Available online at www.sciencedirect.com

ScienceDirect

Neuromuscular Disorders 27 (2017) 439–446

www.elsevier.com/locate/nmd

Establishing a reference dataset for the authentication of spinal muscular atrophy cell lines using STR profiling and digital PCR

Deborah L. Stabley [a], Jennifer Holbrook [a], Ashlee W. Harris [b], Kathryn J. Swoboda [c], Thomas O. Crawford [d,e], Katia Sol-Church [a,f,g], Matthew E.R. Butchbach [b,f,g,h,*]

[a] *Nemours Biomolecular Core Laboratory, Nemours Biomedical Research, Nemours Alfred I. duPont Hospital for Children, Wilmington, DE, USA*
[b] *Center for Applied Clinical Genomics, Nemours Biomedical Research, Nemours Alfred I. duPont Hospital for Children, Wilmington, DE, USA*
[c] *Neurogenetics Research Program, Center for Human Genetics Research, Massachusetts General Hospital, Boston, MA, USA*
[d] *Department of Neurology, Johns Hopkins University, Baltimore, MD, USA*
[e] *Department of Pediatrics, Johns Hopkins University, Baltimore, MD, USA*
[f] *Center for Pediatric Research, Nemours Biomedical Research, Nemours Alfred I. duPont Hospital for Children, Wilmington, DE, USA*
[g] *Department of Pediatrics, Thomas Jefferson University, Philadelphia, PA, USA*
[h] *Department of Biological Sciences, University of Delaware, Newark, DE, USA*

Received 30 November 2016; received in revised form 23 January 2017; accepted 3 February 2017

Abstract

Fibroblasts and lymphoblastoid cell lines (LCLs) derived from individuals with spinal muscular atrophy (SMA) have been and continue to be essential for translational SMA research. Authentication of cell lines helps ensure reproducibility and rigor in biomedical research. This quality control measure identifies mislabeling or cross-contamination of cell lines and prevents misinterpretation of data. Unfortunately, authentication of SMA cell lines used in various studies has not been possible because of a lack of a reference. In this study, we provide said reference so that SMA cell lines can be subsequently authenticated. We use short tandem repeat (STR) profiling and digital PCR (dPCR), which quantifies *SMN1* and *SMN2* copy numbers, to generate molecular identity codes for fibroblasts and LCLs that are commonly used in SMA research. Using these molecular identity codes, we clarify the familial relationships within a set of fibroblasts commonly used in SMA research. This study presents the first cell line reference set for the SMA research community and demonstrates its usefulness for re-identification and authentication of lines commonly used as *in vitro* models for future studies.

Keywords: Spinal muscular atrophy; Cell line authentication; Short tandem repeat profiling; Digital PCR; *SMN1*; *SMN2*

1. Introduction

Translational research requires the use of cell culture models of disease which are established from tissue samples from patients. Disease-specific human cell lines must be correctly identified so as to enhance reproducibility and experimental rigor. Misidentification of cell lines can have a major negative impact on translational research [1–3]. Up to 15% of human cell lines are believed to be either misidentified or contaminated with other cell lines [4,5]. Mislabeling, cross-contamination by another cell line or genetic changes caused by repeated passaging can all lead to cell line misidentification [6]. Cell line authentication, which is meant to prevent incorrect interpretation of data generated from misidentified cell lines, is becoming a requirement for federal funding and for publication. We must have a molecular identity code for a given cell line in order to authenticate it.

Short tandem repeat (STR) profiling is one of the most commonly used techniques for generating a specific molecular identity code [7–10]. This assay amplifies a set of polymorphic microsatellite, or STR, markers and then resolves the PCR products by capillary electrophoresis size fractionation [11]. STR profiling has strong discriminatory power even with degraded DNA samples and is able to resolve mixtures of DNA [12]. Furthermore, the data generated is not dependent on one specific testing platform making this assay very powerful for

* Corresponding author. Nemours Biomedical Research, Nemours Alfred I. duPont Hospital for Children, 240 Rockland Center One, 1600 Rockland Road, Wilmington, DE 19803. Fax: +302 651 6539.
E-mail address: butchbach@nemoursresearch.org (M.E.R. Butchbach).

http://dx.doi.org/10.1016/j.nmd.2017.02.002
0960-8966/

forensic analysis – its original purpose – and for cell line identification or authentication.

While most of the effort in cell line authentication has focused on cancer, identification and authentication of cell lines derived from patients with rare monogenic disorders are equally important to standardize, especially since these biospecimens being so rare are usually shared across many laboratories. One such disorder is spinal muscular atrophy (SMA), an autosomal recessive motor neuron disease that is a leading genetic cause of infant mortality [13,14]. However, as is the case for many neurodegenerative disorders, SMA has a wide range of phenotypes. SMA is caused by the loss of *SMN1* (*survival motor neuron 1*) but retention of the duplicate gene *SMN2*. The clinical severity of SMA is inversely related to the number of *SMN2* copies, i.e. patients with higher *SMN2* copy numbers have a less severe clinical presentation of this disease (reviewed in Ref. [15]). Many important advances in understanding the molecular pathology of SMA and the impact of *SMN2* dosage on severity have been validated using patient-derived cell lines [16–22]. SMA fibroblast lines have been converted into induced pluripotent stem cell lines subsequently used to characterize human-derived *SMN1*-deleted neural cells in culture [23–26]. Drug discovery efforts to identify small molecule inducers of *SMN2* expression and modulators of alternative splicing have relied on the use of SMA patient cell lines [27]. In fact, the discovery of nusinersen (Spinraza, ISIS-SMN_{Rx} or ISS-N1), the first FDA-approved SMA drug, was made using SMA fibroblast lines [28].

Due to their importance as *in vitro* disease models for SMA, it is essential to authenticate patient-derived cell lines in order to ensure reliability and reproducibility. Unfortunately, STR profiles for the commonly used SMA cell lines are not currently available and thus cell line authentication is not possible. Furthermore, accurate copy number determination of *SMN2* – as measured with array digital PCR (dPCR) [21] – has not yet been reported for specific SMA cell lines. In this study, we generate molecular identity codes based on STR profiling and dPCR copy number measurements for SMA fibroblasts and lymphoblastoid cell lines (LCLs) that are commonly used in SMA research. The set of unique molecular identity codes can serve as a reference to re-identify or authenticate cell lines in future studies which will enhance standards for rigor and reproducibility.

2. Material and methods

2.1. Cell lines

Those fibroblast and LCLs cell lines with GM or UMB prefixes were purchased from Coriell Cell Repositories (Camden, NJ) or the NICHD Brain and Tissue Bank for Developmental Disorders at the University of Maryland (Baltimore, MD). The fibroblast lines with a TC prefix were obtained from Johns Hopkins University (Baltimore, MD), those with a KS prefix were from the University of Utah (Salt Lake City, UT) and those with an AIDHC prefix were from Nemours/Alfred I. duPont Hospital for Children (Wilmington, DE). The fibroblast lines generated at Nemours were either obtained from the Molecular Diagnostics Laboratory or established in the Motor Neuron Diseases Research Laboratory using standard procedures [29].

2.2. Ethics statement

For those cell lines obtained from non-commercial sources, biospecimens were obtained after written consent or assent and parental permission. This study was approved by the specific Institutional Review Boards. These cell lines were de-identified so that no protected health information is known for these lines. Available information regarding phenotype and genotype is generally limited to gender, *SMN1* deletion/mutation status, *SMN2* copy number and SMA type.

2.3. Cell line maintenance

In this study, we used 24 type I SMA, 23 type II SMA, 11 type III SMA and 26 healthy control fibroblasts and LCLs. The lowest passage number stock vial was used for each cell line. All cell lines were maintained in a humidified 37 °C incubator with 5% CO_2. Fibroblasts were maintained in Dulbecco's modified essential medium (DMEM; Life Technologies, Grand Island, NY) supplemented with 10% fetal bovine serum (EquaFETAL; Atlas Biologicals, Fort Collins, CO), 2 mM L-glutamine (Life Technologies) and 1% penicillin-streptomycin (Life Technologies). LCLs were maintained in RPMI-1640 medium (Life Technologies) supplemented with 15% EquaFETAL, 2 mM L-glutamine and 1% penicillin-streptomycin. Cell pellets were collected from each line within 3 passages.

2.4. Genomic DNA isolation

Genomic DNA was isolated from cell pellets using the Gentra Puregene Cell Kit (QIAGEN, Germantown, MD) as described previously [21]. Yield was measured with an ND-2000C NanoDrop spectrophotometer (Thermo Fisher Scientific, Waltham, MA). Double-stranded DNA concentrations were measured using the Qubit dsDNA Broad Range Assay (Life Technologies) as recommended by the manufacturer.

2.5. Digital PCR (dPCR)

SMN1 and *SMN2* copy numbers for each sample were determined using the QuantStudio 3D Digital PCR System (Life Technologies) as described previously [21]. *SMN1* and *SMN2* copy numbers were normalized against those for *RPPH1*, as its copy number does not vary in the human population [30]. The testers were blinded to the subject identities of the cell lines until after analysis.

2.6. Short tandem repeat (STR) profiling

2 ng gDNA were processed for STR profiling using the AmpFISTR Identifiler PCR Amplification kit (Life Technologies) according to the manufacturer's direction. This kit includes 15 polymorphic microsatellite markers and the XY chromosome-specific marker *amelogenin* (*AMEL*). After PCR amplification, the samples were analyzed on the ABI 3130xl Genetic Analyzer (Life Technologies) using the GeneMapper