Capillary Electrophoresis (CE)**
  - Addition of Internal Lane Standard
  - CE to separate fragments
  - Fluorescent detection
  - Run allelic ladder in parallel

- **Data Analysis**
  - Calculate size based on Internal Lane Standard
  - Compare fragment sizes to allelic ladders to determine STR alleles
  - Compare to databases

**Requirements**:
- Gene sequencer
- Thermocycler
- Primer sets
- STR database of cell lines
- Experienced technicians

**Sample Electropherogram**
STR DNA polymorphism

A: Homozygous at locus D16S539

- GATA
  - 8 repeating units

B: Heterozygous at locus D16S539

- GATA
  - 10 repeating units
  - 9 repeating units

Unique STR DNA profile for each cell line derived from unrelated individuals
Unrelated human cell lines: STR analysis

2 unrelated cell lines, separate individuals, unique STR DNA profiles
Unrelated mouse cell lines: STR analysis

2 unrelated cell lines, separate individuals, unique STR DNA profiles
Cellular cross-contamination

SK-OV-3
Ovary

SK-OV-3 +
cell line X
## ATCC cell line authentication services

<table>
<thead>
<tr>
<th>ATCC® No.</th>
<th>Designation</th>
<th>Where to order</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human</td>
<td>135-XV™ FTA Sample Collection Kit for Human Cell Authentication Service</td>
<td><a href="http://www.atcc.org/humanSTR">www.atcc.org/humanSTR</a></td>
</tr>
<tr>
<td>Mycoplasma</td>
<td>136-XV™ FTA Sample Collection Kit for PCR-based Mycoplasma Detection Service</td>
<td><a href="http://www.atcc.org/mycotesting">www.atcc.org/mycotesting</a></td>
</tr>
<tr>
<td>Mouse</td>
<td>137-XV™ FTA Sample Collection Kit for Mouse Cell Authentication Service</td>
<td><a href="http://www.atcc.org/mouseSTR">www.atcc.org/mouseSTR</a></td>
</tr>
</tbody>
</table>
Mouse STR Profiling
Goal: Validate an STR multiplex PCR assay to distinguish Mouse cell lines

NIST distributed testing kits containing:
- DNA from 50 ATCC mouse cell lines (blinded)
- Primer Mix
- PCR Master Mix
- Calibrants (allelic ladder)
- Control DNA
- Protocols
Mouse STR Consortium: the assay

- Primers designed to *Mus musculus musculus* (NCBI build 38.1)

- 18 Mouse STR loci & 2 Human loci for contamination detection

- New STR assay builds on NIST’s 9 marker assay published in 2014
  - *Mouse cell line authentication*; J.Almeida Cytotechnology 2014 Jan; 66(1): 133 - 147
Mouse STR Consortium: the results

1) Unique STR profiles obtained for each mouse cell line tested

2) 42 Validated Mouse STR profiles uploaded to NCBI Biosample (≥ 98% concordance)

3) STR Locus 11-1 removed due to abnormal peak morphology
Mouse STR at ATCC
This technology has been licensed to ATCC for commercialization – available NOW!

Methodology
- Samples spotted and shipped on Whatman FTA™ Cards
- Profiles compared to ATCC Mouse STR Database
- Results emailed with three to five business days

Data interpretation
- According to NIST granted US Patent (9,556,482) and 2019 consortium publication
- Database matching follows the Tanabe matching algorithm

\[
\text{% match} = \frac{2 \times \text{number of alleles matching}}{\text{# of query alleles} + \text{# of reference alleles}}
\]
Mouse STR at ATCC

1 Online Order

www.atcc.org/mousestr
Mouse STR at ATCC

2 Spot & Ship
Mix & Spot cells or DNA

Cells

DNA

1x10^6 cells/ml
20 ng/µl

Sample Submission Form

Thank you for placing an order for the Mouse STR Testing Service. Please read this form in its entirety and follow all steps accurately.

Peace of mind in 3 easy steps

1

The Mouse STR Testing Results will be emailed by ATCC to the email address provided below.
Customer Information (All fields are required. Please print information to ensure it is legible)

Name

Institution/Company

Address

City State Zip code Country

Email

(Please print clearly; authentication results will be sent to the email address above)

Prepare your sample according to the Sample Preparation Instructions found on the back of this form. Enter the cell line information and sign the hazard statement below (REQUIRED).

Mouse Cell Line Information

Please verify that the barcode number on your sample card matches the barcode number on this form for each sample.

Mouse Cell Line Name/Designation

Catalog/Item # (if any)

Hazard Statement

ATCC does not accept cultures infected with HIV or BioSafety Level 3 or 4 agents for STR testing. To the best of my knowledge the cell line being submitted is free of hazardous materials, agents, and carcinogens.

Print Name

Signature

Date

Mail your sample and this completed form in the pre-addressed Return Envelope included within the kit.
Mouse STR at ATCC

3 In-Lab Processing & Data Analysis

RAW 264.7 (ATCC® TIB-71™)

RAW 264.7 cells
Image courtesy of Donna Stolz, University of Pittsburgh
# Mouse STR Report

Report e-mailed within 3 to 5 business days

## FTA Barcode: MUSA0001

<table>
<thead>
<tr>
<th>Locus</th>
<th>Query Profile: RAW 264.7</th>
<th>Database Profile: TIB-71 RAW 264.7; Mouse (Mus musculus)</th>
</tr>
</thead>
<tbody>
<tr>
<td>18.3</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>4.2</td>
<td>22.3</td>
<td>22.3</td>
</tr>
<tr>
<td>6.7</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>10.2</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>11.1</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>7.1</td>
<td>25.2</td>
<td>25.2</td>
</tr>
<tr>
<td>1.1</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>3.2</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>8.1</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>2.1</td>
<td>18</td>
<td>16</td>
</tr>
<tr>
<td>15.3</td>
<td>22.3</td>
<td>22.3</td>
</tr>
<tr>
<td>6.4</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>11.2</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>17.2</td>
<td>14</td>
<td>16</td>
</tr>
<tr>
<td>12.1</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>5.5</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>X.1</td>
<td>24</td>
<td>24</td>
</tr>
<tr>
<td>13.1</td>
<td>16.2</td>
<td>16.2</td>
</tr>
</tbody>
</table>

Number of shared alleles between query sample and database profile: 20
Total number of alleles in the query sample profile: 20
Total number of alleles in the database profile: 20
Percent match between the submitted sample and the database profile: 100
Mouse STR Report

Report e-mailed within 3 to 5 business days

<table>
<thead>
<tr>
<th>Result</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>The submitted sample profile is an exact match for the following ATCC cell line(s) in the ATCC mouse STR database: TIB-71</td>
</tr>
<tr>
<td>No</td>
<td>The submitted sample profile is mouse, however a matching reference profile has not previously been established in the ATCC mouse STR database.</td>
</tr>
<tr>
<td>No</td>
<td>The submitted profile is similar to the following ATCC cell line(s):</td>
</tr>
<tr>
<td>No</td>
<td>An STR profile could not be generated from the submitted sample.</td>
</tr>
<tr>
<td>No</td>
<td>Human and/or African Green Monkey Species Detection</td>
</tr>
<tr>
<td></td>
<td>Human and/or African green monkey has been detected in the submitted sample profile (see attached electropherogram at Human D8 &amp; D4 loci).</td>
</tr>
</tbody>
</table>

Mouse STR at ATCC
### Mouse STR Report

Addendum: Comparative Output from ATCC Mouse STR Database

<table>
<thead>
<tr>
<th>ATCC Cat. No.</th>
<th>100</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>Designation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18.3</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>42</td>
<td>22.3</td>
<td>22.3</td>
</tr>
<tr>
<td>6.7</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>19.2</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>12</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>7.1</td>
<td>25.2</td>
<td>25.2</td>
</tr>
<tr>
<td>11.1</td>
<td>15.16</td>
<td>15.16</td>
</tr>
<tr>
<td>3.2</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>6.1</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>2.1</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>15.3</td>
<td>22.3</td>
<td>22.3</td>
</tr>
<tr>
<td>6.4</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>11.2</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>17.2</td>
<td>14.16</td>
<td>14.16</td>
</tr>
<tr>
<td>12.1</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>24.1</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>X-1</td>
<td>24</td>
<td>24</td>
</tr>
<tr>
<td>13.1</td>
<td>16.2</td>
<td>16.2</td>
</tr>
</tbody>
</table>

Definitions of terms used in this report:

**Artifact**: A non-allelic product of the amplification process, an anomaly of the detection process, or a by-product of primer synthesis.

**Pull-up**: A term used to describe when signal from one dye color channel produces artificial peaks in another, usually adjacent, color.

**Spike**: An extraneous peak resulting from dust, dried polymer, an air bubble, or an electrical surge.

**Dye blob**: Free dye not coupled to primer that can be injected into the capillary.
Mouse STR at ATCC – key points

- ATCC worked with NIST to pioneer STR profiling for mouse cell lines (*Available NOW*)
- ATCC authentication services are simple and inexpensive, after placing your order:
  - Spot
  - Dry
  - Mail
  - Receive you results in three to five days

**Report includes:**
- Submitted & Matched Allele Calls
- Contamination check
- Comparative output for database comparison
- PDF of the submitted sample profile
Why authenticate? Journals now require cell authentication

<table>
<thead>
<tr>
<th>Publisher and/or Journal Title</th>
<th>Required</th>
<th>Encouraged</th>
</tr>
</thead>
<tbody>
<tr>
<td>American Association for Cancer Research journals (8)</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>BioMed Central journals (200+)</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Endocrine Society journals (5)</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td><em>Nature</em> journals (approximately 150)</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Society for Endocrinology journals (3)</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td><em>BioTechniques</em></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td><em>Carcinogenesis</em></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td><em>Cell Biochemistry and Biophysics</em></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td><em>Cell Biology International</em></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td><em>International Journal of Cancer</em></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td><em>Investigative Ophthalmology &amp; Visual Science</em></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td><em>In Vitro Cellular &amp; Developmental Biology—Animal</em></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td><em>Journal of Molecular Biology</em></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td><em>Journal of the National Cancer Institute</em></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td><em>Molecular Vision</em></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td><em>Neuro-Oncology</em></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td><em>PLOS ONE</em></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

*Fusenig NE, The need for a worldwide consensus for cell line authentication: Experience implementing a mandatory requirement at the International Journal of Cancer. 2017. https://doi.org/10.1371/journal.pbio.2001438*
Why authenticate? Agencies require cell authentication

NIH revised guidelines to applications for funding - Enhancing Reproducibility through Rigor and Transparency (effective Jan. 25, 2016)

- Authentication of key biological and/or chemical resources
  - NIH expects that key biological and/or chemical resources will be regularly authenticated to ensure their identity and validity for use in the proposed studies


FDA Guidance for Industry: Characterization and Qualification of Cell Substrates and Other Biological Materials

- Provides manufactures of viral vaccines recommendations for the characterization of cell substrates and other biological materials for human use

Docket No: FDA-2006-D-0223; February 2010
When to authenticate?

- At the beginning and conclusion of a study
  - Before submission of grant applications
  - Before manuscript submissions
- After preparing a cell bank
- At regular intervals throughout study
- When in doubt
  - Novel phenotypic behavior observed
Summary

- Cell line misidentification and contamination bears a high cost
- STR technology offers a powerful means to identify cells
- ATCC has helped pioneer STR profiling for mouse and human cells
- Journals and funding agencies require evidence of authentication
- There is no wrong time to authenticate your cell lines
- www.atcc.org/mouseSTR for more information
Cultivating collaboration to support global health

Go to www.atcc.org/mouseSTR for the ATCC Mouse STR Profiling Service

Upcoming webinars:

- Simplifying assay development with molecular standards: Remove culturing from the equation
  September 26, 12:00 ET

www.atcc.org/webinars