White Paper

SUITABILITY OF THE CO1 BARCODE ASSAY FOR ANIMAL CELL IDENTITY TESTING

ABSTRACT:

The author discusses the use of a PCR and sequencing based cytochrome c oxidase subunit 1 (CO1) barcoding assay for establishing the species-level identity of animal cells. The assay targets a semi-conserved region of the mitochondrial CO1 gene for species-level identity testing of animal (mammalian and insect) cells. The "CO1 Barcode Assay" is intended as an alternative to the isoenzyme analysis assay, for which reagents are no longer commercially available. This White Paper provides the opinion of a subject-matter expert on the suitability of the CO1 Barcode Assay platform for species-level identity testing and as a replacement for isoenzyme analysis testing for animal cells.

Author:

Raymond W. Nims, Ph.D.
RMC Pharmaceutical Solutions, Inc.

ABOUT THE AUTHOR:

Raymond Nims developed, implemented, and directed the conduct of the isoenzyme analysis assay for cell line characterization at BioReliance between 1996 and 2006. During this time, he characterized the ability of the assay to detect inter-species cell mixtures[1] and described the optimization of the assay as run at BioReliance[2]. Since 2009, Ray has provided consulting for the pharmaceutical/biologics industries as an employee of RMC Pharmaceutical Solutions, Inc. Ray is a founding member of the International Cell Line Authentication Committee (ICLAC), and has been or is currently a participant in development of three ATCC SDO standards relating to cell line authentication: ASN-0002 "*The Authentication of Human Cell Lines – Standardization of STR Profiling*", ASN-0003 "*Species-Level Identification of Animal Cells through Mitochondrial Cytochrome c Oxidase Subunit 1 (CO1) DNA Barcodes*", and proposed ASN-0004 "*Species-Level Identification and Cross-Contamination Screening in Animal Cells by Multiplex PCR*".

1. Background

Researchers have, until now, been relatively free to conduct experiments using cells and to publish results of their investigations without concerning themselves with cell identity testing. This is beginning to change, as more and more journals and granting agencies are now demanding proof of cell line authenticity. It is becoming increasingly common to see papers being retracted due to cell line misidentification.

In regulated spaces, such as biologics manufacturing and safety testing, there have existed cell substrate characterization guidance documents[3-7] mandating identity testing for more than two decades. Some of these guidance documents mention specific methodologies for establishing identity, while others do not. The primary requirement was, and still remains, demonstration that the production (or testing) substrate is the same cell line identified in the regulatory submission documents, and not some cross-contaminant or mis-identified cell. Historically, diploid cells have been identified by karyotyping and isoenzyme analysis. Aneuploid (i.e., transformed/continuous) cells have been identified most commonly by isoenzyme analysis. The reason for this is that these were the methods in place during the time frame covered by the earliest guidance documents[e.g., 3,4]. It is important to remember that these methods provide species-level identity only. In addition, each of these historical methods is capable of detecting an inter-species cell mixture provided that the contaminant represents ~5-10% of the overall cell population[1]. These methods are therefore considered also to be capable of establishing cell line purity (i.e. purity defined as freedom from contaminating cells of another species).

The FDA does not typically mandate specific methods to be used for a given cell characterization endpoint. There has, until quite recently, therefore been little motivation to update the analytical methods used for establishing cell line identity and purity, with (for instance) one of the newer molecular methods. However, as a result of the unexpected non-availability since early 2015 of isoenzyme analysis reagents from the sole supplier, it has been necessary to identify suitable replacement methods for that cell identity assay. The replacement assays for identifying human cells include short tandem repeat (STR)

profiling[8-10] and single nucleotide polymorphism (SNP) profiling[11]. These are capable of identifying human cells to the individual donor level. From a regulatory standpoint, this must be viewed as a vast improvement in identity testing. The replacement technologies now being considered for animal cell identity testing include DNA fingerprinting and CO1 barcoding. Will these be methods also be acceptable to the regulatory agencies? The answer is that a science-driven rationale for replacing isoenzyme analysis with one of these newer candidate technologies should be acceptable to the agencies.

This White Paper has been written to provide such a science-driven rationale for replacement of the isoenzyme analysis assay with BioReliance's CO1 Barcode Assay for species-level identity testing of animal (mammalian and insect) cells. The document reflects my opinions as well as cited literature and the development report authored by BioReliance for the CO1 Barcode Assay[12].

2. Technical Evaluation of Bioreliance's CO1 Barcode Assay

2.1. The CO1 platform.

DNA barcoding has been used within the zoological community for years and has now been proposed for animal cell identity testing[13]. It is useful since the target (mitochondrial CO1 gene) varies sufficiently between species in order to allow discrimination. A mitochondrial gene is considered to be preferable to a nuclear gene as a target, since nuclear genes display slower rates of evolution. Mitochondrial genes mutate at a higher rate than nuclear genes, and therefore provide better differentiation of species separated in a comparatively shorter evolutionary time frame, while still showing good conservation among individual members of a species[14,15]. Unlike nuclear genes, animal mitochondrial genes rarely contain introns, facilitating PCR amplification. The mitochondrial genome is inherited in a haploid pattern and undergoes limited or no recombination, and thus is associated with relatively less complexity and easier data analysis compared with nuclear genes. Mitochondria are present at high copy number within cells, allowing for a relatively high yield of specific target gene DNA.

As an identity testing method, the DNA barcoding platform has some limitations. The rates of evolution of mitochondrial genes can vary between animal species, which in some cases can lead to overlap of intra- and inter-species distances. DNA barcoding will not distinguish among cell lines from different individuals of the same animal species, or cell lines derived from different tissues of the same donor. CO1 barcodes are not suitable for resolving non-animal species, including fungi and plants. It should be noted, however, that certain of these limitations are shared by the historical methods. These include the inability to distinguish between different tissues of the same donor organism and the inability to distinguish between different donor organisms derived from the same animal species.

2.2. BioReliance's CO1 Barcode Assay.

The CO1 Barcode Assay offered by BioReliance is a modification of the standard CO1 barcoding platform. Development of the assay[12] involved design and screening of a set of mammalian and insect universal forward and reverse primers, the optimization of the PCR amplification conditions, and the sequencing qualification for mammalian and insect species commonly employed in regulated industries. The mammalian cells used to qualify the sequencing assay included Chinese hamster (CHO), Syrian hamster (BHK), human (MRC-5), mouse (NIH/3T3), and Cercopithecus monkey (Vero). The sequencing portion of the assay was qualified for insect cells using the *Spodoptera frugiperda* cell line (Sf9).

The optimized primer sets and reaction conditions result in an assay which generates amplicons in response to the presence of the appropriate target mitochondrial DNA. Sequencing results indicate that the assay of the six mammalian and insect cell samples generates amplicons with DNA sequence expected (at least 99% similarity) for the corresponding species of origin. This indicates that the universal primers are performing as expected and that the assay has the requisite specificity for an identity test.

The sequencing assay was qualified using cells derived from six mammalian and insect species, but on theoretical grounds, the assay is capable of detecting each of the mammalian and insect cell species from which cell lines have been derived. As of May 2015, the Barcode of Life Data systems (BOLD) database contained 3.3 million public barcode sequences representing >130,000 named animal species[16].

The criterion for making a species assignment from the amplicon DNA sequencing result is that at least 98% match with a BioReliance Reference Sequence is obtained with a Quality Value of >20[17]. Quality values are assigned by the SeqScape® software and represent an assessment of the probability that the wrong base assignment is made by chance.

While BioReliance's CO1 Barcode Assay is designed as an identity test, it may also have utility as a purity test for cell line characterization. A study[12] performed for informational purposes indicated that cell mixtures containing 90% human cells and

10% mouse cells or 90% mouse cells and 10% human cells (i.e., 10% inter-species cross-contaminant) generated amplicons that were detected by the assay as mixed sequences. In such cases, the SeqScape® software indicates that a higher than expected number of base changes are detected and an assignment of "inconclusive" results. The inconclusive result will then trigger an investigation.

3. Technical Evaluation of the Isoenzyme Analysis Assay

3.1. Historical acceptance of the isoenzyme analysis method.

Consideration of the various cell substrate regulatory guidance documents[3-7] reveals that isoenzyme analysis is one of a variety of technologies that may be used to establish species-level identity of cell lines. Isoenzyme analysis is specifically mentioned as an appropriate identity test in ICH Q5D[4], WHO[5], and European Pharmacopoeia 5.2.3[6], although in each case it is clearly indicated that other technologies can be employed for this purpose. For instance, in the relatively new WHO guidance[5], it is stated that: "Other tests that may be used but tend to be less specific include isoenzyme analysis and karyology, which may be particularly useful where there are characteristic marker chromosomes. However, where more specific genetic markers are available, they should be considered."

3.2. The isoenzyme analysis assay platform.

Isoenzyme analysis is a cytosolic protein electrophoretic method that provides species-level identity *verification* for cells derived from the animal species most commonly used in cell culture. This platform is based on the genetic polymorphisms of cytosolic enzymes such as peptidase B, nucleoside phosphorylase, glucose-6-phosphate dehydrogenase, malate dehydrogenase, aspartate aminotransferase, mannose phosphate isomerase, and lactate dehydrogenase. These polymorphisms result from point mutations of the encoding genes during species evolution. The resulting isoenzymes exhibit differences in electrophoretic mobilities that are used to discriminate between cell lines derived from different animal species.

This method requires optimization of the testing strategy with regard to color development time, correction of migration distances relative to the mobility of control reagents, and selection of the appropriate enzyme or enzyme combinations for confirming an expected animal species of origin[2]. In some instances, closely migrating bands for a given enzyme may leave two or more animal species unresolved. For this reason, an assignment of species of origin for a test sample is made on the basis of mobility information obtained through analysis of four or more target enzymes[1,2]. The reagents required to perform this assay are available only from a single supplier, and the recent (since March 2015) non-availability of reagents has severely limited the use of this method worldwide.

Enzyme mobility results obtained for a test sample must be compared to a set of expected migration distances for known animal species that is provided by the reagent manufacturer[18]. This set of tabular migration distances consists of 25-30 animal species, including most of the commonly employed mammalian cell types and at least two insect species. Since not all animal and insect species are represented by the tabular migration distance values provided, isoenzyme analysis results are most appropriately interpreted and reported as either *consistent with* or *not consistent with* the expected animal species of origin. Because of this, the results are not able to exclude other possible animal species of origin (i.e., species for which data have not been provided) and therefore cannot definitively establish the species of origin for a test cell.

The isoenzyme analysis method is capable of detecting inter-species mixtures of cells, albeit with low sensitivity. The presence of an inter-species cross contaminant may be detected if it represents ≥10% of the total cell population[1]. The method cannot detect mixtures of cells derived from the same animal species or cells derived from different tissues of the same donor organism.

It should be noted that isoenzyme analysis has been acceptable to regulatory agencies as a species-level identity test for cell substrates in the past, despite the limitations discussed above.

4. Direct Comparison of The CO1 Barcoding and Isoenzyme Analysis Platforms

4.1. Specificity.

For animal species identity testing, the CO1 barcoding platform is far superior to the isoenzyme analysis platform in terms of specificity (i.e., able to resolve cells derived from different species). The various factors determining specificity are displayed in Table 1.

As displayed in Table 1, the specificity advantages of the CO1 barcoding platform include not only the much greater number of reference species for which information is available, but also the minimal subjectivity for this platform relative to the isoenzyme analysis method and the ability of the barcoding assay to definitively establish the animal species of origin for the test sample.

Table 1. Comparison of determinants of specificity

Determinant	Isoenzyme Analysis	CO1 Barcoding
Reference Species Available	25-30 species [18]*	>130000 species [16]**
Subjectivity	Choice of optimal enzymes for resolving species is operator-dependent; scoring of gels to establish migration distances is somewhat subjective and is dependent on use of proper color development time	No subjectivity in test design or interpretation
Interpretation of Result	Data are consistent with or not consistent with expected animal species of origin	Species of origin is provided, or inconclusive†

* Information for 23 animal species (including two polymorphisms for human cells) is listed in the cited source. Additional species information has in the past been available from Innovative Chemistry.

** As of May 2015 from the BOLD database.

† An inconclusive result which derives from detection of mixed sequences will trigger an investigation.

4.2. Establishing cell line purity (freedom from cross-contaminants).

For demonstrating cell line purity (defined as freedom from cross-contaminating cells), the CO1 barcoding platform is as capable as the isoenzyme analysis platform. The various factors determining capability as a purity assay are displayed in Table 2.

Table 2. Comparison of capability as a cell line purity assay

Determinant	Isoenzyme Analysis	CO1 Barcoding
Ability to detect the presence of an intra-species contaminant	Limited to ability to resolve human polymorphism (A vs. B) for glucose-6-phosphate dehydrogenase	No capability
Sensitivity for determining the presence of an inter-species contaminant	The contaminating cell must represent ≥10% of the total cell population[1]	The contaminating cell must represent ≥10% of the total cell population[12]
Ability to identify the contaminant directly	In most cases, the mobility information for the extra bands attributed to the cross-contaminant may be used to identify the species of the contaminant	The presence of a mixed sequence for the amplicon triggers an investigation. A separate assay must then be used to confirm and identify the cross-contaminant

As shown in Table 2, the isoenzyme analysis and the CO1 barcoding platforms each have some capability as a cell line purity assay. Neither assay can differentiate between cells derived from different tissues of the same donor organism. This characteristic is shared by all available identity test methods. Isoenzyme analysis has only extremely limited ability to resolve intra-species contamination (i.e., human type A vs. human type B polymorphs at the glucose-6-phosphate dehydrogenase locus). The CO1 barcoding platform has essentially no capability for detecting intra-species contamination. The platforms have similar sensitivities for detecting inter-species cross-contamination. In either case, the cross-contaminant must represent ~10% of the total cell population in order to be detected. In reality, true co-cultivations of two cell types are rare. More typically, one of the pair propagates faster and rapidly replaces the other cell type in the co-cultivation[e.g., 1]. The ability of the isoenzyme analysis platform to identify the cross-contaminant directly is not a major advantage. For investigation of an inconclusive result in a CO1 Barcode Assay, BioReliance has at its disposal investigational tools that will serve this purpose, including Next Generation Sequencing.

5. Discussion

The isoenzyme analysis assay has served as the prototypic identity test for continuous animal cell lines for a number of decades, complementing and in some cases superseding the more labor-intensive karyotyping method historically used for diploid cells. Advances in analytical science have paved the way for replacement of isoenzyme analysis with STR-profiling or SNP profiling for human cells. The specificity gains associated with the molecular methods, as well as the existence since 2012 of an ANSI standard for STR profiling[10], have helped to overcome the inertia associated with longstanding use of the electrophoretic method.

In the case of animal cell identity testing, STR profiling and SNP profiling are not (yet) suitable for most animal species. This is a reflection of the relatively great numbers of species for which identity testing is needed and the lack of profiling primers and datasets required for such testing. In the case of identity testing for animal cells, CO1 barcoding and DNA fingerprinting appear now to be the preferable technologies. Despite these analytical advances, it is likely that isoenzyme analysis would have continued to be the method most commonly used for cell identity testing within regulated industries. It is only the non-availability of the reagents required for the conduct of isoenzyme analysis that is now driving the consideration of replacement methods.

The fact that the newer technologies have not already replaced isoenzyme analysis for animal cell identity testing should not be interpreted to mean that the latter technology is better suited for the purpose. The opposite is true. In fact, the method update has not occurred previously due to the inertia present in the regulated industries (i.e., the need for revision of release specifications for cell banks, or the need to file for approval of method changes, etc.) and the fact that there has been essentially no regulatory advocation for moving to the newer methods.

In this White Paper, the CO1 barcoding platform, and the BioReliance CO1 Barcode Assay in particular, have been assessed in parallel with the historical isoenyzme analysis procedure that must now be replaced. It has been shown that specificity (the most critical attribute for an identity test) is much superior for CO1 barcoding relative to isoenzyme analysis. The ability of CO1 barcoding to serve as a cell line purity assessment is approximately equivalent to the ability of isoenzyme analysis. Both can detect inter-species cell line cross-contamination, albeit with limited sensitivity.

In conclusion, the CO1 Barcode Assay for animal cell identity testing clearly represents an improvement over the historical methods, including isoenzyme analysis, in terms of specificity. There is no reason to expect that the regulatory authorities will have any objection, therefore, to replacing isoenzyme analysis with the CO1 Barcode Assay for the identity component of cell line characterization.

6. References

1. Nims RW, et al. Sensitivity of isoenzyme analysis for the detection of interspecies cell line cross-contamination. In Vitro Cell Dev Biol Anim. 34, 35-39, 1998.
2. Nims RW and Herbstritt CJ. Cell line authentication using isoenzyme analysis. BioPharm Int 18:76-82, 2005.
3. US FDA Points to Consider in the Characterization of Cell Lines used to Produce Biologicals. 1993
4. International Council on Harmonisation (ICH) Q5D Derivation and Characterisation of Cell Substrates used for Production for Biotechnological/Biological Products.
5. World Health Organization (WHO) Recommendations for the evaluation of animal cell cultures as substrates for the manufacture of biological medicinal products and for the characterization of cell banks.
6. European Pharmacopoeia 5.2.3 Cell substrates for production of vaccines for human use
7. US FDA Characterization and Qualification of Cell Substrates and Other Biological Materials Used in the Production of Viral Vaccines for Infectious Disease Indications. 2010.
8. Barallon R, et al. Recommendation of short tandem repeat profiling for authenticating human cell lines, stem cells, and tissues. In Vitro Cell Dev Biol Anim 46:727-732, 2010.
9. Kerrigan L and Nims RW. Authentication of human cell-based products: the role of a new consensus standard. Regen Med. 6:255-260, 2011.
10. ANSI/ATCC ASN-0002-2012 Authentication of human cell lines: Standardization of STR profiling. ANSI, 2012.
11. Pakstis AJ, et al. SNPs for a universal individual identification panel. Human Genet. 127: 315-24; 2010.
12. BioReliance Development Report. Development of a PCR and Sequencing Assay that Targets a Semi-conserved Region of the Mitochondrial Cytochrome c Oxidase Subunit 1 (CO1) Gene for Species Level Identification of Mammalian and Insect Cells. 2015.
13. ANSI/ATCC ASN-0003 Species-Level Identification of Animal Cells through Mitochondrial Cytochrome c Oxidase Subunit 1 (CO1) DNA Barcodes. Currently under review by the ATCC SDO Steering Committee.
14. Hebert PD, et al. Identification of Birds through DNA Barcodes. PLoS Biol 2(10): p. e312., 2004
15. Savolainen V, et al. Towards writing the encyclopedia of life: an introduction to DNA barcoding. Philos Trans R Soc Lond B Biol Sci. 360(1462): p. 1805-11, 2005.
16. http://www.boldsystems.org/
17. BioReliance VPPO0309.R00 Validation Report for Cell Line Identification using Cytochrome Oxidase 1 Universal Primers by Sanger Sequencing. 2015.
18. Ziegenmeyer J. The AuthentiKit™ System Handbook for Cell Authentication and Identification. Innovative Chemistry, Inc., 1988.

For more information about cell identity testing and cell line characterization, visit:
bioreliance.com

BioReliance® by SAFC

RUA
O-1030815
REPORT NUMBER: RMC-TR15003 DATE: 05 AUGUST 2015

©2015 Sigma-Aldrich Co. LLC. All rights reserved. SAFC and SIGMA-ALDRICH are trademarks of Sigma-Aldrich Co. LLC, registered in the US and other countries. Purchaser must determine the suitability of the product(s) for their particular use. Additional terms and conditions may apply. Please see product information on the SAFC website at sigma-aldrich.com/safc and/or on the reverse side of the invoice or packing slip.