DNA Fingerprinting: A Quality Control Case Study for Human Biospecimen Authentication

Olga A. Kofanova,¹ William Mathieson,² Gerry A. Thomas,² and Fotini Betsou¹

This case study illustrates the usefulness of the DNA fingerprinting method in biobank quality control (QC) procedures and emphasizes the need for detailed and accurate record keeping during processing of biological samples. It also underlines the value of independent third-party assessment to identify points at which errors are most likely to have occurred when unexpected results are obtained from biospecimens.

Introduction

IN MANY CASES, PUBLISHED SCIENTIFIC RESEARCH results are not reproducible and the ability to verify new discoveries fails. It has been estimated that up to one-third of all cell lines have an origin other than that assumed by the investigators,¹ and research articles have been published with misidentified cell lines leading to invalid results.^{2,3} Cell line contamination and authentication has been a huge issue for many decades, with reports regularly emerging ever since the first (HeLa) cell line was established in the 1950s.^{4–6} For example, it was reported by Boonstra et al.¹ that the frequently used cell lines SEG-1, BIC-1, and SK-GT-5 were from other tumor types. Experimental results based on these contaminated cell lines have led to clinical trials recruiting patients, to more than 100 scientific publications, and to at least three National Institutes of Health cancer research grants and 11 US patents.¹

In the recent discussion by Reid and Mintzer,⁷ the importance of this problem for biobanks has been mentioned. However, biobanks contain not only cell lines, but also different types of samples and human biological material, such as tissue, blood, urine, extracted cells, and nucleic acids. Such samples are also used in wide-ranging projects designed to identify biomarkers, underlying causes of diseases, or genes associated with a particular disease. The results of these studies can potentially lead to the development of new treatments or prognostic and diagnostic biomarkers. During handling and processing, samples and associated data may undergo different processes, possibly performed at different institutions or locations and by different operators. Therefore it is imperative to maintain detailed and accurate records of the samples and their associated data during all the processes they undergo.

Whenever a doubt relating to sample authenticity arises in a biobank, it is usually difficult to trace the sample back and categorically establish that it originates from the correct donor, because the sample has been anonymized and the patient's details are therefore unknown. In such cases, the biobank will struggle to certify that the samples have been correctly supplied to the research group. It is therefore crucial for biobanks to implement a quality control (QC) assay/procedure to guarantee authentication of stored biospecimens.

To overcome the sample traceability problem, the ISBER Biospecimen Science Working Group has proposed that each biobank uses an allocated and unique structured code—the SPREC.⁸ However, it is still necessary to overcome the sample authentication and traceability problem. In this report, we describe a case study of biospecimen authentication from a tissue bank using DNA fingerprinting. This technique is regularly used in forensic science and in cell line discrimination/authentication and allows sample-discrimination with extremely high confidence (in the order of one in several billion).⁹ In terms of biobanking, DNA fingerprinting could be a useful QC assay procedure, since the technique can be applied to any derivatives, such as cell line collections, cells isolated from biological fluids, or any DNA obtained from any biospecimen type.

In the current study, a researcher using nucleic acid obtained from fresh-frozen tissue blocks supplied by a tissue bank, for “omics” technology, reported that there appeared to be a mismatch between samples from three different cases. All the samples were collected by a single Institute (Institute A), stored there and then transported to a second Institute (Institute B) where the DNA and RNA were extracted in 2009 and 2011. Care was taken to document the

¹Integrated Biobank of Luxembourg (IBBL), 6 rue Nicolas Ernest Barblé, 1210 Luxembourg, Luxembourg.

²Human Cancer Studies Group, Department of Surgery and Cancer, Imperial College London, Charing Cross Hospital, London, United Kingdom.

extraction procedure and it was ascertained that it was unlikely that the samples had been mislabeled during the extraction process, as they were extracted by different operators on different dates.

Two possibilities therefore remained: mislabeling at the collection center, or mislabeling at the researcher's institute (Institute C). The "omics" technique involved multiple steps and there was therefore a third possible scenario—the potential mislabeling, or contamination of tubes could have occurred during the analysis.

Institute A had initially provided Institute B with fresh-frozen blocks of patient-matched tumor and normal tissue for DNA extraction. In total, there were three patients (therefore six blocks) for which the origin had been questioned by Institute C. There was also DNA available from extractions, also carried out by Institute A, but using duplicate blocks for two of the patients (three blocks). So, in total, aliquots of nine DNA samples were still retained at Institute B. These were sent to the Integrated Biobank of Luxembourg (IBBL), which had agreed to act as an independent assessor and perform DNA fingerprinting on the samples.

Materials and Methods

Ethics statement

The sample collection used in the study was approved by the Imperial College Research Ethics Committee in UK (protocol number ICREC-8-2-4) and the U.S. National

Cancer Institute (NCI IRB protocol number is 0H00-C-N024) and patient data were analyzed anonymously.

DNA samples

DNA was extracted from the frozen tissue samples using a QIAamp Mini DNA extraction kit (Qiagen) according to a slightly modified protocol (the modifications principally consisting of the incorporation of additional wash steps).

A single 1 µg aliquot of double-stranded DNA, assayed using the QuBit ds BR assay (Invitrogen) from each of the extractions, was sent to the IBBL in January 2013. IBBL was blinded to the sample identities.

DNA fingerprinting

Using an in-house developed IBBL biospecimen QC assay for authentication of biospecimen identity, all samples were tested for different single- and multi-locus variable numbers of tandem repeats (VNTRs) based on polymerase chain reaction (PCR) amplification of highly polymorphic minisatellite VNTR loci with generation of banding patterns and a specific DNA profile for each sample. For that procedure, the PCR amplification of six minisatellite VNTR loci (Apo-B1: with expected product length 522–909 bp; Col2A1: 600–850 bp; D17S5: 168–1080 bp; D1S80: 400–940 bp; D2S44: 600–>5000 bp; PAH: 370–760 bp), and one microsatellite locus (GTG_N: 10–>15000 bp) was performed.^{10,11} After size determination of the amplicons from each loci on an agarose

FIG. 1. A relationship between the examined samples.

gel using gel-analyzing software (GE HealthCare, ImageQuant TL), the DNA profiles of all samples were compared to each other.

Results and Discussion

Sample mislabeling or cross-contamination may arise during initial collection or later processing as a result of pipetting, labeling, or other manipulation errors, leading to incorrect assessment of samples in downstream analytical applications.¹ In our current study, we report the use of DNA fingerprinting as a QC assay for biospecimen authentication.

DNA extracted from fresh-frozen tissue from nine unknown samples was analyzed by DNA fingerprinting. All nine samples were successfully analyzed and three different profiles were obtained, proving that the samples originated from three different patients: samples labeled as “1” and “5” were identical; samples labeled as “2”, “3”, “7”, and “9” were identical; and samples labeled as “4”, “6”, and “8” were identical (Fig. 1). The DNA fingerprinting QC assay report was sent to Institute B.

The DNA fingerprinting analysis showed that there had been no mismatch of samples between individuals, either at Institute A (during sample collection) or at Institute B during the extraction procedures. The researcher at Institute C was finally provided with one extra aliquot of DNA from tumor and normal tissue from each of the cases in question, where duplicate blocks had been extracted, in order to re-run their analyses.

The number of alleles of the forensically-relevant short tandem repeat (STR) loci ranges from 5 to 20 common alleles. There are a variety of commercial kits available for the robust multiplex amplification of these core STR loci.^{4,12} In spite of being the most commonly and routinely used markers in forensics, STRs have some shortcomings. The application of single-nucleotide polymorphism (SNP) markers additionally offers a useful and increasingly important extension to a routine STR-based DNA profiling.¹²

In the current study, the multiplexed PCR amplification fragment length polymorphism (AmpFLP) of six prominent and highly polymorphic minisatellite VNTR loci and one additional microsatellite locus for sex determination using the detection of the *SRY* gene on the Y chromosome was performed. This combination allows discrimination of one biospecimen from another at the level of 10^6 .^{10,11} This specificity is in our view sufficient for biobanking authentication QC purposes. Therefore, this is a rapid but highly reliable and robust method, enabling easy verification of biospecimen identity.

This scenario emphasizes the need for detailed and accurate record keeping during processing of biological samples, and the value of independent third-party assessment to identify points at which errors are most likely to have occurred when unexpected results are found. This case study, together with another recently published study on a new DNA fingerprinting technique,¹³ demonstrate the value of the DNA fingerprinting methods in biobank QC procedures

and exemplifies the use of this quality control method in a wide range of sample authentication situations.

Author Disclosure Statement

No competing financial interests exist.

References

1. Boonstr JJ, Van Marion R, Beer G, et al. Verification and unmasking of widely used human esophageal adenocarcinoma cell lines. *J Natl Cancer Inst* 2010;102:271–274.
2. Drexler HG, MacLeod RA, Dirks WG. Cross-contamination: HS-Sultan is not a myeloma but a Burkitt lymphoma cell line. *Blood* 2001;98:3495–3496.
3. Drexler HG, Dirks WG, MacLeod RA. False human hematopoietic cell lines: Cross-contaminations and misinterpretations. *Leukemia* 1999;13:1601–1607.
4. ATCC SDO Workgroup ASN-0002. Cell line misidentification: The beginning of the end. *Nature Rev Cancer* 2010;10:441–448.
5. Nelson-Rees WA, Daniels DW, Flandermeyer RR. Cross-contamination of cells in culture. *Science* 1981;212:446–452.
6. Fogh J, Wright WC, Loveless JD. Absence of HeLa cell contamination in 169 cell lines derived from human tumors. *J Natl Cancer Inst* 1977;58:209–214.
7. Reid Y, Mintzer J. The current state of cell contamination and authentication—And what it means for biobanks. *Biopreserv Biobank* 2012;10:236–238.
8. ISBER Working Group on Biospecimen Science. Standard Preanalytical Coding for Biospecimens: Review and Implementation of the Sample Preanalytical Code (SPREC). *Biopreserv Biobank* 2012;10:366–374.
9. Kayser M, De Knijff P. Improving human forensics through advances in genetics, genomics and molecular biology. *Nat Rev Genet* 2011;12:179–192.
10. Dirks WG, Drexler HG. Authentication of scientific human cell lines: Easy-to-use DNA fingerprinting. *Methods Mol Biol* 2005;290:35–50.
11. Dirks W, MacLeod RA, Jäger K, et al. First searchable database for DNA profiles of human cell lines: Sequential use of fingerprint techniques for authentication. *Cell Mol Biol* 1999;45:841–853.
12. Ziętkiewicz E, Witt M, Daca P, et al. Current genetic methodologies in the identification of disaster victims and in forensic analysis. *J Appl Genet* 2012;53:41–60.
13. Palmirotta R, De Marchis ML, Ludovici G, et al. A reliable and reproducible technique for DNA fingerprinting in biorepositories: A pilot study from BioBIM. *Int J Biol Markers* 2013;28:398–404.

Address correspondence to:

Dr. Olga Kofanova
Integrated Biobank of Luxembourg (IBBL)
6 rue Nicolas Ernest Barblé
1210 Luxembourg
Luxembourg

E-mail: olga.kofanova@ibbl.lu