Species identification in cell culture: a two-pronged molecular approach

**Jason K. Cooper · Greg Sykes · Steve King ·
Karin Cottrill · Natalia V. Ivanova · Robert Hanner ·
Pranvera Ikonomi**

Abstract Species identification of cell lines and detection of cross-contamination are crucial for scientific research accuracy and reproducibility. Whereas short tandem repeat profiling offers a solution for a limited number of species, primarily human and mouse, the standard method for species identification of cell lines is enzyme polymorphism. Isoezymology, however, has its own drawbacks; it is cumbersome and the data interpretation is often difficult. Furthermore, the detection sensitivity for cross-contamination is low; it requires large amounts of the contaminant present and cross-contamination within closely related species may go undetected. In this paper, we describe a two-pronged molecular approach that addresses these issues by targeting the mitochondrial genome. First, we developed a multiplex PCR-based assay to rapidly identify the most common cell culture species and quickly detect cross-contaminations among these species. Second, for speciation and identification of a wider variety of cell lines, we amplified and sequenced a 648-bp region, often described as the "barcode region" by using a universal primer mix targeted at conserved sequences of the cytochrome C oxidase I gene (COI). This method was challenged with a panel of 67 cell lines from 45 diverse species. Implementation of these assays will accurately determine the species of cell lines and will reduce the problems of misidentification and cross-contamination that plague research efforts.

J. K. Cooper · G. Sykes · S. King · K. Cottrill · P. Ikonomi (✉)
ATCC (American Type Culture Collection),
10801 University Boulevard,
Manassas, VA 20110, USA
e-mail: pikonomi@atcc.org

N. V. Ivanova · R. Hanner
Biodiversity Institute of Ontario, University of Guelph,
Guelph, ON, Canada, N1G 2W1

Introduction

Researchers rely on cell lines as model systems for basic research, standards, and controls. A significant portion of this research is however misleading because cell lines are of a different origin from the one being claimed (Stacey 2000). Several reports have demonstrated evidence of interspecies and intraspecies contamination (Povey et al. 1976; Nelson-Rees et al. 1981). Most notably, cross-contamination and subsequent overgrowth of HeLa cells invalidated many cell culture-based experiments (Nelson-Rees et al. 1981; Masters 2002). In another example, a series of early reports on the establishment and characterization of Hodgkin's disease in human cell cultures was marred by misidentification; three cell lines proved to be nonhuman and, in fact, were derived from owl monkey (Harris et al. 1981). Therefore, as the awareness of the contamination or misidentification of cell lines increases (Chatterjee 2007), researchers have been implored to ensure the identity and the purity of their cell lines (Langdon 2004; Lincoln and Gabridge 1998; Markovic and Markovic 1998).

Emerging genetic methodology offers the most promise in species verification because these methods are easy to use and cost-effective. Analysis of short tandem repeats (STR), although limited to a small number of species origins, has become a valuable tool to establish the identity of cell lines. For interspecies identification of cell cultures, several PCR-based methods have recently been described (Parodi et al. 2002; Liu et al. 2003; Steube et al. 2003). However, the biochemical analysis of isoenzyme polymorphism, developed over 35 years ago, remains the most common test used, if any testing is performed at all.

Isoenzymology suffers from three major limitations. First, commercially available kits provide identification information for only 20 species. Although this range can be extended, difficult data interpretation makes it challenging to distinguish closely related species. Second, isoenzymology is not used at many of the research laboratories that work with cell lines. For example, nearly 90% of polled active cell culture workers reported that they had never used isoenzymology to assess cell line purity (Buehring et al. 2004). Finally, isoenzymology lacks sufficient sensitivity; cross-contamination is only detected when the contaminating cell lines comprise at least 25% of the total cell population (Hay et al. 1992). A genetic approach addresses all of these issues: PCR is both simple and extremely sensitive. Furthermore, by selecting the correct target gene, a genetic approach is capable of identifying an enormous species range with fine resolution.

Conserved mitochondrial protein coding genes represent excellent targets for species identification (Hebert et al. 2003a). Hebert et al. (2003b) have demonstrated that the first 648-bp of the cytochrome C oxidase subunit I (COI) gene contain substantial interspecies variation whereas intraspecies variation remains remarkably low in most animals. Using this approach, often referred to as "DNA barcoding", a wide array of field specimens has been accurately identified (Hebert et al. 2003a; 2004; Smith et al. 2005; Cywinska et al. 2006; Kerr et al. 2007). Based on these findings, a major international initiative called the Consortium for the Barcode of Life (CBOL; http://barcoding.si.edu/) was formed. CBOL advocates the use of this region of COI as a universal barcode system for the genetic identification of all animal life.

In this paper, we suggest a two-pronged approach for species identification and detection of cross-contamination. First, we used the COI and cytochrome B mitochondrial gene sequences as targets for species-specific PCR amplification. For this, primers are designed to function in a multiplex PCR assay and to generate a size-specific amplicon for each of the species detected. This assay provides a rapid, simple, and cost-effective method for species identification and detecting cross-contamination.

Second, for identification of a wider variety of cell lines and for making finer distinctions between closely related species, we adapted the COI barcode method (Lorenz et al. 2005). We developed a common platform to identify a broad cell line panel comprising mammals, fish, birds, amphibians, and insects. Our method compares barcode sequences from cell lines to reference sequences derived from expert-identified voucher specimens.

Materials and Methods

Template preparation. Three types of DNA templates were used throughout this study: cell lysates, purified genomic DNA, and cultured cells dried onto FTA cards.

Cell lysates: 10^3–10^6 cultured cells were harvested and centrifuged for 3 min at $13,000 \times g$. The supernatant was discarded and the pellet resuspended in 100μl of lysis buffer containing: 40 mM Tris acetate pH 7.6, 1 mM EDTA, 0.5% Igepal CA-630 (nonionic detergent). The lysate was incubated for 15 min at 37° C in a heat block, followed by 10 min at 95° C to inactivate proteinase K. After spinning down the lysate for 5 mins at $13,000 \times g$, 5μL of the supernatant was used as a template for PCR.

Purified genomic DNA: DNA was extracted from 10^6 cells using the UltraClean Tissue DNA kit (MoBio, Carlsbad, CA, USA). These cells include ATCC® numbers: CCL-1™, CCL-60™, CRL-1430™, CRL-2032™, CRL 1601™, CL-101™, CCL-81™, CCL-2™, CCL-57™, CRL-1633™, CCL-209™, CRL-6306™, CCL-73™, and CCL-39™ (ATCC, Manassas, VA, USA). K562 was also used (Promega, Madison, WI, USA). 1μL of DNA was used as a template for PCR.

Cell cultures dried in FTA cards: Sixty-seven cell lines used for barcode analysis (Table 1) were expanded under optimal conditions to a density of 10^5–10^7 cells/ml and frozen as glycerol stocks. Twenty microliters of cell culture was applied to FTA cards (Whatman, Middlesex, UK), dried for 1 h at room temperature, and stored at room temperature. Samples sent to the University of Guelph were transported by conventional air transportation. Before PCR, a sample was removed from each card with a 2-mm Harris punch. Punches were washed three times with 200μL of FTA reagent and once with 200μL TE pH 8.0, according to the manufacturer's instructions. The punches were dried for 1 h and used directly for PCR.

Oligonucleotide design multiplex assay. Species-specific primer sets were designed to amplify a specifically sized product only in the presence of the target species. The primer sets used in this assay include five that were either fully or partially designed by Parodi et al. (2002) and nine sets designed *de novo* for this study. Primers were designed with Oligo 6 (Molecular Biology Insights, Cascade, CO) and synthesized by Integrated DNA Technologies (IDT, Coralville, IA). A complete list of the oligonucleotide sequences is shown in Table 2. The oligonucleotides were used at specific final concentrations in the multiplex as described in Table 2.

DNA COI barcode: To amplify the 648-bp COI barcode region, previously published primers were used (Hebert et al. 2004; Ward et al. 2005; Ivanova

Table 1. Cell lines used for barcode experiment

ATCC® No.	Identification	Common name		ATCC® No	Identification	Common name	
CCL-126™	*Aedes albopictus*	mosquito, Asian Tiger		CRL-2756™	*Ictalurus punctatus*	channel catfish	
CCL-141™	*Anas platyrhynchus domesticus*	duck, Pekin	x	CCL-108™	*Iguana iguana*	iguana	
CRL-1556™	*Aotus trivirgatus*	monkey, Owl	x	CRL-6299™	*Macaca mulatta*	monkey, rhesus	x
CCL-22™	*Bos taurus*	cow	x	CRL-6300™	*Macaca mulatta*	monkey, rhesus	
CRL-1395™	*Bos taurus*	cow		CCL-208™	*Macaca mulatta*	monkey, rhesus	
CRL-2048™	*Bos taurus*	cow	x	CRL-8135™	*Meleagris gallopavo*	turkey	x
CRL-1926™	*Calloscirus notatus*	plantain squirrel	x	CRL-1632™	*Mesocricetus auratus*	hamster, Syrian golden	
CRL-1430™	*Canis familiaris*	dog		CRL-2279™	*Mus musculus*	mouse	x
CCL-34™	*Canis familiaris*	dog		CCL-1™	*Mus musculus*	mouse	
CCL-71™	*Carassius auratus*	goldfish	x	TIB-68™	*Mus musculus*	mouse	
CRL-1405™	*Cavia porcellus*	guinea pig	x	CRL-2674™	*Mustela vison*	mink	x
CRL-1587™	*Cercopithecus aethiops*	monkey, African green		CRL-6193™	*Odocoileus hemionus hemionus*	Columbian black tail deer	
CCL-70™	*Cercopithecus aethiops*	monkey, African green		CRL-2301™	*Oncorhynchus mykiss*	trout, rainbow	
CRL-1650™	*Cercopithecus aethiops*	monkey, African green		CRL-1681™	*Oncorhynchus tshawytscha*	salmon, Chinook	x
CRL-2750™	*Clupea pallasi*	Pacific herring		CCL-60™	*Oryctolagus cuniculus*	rabbit, European domestic	x
CRL-1708™	*Coturnix coturnix japonica*	quail, Japanese	x	CRL-6497™	*Oryctolagus cuniculus*	rabbit, European domestic	
CRL-2532™	*Coturnix coturnix japonica*	quail, Japanese		CRL-1700™	*Ovis aries*	sheep	x
CRL-11397™	*Cricetulus griseus*	hamster, Chinese	x	CRL-1868™	*Pan troglodytes*	chimpanzee	
CCL-93™	*Cricetulus griseus*	hamster, Chinese		CRL-6314™	*Pan troglodytes*	chimpanzee	
CCL-39™	*Cricetulus griseus*	hamster, Chinese		CRL-1495™	*Papio hamadryas*	baboon, African	x
CCL-195™	*Cricetulus migratorius*	hamster, Armenian	x	CRL-6488™	*Pecacari tajacu*	peccary	
CRL-2298™	*Danio rerio*	zebrafish		CRL-1850™	*Pongo pygmaeus*	orangutan	x
CRL-2296™	*Danio rerio*	zebrafish		CCL-74™	*Procyon lotor*	raccoon	
CRL-6009™	*Dasypus novemcinctus*	armadillo, nine-banded		CCL-192™	*Rattus norvegicus*	rat	x
CRL-10191™	*Drosophila melanogaster*	fruit fly	x	CRL-2256™	*Rattus norvegicus*	rat	x
CRL-1963™	*Drosophila melanogaster*	fruit fly	x	CRL-1721™	*Rattus norvegicus*	rat	
CCL-57™	*Equus caballus*	horse	x	CCL-194™	*Saimiri boliviensis peruviensis*	monkey, bolivian squirrel	x
CRL-2032™	*Felis catus*	cat		CRL-1711™	*Spodoptera frugiperda*	fall armyworm	x
CRL-2112™	*Gallus gallus*	chicken		CL-101™	*Sus scrofa*	pig	
CRL-2118™	*Gallus gallus*	chicken	x	CCL-33™	*Sus scrofa*	pig	
CRL-2117™	*Gallus gallus*	chicken		CRL-2641™	*Takifugu rubripes*	torafugu	x
CRL-2522™	*Homo sapiens*	human		CRL-6001™	*Tamandua tetradactyla*	anteater	
				CCL-168™	*Urocyon cinereoargenteus™*	grey fox	
				CCL-102™	*Xenopus laevis™*	frog, South African clawed	

[a] K-562 (*Homo sapiens*) from Promega (Madison, WI) was also used. An 'x' in the shaded column shows cell lines sequenced at both ATCC and Guelph.

et al. 2006). Primers were synthesized at IDT. 1μL of the following 10μM primer mixes were used:

Forward mix (C_VF1di):

1 part VF1: TTCTCAACCAACCACAAAGACATTGG,
1 part VF1d: TTCTCAACCAACCACAARGAYATYGG, and
2 parts VF1i: TTCTCAACCAACCAIAAIGAIATIGG.
Reverse mix (C_VR1di):

1 part VR1: TAGACTTCTGGGTGGCCAAAGAATCA,
1 part VR1d: TAGACTTCTGGGTGGCCRAARAAYCA,
2 parts VR1i: TAGACTTCTGGGTGICCIAAIAAICA

For insect cell lines, the following primers were used:

LepF1: ATTCAACCAATCATAAAGATATTGG,
LepR1: TAAACTTCTGGATGTCCAAAAAATCA.

PCR amplification. The PCR buffer consisted of: 20 mM Tris–Cl at pH 8.4, 50 mM KCl, 2.0 mM MgCl$_2$, 5 mM dNTPs, 0.5% glycerol, 0.006% NP40/Tween (1:1), 0.5 U Platinum *Taq* (Invitrogen, Carlsbad, CA), and molecular grade water (ATCC cat no. 60–2450) to make a 50μL total reaction.

The thermocycling conditions were as follows:

Multiplex cycling conditions: One cycle of 95° C for 5 min; 30 cycles of 95° C for 30 s, 60° C for 15 s, 72° C for 30 s; 1 cycle of 72° C for 7 min; and indefinite hold at 4° C.
DNA barcode cycling conditions: One cycle of 95° C for 5 min; 30 cycles of 95° C for 30 s, 45° C for 15 s, 72° C for 30 s; 1 cycle of 72° C for 7 min; and indefinite hold at 4° C.

PCR products were visualized on 4% (multiplex) or 2% (barcode) precast gels stained with ethidium bromide (Cambrex, East Rutherford, NJ).

Sequence analysis. A QIAquick PCR purification kit was used to clean the PCR products (QIAGEN, Hilden, Germany). VF1d and VR1d primers were used as sequencing primers. For insect, cell lines LepF1 and LepR1 were used. Sequencing was done with a CEQ 8000 genetic analyzer following manufacturer's instructions (Beckman

Table 2. Oligonucleotide primer sequences

Size (bp)	Species	Final concentration	Forward primer Name	Sequence	Reverse primer Name	Sequence
460	*Sus scrofa* (pig)	200 nM	Ss-F	CT ACT ATC CCT GCC AGT T	Ss-R	GAA TAG GAA GAT GAA GCC C
391	*Homo sapiens* (human)	100 nM	Hs-F	TAG ACA TCG TAC TAC ACG ACA CG	Hs-R	TCC AGG TTT ATG GAG GGT TC
341	*Felis catus* (cat)	200 nM	Fc-F	TAT TGC CAT TCC TAC CGG GGT G	Fc-R	ACG TTA TAT TGA CTC CTA CAA ACA TAA TC
315	*Cricetulus griseus* (Ch. Hamster)	200 nM	Cg-F	ACTAACCCGCTTCTTCGCATTC	Cg-R	GCG TAG GCG AAC GGA AGT ATC
287	*Macaca mulatta* (Rhesus monkey)	400 nM	Mmul-F	CCCACCCAGTTCAACTAAGC	Mmul-R	AATGGTGAAGGATGGGTCG
267	*Ovis aris* (sheep)	200 nM	Oa-F	CGA TAC ACG GGC TTA CTT CAC G	Oa-R	AAA TAC AGC TCC TAT TGA TAA T
243	*Equus caballus* (horse)	200 nM	Ec-F	CTG*CCCTAAGCCTCCTAAT*	Ec-R	AGAAGT*AGGAATGATGGGGG*
222	*Cercopithecus aethiops* (Gr. Monkey)	400 nM	Ca-F	CTT*CTTTCCTGCTGCTAATG*	Ca-R	*TTTGATACTGGGATATGGCG*
196	*Rattus norvegicus* (rat)	100 nM	Rn-F	*CGGCCACCCAGAAGTGTACATC*	Rn-R	*GGCTCGGGTGTCTACATCTAGG*
172	*Canis familiaris* (dog)	200 nM	Cf-F	*GAACTAGGTCAGCCCGGTACTT*	Cf-R	TTCGGGGGAATGCCATGTC
150	*Mus musculus* (mouse)	140 nM	Mmus-F	*ATTACAGCCGTACGCTCCTAT*	Mmus-R	*CCCAAAGAATCAGAACAGATGC*
136	*Oryctolagus cuniculus* (rabbit)	200 nM	Oc-F	CGCC TAT ACA ATA TGA AAT ACT GTT	Oc-R	TGTGG TTG TTA GTT CAA TAG TCT
117	*Capra hircus* (goat)	200 nM	Ch-F	ATA TCA ATC GGG TTT CTA GGA TTT ATT	Ch-r	AGT TGG GAT AGC GAT AAT TAT GGT AGT
102	*Bos taurus* (cow)	200 nM	Bt-F	GCTATTCC AAC CGG GGT AAA AGT C	Bt-R	GAAAAT AAA GCC TAG GGC TCA C
70	IC	40 nM	IC-F	CGG GGA ATY AGG GTT CGA TTC	IC-R	GCC TGC TGC CTT CCT TKG ATG

[a] Italicized primer sequences are from Parodi et al. (2002).

Coulter, Fullerton, CA) except that Performa® DTR plates were used for dye terminator removal (Edge BioSystems, Gaithersburg, MD). Sequences were analyzed for quality and trimmed using CodonCode Aligner (Dedham, MA). Overlapping forward and reverse sequences were then assembled.

Sequencing at the University of Guelph was performed on a 3730 DNA Analyzer (Applied Biosystems/Hitachi, Foster City, CA) with BigDye terminator v3.1 (Applied Biosystems). Bidirectional sequences were assembled in SeqScape v. 2.1.1 (Applied Biosystems) and manually edited. At the Guelph facility, different primer sets and reaction conditions were used as published previously; for birds (Kerr et al. 2007), for fish (Ward et al. 2005; Ivanova et al. 2007) using M13-tailed cocktails (Ivanova et al. 2007), and for insects (Folmer et al. 1994; Hebert et al. 2004).

All sequence results were stored and analyzed at the Barcode of Life Data (BOLD) systems identification engine (Ratnasingham and Herbert 2007).

Results

Multiplex assay. To verify the identity of cell lines commonly used in research laboratories, we developed a rapid PCR-based assay with species-specific primer sets for the detection of 14 species. Species-specific primers have been developed recently by several groups (Parodi et al. 2002; Liu et al. 2003; Steube et al. 2003). Our primers, however, differ in that they are designed to generate amplicons of distinct size between species most commonly used in basic and applied research. The primers are also

designed to function in a multiplex PCR assay, using amplicon size to distinguish between species. In this way, the size of the amplified product becomes a signature for the presence of a particular species in a sample (Table 2).

We then combined the primer sets into a multiplex assay (described in the "Materials and Methods" section) and challenged the assay with DNA extracts from all 14 species for which specific primers were designed. Analysis was performed on each extract individually as well as on the pooled extracts (Fig. 1a, b). In all cases, detection revealed the expected species signature bands. For increased sensitivity, we recommend using smaller, targeted, combinations of these 14 primer sets. For example, a laboratory that maintains cell lines from five or six species could target those specific species (Fig. 1c). We found that any combination of the described primers can be used successfully in a multiplex assay (data not shown).

In a PCR assay, the presence of controls is necessary as numerous factors such as inactive polymerase enzyme or improper cycling conditions can render a PCR ineffectual. Therefore, a negative result has little diagnostic meaning. The template quality and preparation method is also important for adequate assay functionality. We designed two specific oligonucleotides to amplify a conserved region in the 18S rRNA gene. These primers serve as an internal control and produce a 70-bp band (Fig. 1a, b, c).

To test our method in detecting low-level cross-contamination, we created cell mixes from mouse and human cell lines as they represent extremely common cell lines for laboratory research. Mixed templates were prepared by mixing a human cell line (ATCC® CCL-2™) and a mouse cell line (ATCC® CCL-1™) at fixed ratios. A total of 11 cell mixtures were prepared, each containing 1×10^6 total cells as starting material but consisting of different ratios
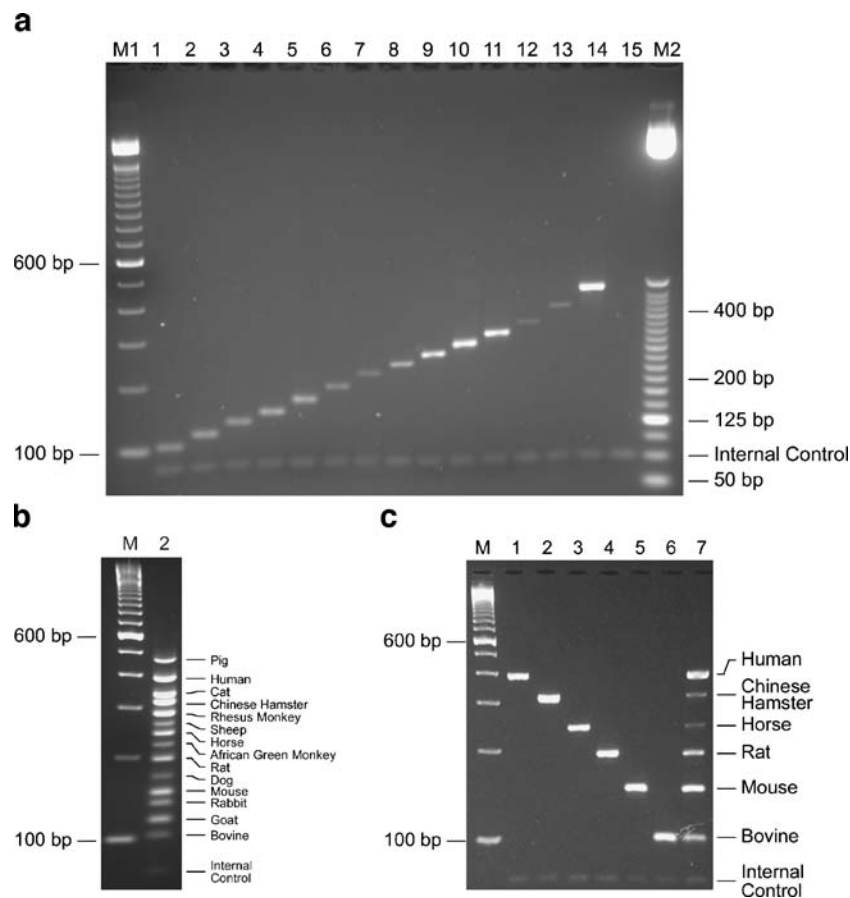

Figure 1. Detection of 14 different species by the multiplex PCR assay. A mix of 1 ng of purified DNA from all the 14 species for which primer sets have been designed is used as a template (a) Detection of species-specific amplified products in a monoplex PCR assay. M: 100 bp DNA ladder. 1–14: Pig, human, cat, Chinese hamster, Rhesus monkey, sheep, horse, African green monkey, rat, dog, mouse, rabbit, goat, and bovine. M2: 25 bp DNA ladder. 15: Internal control. (b) Detection of species-specific amplified products in a multiplex PCR assay. M: 100 bp DNA ladder. 2: Mixed DNA template. Species signature bands for all species are seen descending by size. (c) Targeted mix of primers used in multiplex PCR assay. A subset of seven primers (human, Chinese hamster, horse, rat, mouse, bovine, and internal control) is used in a multiplex PCR assay. M: 100 bp DNA ladder. 1–6: Purified DNA of each of the six species used as positive controls. 7: Mixed DNA template from all the 14 species.
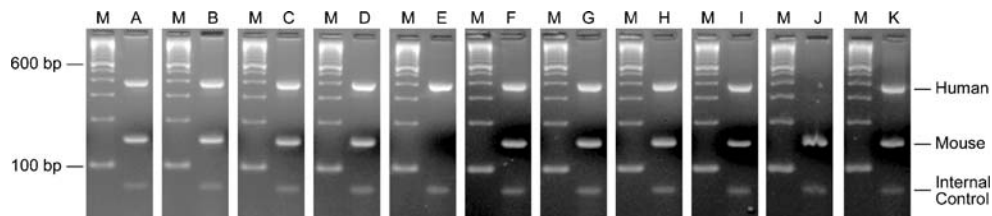
Figure 2. Detection of interspecies contamination using small ratios of contaminating cells. Mouse cell line (ATCC® CCL-1™) and human cell line (ATCC® CCL-2™) were mixed together at fixed ratios designed to mimic cross-contamination (lanes A–K). M: 100 bp DNA ladder shown in each lane. The human and mouse template mixtures vary across lanes as follows: (A) human/mouse 1:99, (B) human/mouse 5:95, (C) human/mouse 10:90, (D) human/mouse 20:80, (E) human/mouse 100:0, (F) human/mouse 80:20, (G) human/mouse 90:10, (H) human/mouse 95:5, (I) human/mouse 99:1, (J) human/mouse 0:100, (K) human/mouse 50:50.

of human and mouse cells. For this assay, only a subset of primer sets from six species (bovine, mouse, rat, horse, Chinese hamster, and human) and the internal control was used (Fig. 2; this multiplex is shown in Fig. 1c).

The signature band for human at 391 bp is seen in all panels with one exception, which corresponds to a sample of only mouse cells (Fig. 2 lane J). When human cells were mixed in a ratio of 1:99 in a mouse cell background, the presence of human DNA was detected (Fig. 2 lane A). Likewise, we detected a signature band of 150 bp for mouse in all panels with one exception, which contained only human cells (Fig. 2 lane E). Mouse cells mixed in a ratio of 1:99 in a human cell background was easily detected (Fig. 2 lane I). The internal control can be seen at 70 bp in all panels. We detected no nonspecific products caused by misannealing of oligonucleotides.

COI barcode analysis. For the identification of cell lines from a larger variety of species, we tested the use of the COI barcode method. For this, we sequenced a 648-bp region of the COI gene, previously recognized as the barcode region (Folmer et al. 1994; Hebert et al. 2003a, b; 2004; Cywinska et al. 2006), from a panel of 67 cell lines comprising 45 unique species (Table 1). Many of these cell lines came from less common species not detected by our multiplex assay. Others, from more common species,
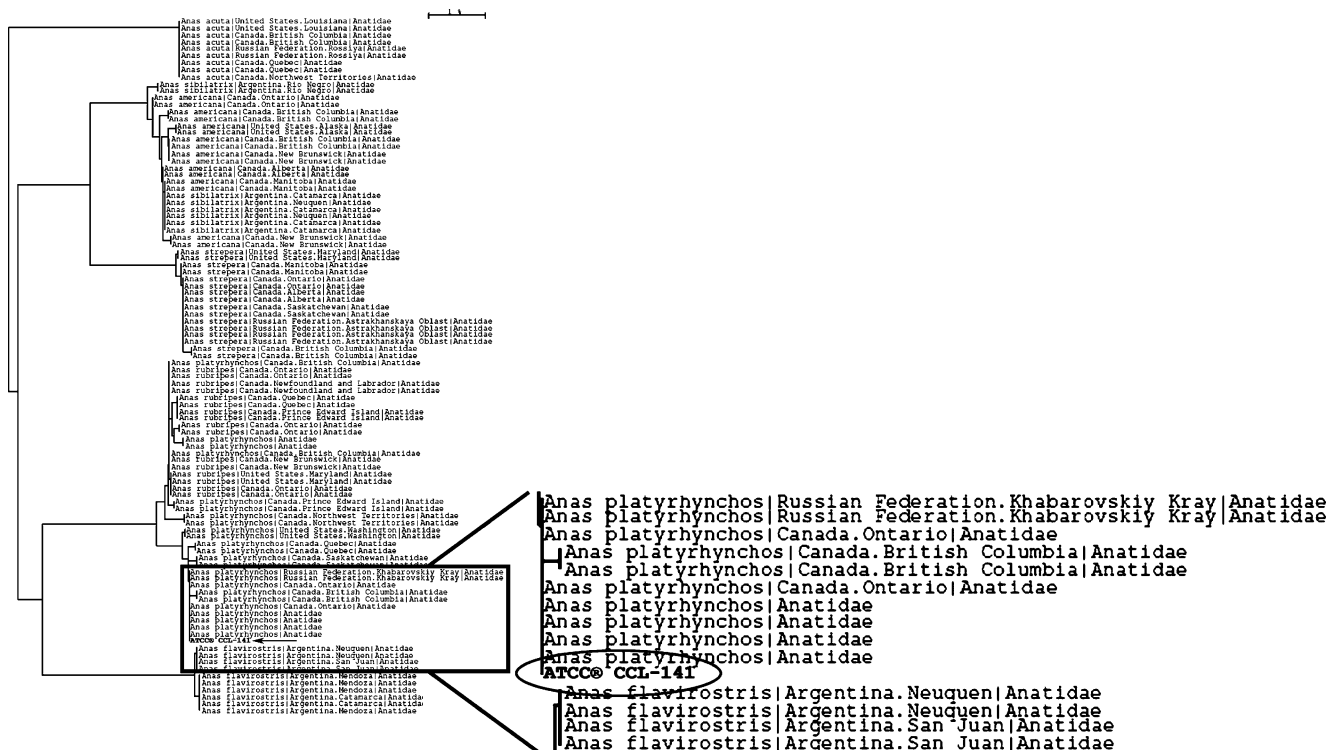

Figure 3. Definitive species identification of cell culture by COI barcodes. The NJ tree shows how sequence data from cell line CCL-141™ (*Anas platyrhynchus*) compares with sequence data obtained from closely related field tissue specimens. The species identification engine at the BOLD systems database at the University of Guelph generated the tree. (http://www.barcodinglife.com/).

served as positive controls for the barcode method. By optimizing the conditions for PCR, we were able to generate all necessary COI sequences with the mix of primers described in the "*Materials and Methods*" section. This mix contains one set of universal primers and another set specific for identification of cell lines of insect origin. Combined with our use of a common cycling condition, this assay is adaptable to a 96-well format.

We identified all 67 cell lines using the BOLD (Ratnasingham and Herbert 2007) systems animal identification engine (http://www.barcodinglife.com/) hosted by the Canadian Centre for DNA Barcoding at the Biodiversity Institute of Ontario. Figure 3 illustrates this method for one cell line in our sample (ATCC® CCL-141™), from duck embryo fibroblasts (*Anas platyrhynchos*) (Marcovici and Prior 1968), which has been used in numerous experiments (Wolf et al. 1974; Alexander et al. 1998; Sick et al. 1998; Farris et al. 1999). A neighbor-joining (NJ) tree generated by BOLD shows how the sequence from this cell line clusters with sequences from tissue samples from various *Anus* species (Fig. 3). In this way, BOLD becomes the portal linking the cell biologist and the field taxonomist. BOLD currently contains 209,657 barcode sequences covering 26,209 different species. As the database grows, so will its sensitivity and reach.

To test for the reproducibility and robustness of the COI barcode method as a tool for cell line verification, extracted DNA from 28 cell lines stored on Whatman FTA® cards was sent to the University of Guelph for sequence analysis (Table 1; shaded column). The barcode sequences generated at both facilities show 100% identity correlation.

Discussion

We have developed a simple, sensitive, and rapid PCR assay that can be used to efficiently identify common cell cultures. As most eukaryotic culture systems offered by major culture collections are from human, mouse, rat, or Chinese hamster, the described multiplex assay can greatly reduce the risk of most interspecies contamination in a single PCR reaction and is therefore recommended for labs frequently using cell lines in their research. For identification and authentication of a wider variety of cell lines, our multiplex assay described in this paper is complemented by COI barcode sequence analysis. By employing these assays, species identification of cell cultures becomes a rapid, cost-effective routine procedure, and it overcomes the limitations of the currently used methodology.

Although major culture collections routinely test for interspecies cross-contamination, many cell lines are created in individual labs and are shared without thorough testing.

This creates a large-scale potential problem. Many published articles on cell biology may be based on cell lines that have not been properly validated. In one estimate, Charles Patrick Reynolds of the University of Southern California and the Children's Hospital Los Angeles' Institute for Pediatric Clinical Research predicts that fully 35% to 40% of published cell biology papers would have to be retracted because of invalid data (Chatterjee 2007).

One of the main weapons in combating this problem is short tandem repeat (STR) DNA fingerprinting. STR provides a unique identifier of a cell line and is used to distinguish one cell line from another within a species. STR is most commonly used for distinguishing between human cell lines with applications extending from forensics, paternity testing, cell culture, and others. STR has also been used to identify cell lines within other common species from which many cell lines derive such as mouse. However, STR cannot distinguish between species. Thus, the assays described in this paper for species identification are essential complements to STR providing a fuller picture of the identity of a cell line.

In fact, this complement with STR extends to the practical level at the lab bench. As both STR and the assays described in this paper are PCR-based, they can share a common pipeline in the lab. At ATCC, cells lines spotted onto FTA cards are used as templates for the multiplex assay, COI barcoding, and STR. Thus, the DNA is collected, archived, and could be easily isolated before the assay. Should a question ever arise about authenticity or contamination, the DNA can be retrieved and analyzed by these PCR-based assays to help pinpoint when and how the problem occurred.

As awareness of the scope of this problem of cell line authenticity grows, the availability of a simple, economic testing platform for species-identification to complement STR, such as the one described in this paper, is important to meet that demand.

In the near future, cell lines will likely play a role in the preservation of DNA from endangered species (Ryder et al. 2000). By embracing the use of the COI locus as an identification system, cell lines could be employed, beyond their traditional application as disease models, as safeguards providing a renewable source of DNA standards for these endangered animals. Such standards could be used to resolve a number of molecular questions and to determine evolutionary relationships between species even after extinction.

References

Alexander, D.S.; Sipley, J.D.; Quigley, J.P. Autoactivation of avian urokinase-type plasminogen activator (uPA). A novel mode of initiation of the uPA/plasmin cascade. J Biol Chem 273 (13):7457–61; 1998.

Buehring, G.C.; Eby, E.A.; Eby M.J. Cell line cross-contamination: how aware are Mammalian cell culturists of the problem and how to monitor it? In Vitro Cell Dev Biol Anim 40(7):211–5; 2004.

Chatterjee, R. Cell biology. Cases of mistaken identity. Science 315 (5814):928–31; 2007.

Cywinska, A.; Hunter, F.F.; Hebert, P.D. Identifying Canadian mosquito species through DNA barcodes. Med Vet Entomol 20 (4):413–24; 2006.

Farris, A.D.; Koelsch, G.; Pruijn, G.J.; van Venrooij, W.J.; Harley, J.B. Conserved features of Y RNAs revealed by automated phylogenetic secondary structure analysis. Nucleic Acids Res 27 (4):1070–8; 1999.

Folmer, O.; Black, M.; Hoeh, W.; Lutz, R.; Vrijenhoek, R. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol Mar Biol Biotechnol 3(5):294–9; 1994.

Harris, N.L.; Gang, D.L.; Quay, S.C.; Poppema, S; Zamecnik, P.C.; Nelson-Rees, W.A.; O'Brien, S.J. Contamination of Hodgkin's disease cell cultures. Nature 289(5795):228–30; 1981.

Hay, R.J.; Chen, T.J.; Macy, M.L.; Reid, Y.A. Replay to "Cell, lines and DNA fingerprinting". In Vitro Cell Dev Biol 28A:593–594; 1992.

Hebert, P.D.; Cywinska, A.; Ball, S.L.; deWaard, J.R. Biological identifications through DNA barcodes. Proc Biol Sci 270 (1512):313–321; 2003a

Hebert, P.D.; Ratnasingham, S.; deWaard, J.R. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Proc Biol Sci 270 Suppl 1:S96–9; 2003b.

Hebert, P.D.; Penton, E.H.; Burns, J.M.; Janzen, D.H.; Hallwachs, W. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly Astraptes fulgerator. Proc Natl Acad Sci U S A 101(41):14812–7; 2004

Ivanova, N.V.; deWaard, J.R.; Hebert, P.D.N. An inexpensive, automation-friendly protocol for recovering high-quality DNA. Molecular Ecology Notes 6:992–1002; 2006.

Ivanova, N.V.; Zemlak, T.S.; Hanner, R.H.; Hebert, P.D.N. Universal primer cocktails for fish DNA barcoding. Molecular Ecology Notes 7(4):544–548; 2007.

Kerr, K.C.R.; Stoeckle, M.Y.; Dove, C.J.; Weight, L.A.; Francis, C.M.; Hebert, P.D.N. Comprehensive DNA barcode coverage of North American birds. Mol Ecol Notes 7(4):535–543; 2007.

Langdon, S.P. Cell culture contamination: an overview. Methods Mol Med 88:309–17; 2004.

Lincoln, C.K.; Gabridge M.G. Cell culture contamination: sources, consequences, prevention, and elimination. Methods Cell Biol 57:49–65; 1998.

Liu, M.Y.; Lin, S.C.; Liu, H.; Candal, F.; Vafai, A. Identification and authentication of animal cell culture by polymerase chain reaction amplification and DNA sequencing. In Vitro Cell Dev Biol Anim 39(10):424–7; 2003.

Lorenz, J.G.; Jackson, W.E.; Beck, J.C.; Hanner, R. The problems and promise of DNA barcodes for species diagnosis of primate biomaterials. Philos Trans R Soc Lond B Biol Sci 360 (1462):1869–77; 2005.

Marcovici, M.; Prier, J.E.. Enhancement of St. Louis arbovirus plaque formation by neutral red. J Virol 2(3):178–81; 1968.

Markovic, O.; Markovic, N. Cell cross-contamination in cell cultures: the silent and neglected danger. In Vitro Cell Dev Biol Anim 34 (1):1–8; 1998.

Masters, J.R. HeLa cells 50 years on: the good, the bad and the ugly. Nat Rev Cancer 2(4):315–9; 2002.

Nelson-Rees, W.A.; Daniels, D.W.; Flandermeyer, R.R. Cross-contamination of cells in culture. Science 212(4493):446–52; 1981.

Parodi, B.; Aresu, O.; Bini, D.; Lorenzini, R.; Schena, F.; Visconti, P.; Cesaro, M.; Ferrera, D.; Andreotti V.; Ruzzon T. Species identification and confirmation of human and animal cell lines: a PCR-based method. Biotechniques 32(2):432–4, 436, 438–40; 2002.

Povey, S.; Hopkinson, D.A.; Harris, H.; Franks, L.M. Characterisation of human cell lines and differentiation from HeLa by enzyme typing. Nature 264(5581):60–3; 1976.

Ratnasingham, S.; Hebert, P. The Barcode of Life Database. Mol. Ecol. Notes 7(3):355–364; 2007.

Ryder, O.A.; McLaren, A.; Brenner, S.; Zhang, Y.P.; Benirschke, K. DNA banks for endangered animal species. Science 288 (5464):275–7; 2000.

Sick, C.; Schultz, U.; Munster, U.; Meier, J.; Kaspers, B.; Staeheli, P. Promoter structures and differential responses to viral and nonviral inducers of chicken type I interferon genes. J Biol Chem 273(16):9749–54; 1998.

Smith, M.A.; Fisher, B.L.; Hebert, P.D. DNA barcoding for effective biodiversity assessment of a hyperdiverse arthropod group: the ants of Madagascar. Philos Trans R Soc Lond B Biol Sci 360 (1462):1825–34; 2005

Stacey, G.N. Cell contamination leads to inaccurate data: we must take action now. Nature 403(6768):356; 2000.

Steube, K.G.; Meyer, C.; Uphoff, C.C.; Drexler, H.G. A simple method using beta-globin polymerase chain reaction for the species identification of animal cell lines—a progress report. In Vitro Cell Dev Biol Anim 39(10):468–75; 2003.

Ward, R.D.; Zemlak, T.S.; Innes, B.H.; Last, P.R.; Hebert, P.D. DNA barcoding Australia's fish species. Philos Trans R Soc Lond B Biol Sci 360(1462):1847–57; 2005.

Wolf, K.; Burke, C.N.; Quimby, M.C. Duck viral enteritis: microtiter plate isolation and neutralization test using the duck embryo fibroblast cell line. Avian Dis 18(3):427–34; 1974.