JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Article

pubs.acs.org/JAFC

Pyrosequencing as a Tool for Rapid Fish Species Identification and Commercial Fraud Detection

Cristian De Battisti,*[,†] Sabrina Marciano,[†] Cristian Magnabosco,[‡,⊥] Sara Busato,[†] Giuseppe Arcangeli,[‡] and Giovanni Cattoli[†]

[†]Research & Innovation Department, Istituto Zooprofilattico Sperimentale delle Venezie, Viale dell'Università 10, 35020 Legnaro, Padova, Italy

[‡]Istituto Zooprofilattico Sperimentale delle Venezie, Via L. da Vinci, 39, 45011 Adria, Rovigo, Italy

Supporting Information

ABSTRACT: The increased consumption of fish products, as well as the occurrence of exotic fish species in the Mediterranean Sea and in the fish market, has increased the risk of commercial fraud. Furthermore, the great amount of processed seafood products has greatly limited the application of classic identification systems. DNA-based identification allows a clear and unambiguous detection of polymorphisms between species, permitting differentiation and identification of both commercial fraud and introduction of species with potential toxic effects on humans. In this study, a novel DNA-based approach for differentiation of fish species based on pyrosequencing technology has been developed. Raw and processed fish products were tested, and up to 25 species of fish belonging to *Clupeiformes* and *Pleuronectiformes* groups were uniquely and rapidly identified. The proper identification based on short and unique genetic sequence signatures demonstrates that this approach is promising and cost-effective for large-scale surveys.

KEYWORDS: *Clupeiformes, Pleuronectiformes, commercial fraud, fish species identification, pyrosequencing*

INTRODUCTION

The consumption of fish products in Europe has increased in recent decades mainly because of the growing awareness of the importance of a healthy diet and the consequent demand for new sources of healthy food.[1] This factor, in combination with the globalization of the fish markets, has contributed to the introduction of novel fish species in the European fish markets and to an increased risk of commercial and/or sanitary fraud. In this regard, a commercial fraud consists of the illegal substitution of one species with another.[2] These economic corruptions can greatly affect the seafood commerce. Considering an estimated conservative worldwide substitution rate of 10%, the economic losses for worldwide fisheries is about US$ 24 billion/year.[3] For instance, in the case of red snapper (*Lutjanus campechanus*) fraudulent substitutions can account for three-quarters of the fish sold,[4] whereas for fish products sold as Alaska pollock (*Theragra chalcogramma*) more than 80% of the analyzed samples were prepared with species different from the one indicated on the label.[5] In the Mediterranean sea, high levels of market mislabeling were reported in different countries, to quote but a few, hake in Spain and Greece[6] and cod in Italy.[7] Besides the economic impact, these substitutions could also introduce species with potential toxic effects on humans.[8] All these issues combined with the increase in seafood consumption may well explain why it is mandatory to assign a unique market name to a given fish species in order to be unambiguously identified.[1,8] The European Union (EU) regulation 104/2000 imposes that commercial name, method of production (Directive 2200/13/EC), and capture area (EU Commission Regulation No 2065/2001) must be clearly indicated on the product label prior to its submission to the commercial circuit.[9] Despite these regulations, the process of labeling can still have some drawbacks.[10]

Identification of fish species is traditionally based on external morphological features (i.e., body shape, pattern of colors, position of fins, and number); otoliths count and shape analysis are also occasionally used. However, these identification systems are not considered useful or reliable for processed products subject to filleting, beheading, and/or skinning due to the obvious lack of morphological distinct features.[11] In more recent times, methods based on the separation and characterization of specific proteins using electrophoretic techniques, such as isoelectric focusing (IEF),[12] have been officially recognized and adopted to identify and assign species to fish products characterized by poor morphological integrity.[13] Some processing procedures often remove or damage key species signatures, as might be the case for proteins denatured by cooking,[13,14] making identification of species by morphological or protein-based taxonomic means extremely difficult if not impossible. The percentage of processed seafood related to the total seafood consumed in Europe can be relevant, for instance, up to 80% in the German market according to Horstkotte and Rehbein,[15] or to about 50% in Italy (ISMEA, 2007, http://www.ismea.it/flex/cm/pages/ServeBLOB.php/L/IT/IDPagina/2756). Taking these data into account, traditional

Received: August 9, 2013
Revised: November 19, 2013
Accepted: December 11, 2013

   dx.doi.org/10.1021/jf403545m | J. Agric. Food Chem. XXXX, XXX, XXX−XXX

Table 1. Samples Analyzed in This Study[a]

family	genera species (number of samples)	phase 1	phase 2
Engraulidae	*Engraulis encrasicolus* (11)	not unambiguously identified	*E. encrasicolus* 77.8−100
	Engraulis japonicus (0)	no samples available	
	Coilia spp (3)	*Coilia spp* 88.8−92.9	
Clupeidae	*Sardinops sagax* (1)	*Sardinops sagax* 96.5	
	Sprattus sprattus (2)	*Sprattus sprattus* 95.6−100	
	Sardinella aurita (4)	*Sardinella aurita* 100	
	Sardinella fimbriata (1)	*Sardinella fimbriata* 93.4	
	Sardina pilchardus (15)	Sardina pilchardus 66.8−93	
	Clupea harengus (12)	*Clupea harengus* 95.2−100	
	Sardinella jussieu (1)	*Sardinella jussieu/gibbosa/fimbriata* 93.4	
	Alosa spp (7)	*Alosa spp* 100	
Pleuronectidae	*Pleuronectes platessa* (20)	not unambiguously identified	*Pleuronectes platessa* 86.1−100
	Limanda limanda (2)	not unambiguously identified	*Limanda limanda* 93.1−93.6
	Atheresthes stomias (1)	*Atheresthes stomias* 100	
	Reinhardtius hippoglossoides (2)	*Reinhardtius hippoglossoides* 95.5−100	
	Paralichthys patagonicus (1)	*Paralichthys patagonicus* 100	
	Lepidopsetta polyxystra (5)	not unambiguously identified	*Lepidopsetta polyxystra* 93.6−100
	Platichthys flesus (3)	not unambiguously identified	*Platichthys flesus* 79.4−87.2
	Eopsetta jordani (1)	*Eopsetta jordani 100*	
Soleidae	*Solea solea* (6)	*Solea solea 96.8−100*	
	Solea senegalensis (1)	*Solea senegalensis* 100	
	Solea lascaris (3)	*Solea lascaris* 83.1−86.6	
	Synaptura lusitanica (8)	*Synaptura lusitanica* 94.2−100	
	Microchirus azevia (2)	*Microchirus azevia 100*	
Psettodidae	*Psettodes belcheri* (1)	*Psettodes belcheri* 78.6	
Bothidae	*Arnoglossus spp* (3)	*Arnoglossus laterna* 100	

[a]Column 2, species genera analyzed and number of samples tested. Column 3, identifications of phase 1. Column 4, identifications of phase 2. The identification score for the different species is also reported.

morphological and protein-based methods are not sufficient for correct species identification.[8]

The advancement of genomic technologies in recent years has made the DNA-based identification system a new and powerful option to accurately identify the fish species of high commercial value or of those frequently involved in commercial fraud. Generally speaking, this system is based on the presence of species-specific polymorphisms naturally occurring in the genomes.[16] The DNA-based methods present several advantages, such as high sensitivity, high specificity, large scale throughput, and the possibility to apply different types of analyses on the same specimen (e.g., PCR, sequencing, cloning, phylogenetic analysis, etc.). A variety of DNA-based identification approaches have been developed, as reported by refs 13 and 17, mainly based on PCR reaction: PCR-RFLP,[18−20] PCR-SSCP,[21] real time PCR,[22] PCR-RAPD,[23] and PCR-AFLP.[24] Besides the DNA-based approaches, DNA barcoding is the most widely adopted method for species identification.[25,26] Over the past decade, Sanger sequencing technology has become largely available, and at present it is considered one of the best methodologies to overcome the limitations of the morphology-based approach.[8,27−30] Microarray and next generation sequencing (NGS) technology has been also considered but not implemented to date for fish species identification.[11]

A novel and powerful DNA-based technology with a potential for species identification is pyrosequencing.[17,31] Using this technique, short stretches of nucleotides (approximately 30−40 nucleotides in length) downstream from a sequencing primer can be sequenced with high efficiency and accuracy.[32,33]

In this study, we evaluated the application of an alternative and modular technique to the Sanger sequencing (DNA barcoding) based on PCR followed by pyrosequencing for the rapid identification of two groups of fish commonly present in fish markets and frequently involved in commercial fraud. The first group includes species belonging to the genus *Clupeida*e as well as other species potentially used as fraudulent substitutes of *Clupeids*. The second group includes species belonging to the genus *Pleuronectidae* as well as other flat fish, such as *Solea solea*, *Hipoglossus hipoglossus*, etc.

It is interesting to note that products derived from these species are often commercialized as fillets in various forms, for example, fresh, frozen, marinated, salted, or breaded, all of which are unsuitable for morphological means of identification and sometimes even for protein-based species identification. Although pyrosequencing was previously applied for the identification of genetic lineages of *Salmo trutta* species,[34] to date, this is the first application of pyrosequencing-based analysis that aims at the identification of commercial fraud involving fish species.

MATERIALS AND METHODS

Sample Collection and DNA Extraction. One-hundred and sixteen (57 *Clupeiformes* and 59 *Pleuronectiformes*) specimens consisting of whole fishes and processed seafoods were collected from the fish market in Chioggia (Venice, Italy) and from local shops. Scientific and common names of fish species (11 *Clupeiformes* and 15 *Pleuronectiformes*) were assigned following identification by species-specific morphological traits following classification criteria and Sanger sequence DNA analysis; species names are listed in Table 1. DNA was extracted from 25 to 50 mg of fish muscles or processed seafood with the High Pure PCR Template Preparation kit (Roche Diagnostics,

dx.doi.org/10.1021/jf403545m | J. Agric. Food Chem. XXXX, XXX, XXX−XXX

Mannheim, Germany) according to the manufacturer's instructions. Specimens represented the most common product types present on the market in the commerce, such as fresh (skinned, filleted, eviscerated, etc.) or processed products (dried, salted, marinated, smoked, etc.). In detail, the samples analyzed were frozen fillet (59 samples), smoked (7 samples), cooked/precooked (12 samples), whole fresh fish (5 samples), marinated (6 samples), fermented (2 samples), salted in oil (14 samples), and frozen in breadcrumbs fillets (11 samples).

Preliminary Identification. *Morphological Analysis.* The specimens represented by the whole fish (n = 5) included in this study were at first identified by morphological analysis. Whole fish, with intact skin and fins, were identified using the dichotomous keys approach proposed by the FAO and available on the Web site http://www.fao.org/fishery/species/search/en, which refers to specific monographs (ref 35 for the Mediterranean Sea).

Sanger Sequence-Based Identification. All the specimens were identified by using cytochrome c oxidase I (COI) sequence. Briefly, the samples were amplified by using the PCR reaction previously described.[36] Sanger sequencing method was applied on the obtained PCR products. The sequences were generated using the Big Dye Terminator v3.1 cycle sequencing kit (Applied Biosystem, Foster City, CA). The products of the sequencing reactions were cleaned-up using PERFORMA DTR Ultra 96-Well kit (Edge BioSystems, Gaithersburg, MD) and sequenced by using ABI PRISM 3130xl Genetic Analyzer (Applied Biosystem, Foster City, CA). COI sequences obtained were compared to the COI sequences deposited in the GenBank database.

PCR Design and PCR Amplification for Pyrosequencing. Sequences of conserved regions of mitochondrial DNA (mtDNA) from different species present in the public database GenBank were aligned using the software MEGA 5.[37] On the basis of the analysis of the alignment, a set of PCR primers targeting 16S rRNA conserved region (16SForbio 5′-biotin-ACGAGAAGACCCTDTGGAG-3′, 16SRev 5′-TGTTATCCCTAGGGTAACTTG-3′) and a sequencing primer (16Sseq 5′-GTCGCCCCAACCGAAGA-3′) for pyrosequencing were designed. This set of primer was developed to amplify all the commercial species of interest in this study. On the basis of the *in silico* analysis, not all the originated pyrosequences showed the polymorphisms needed to be correctly and uniquely differentiated as in the case of *Cluipeidae*, that is, *Engraulis encrasicolus* and *Engraulis japonicus.* As for *Pleuronectidae*, *Hippoglossoides elassodon*, *Reinhardtius hippoglossoides*, *Limanda limanda*, *Limanda aspera*, *Pleuronectes quadrituberculatus*, *Pleuronectes platessa*, *Platichtys flesus*, and *Lepidopsetta bilineata* were not uniquely identified. Then, two other set of primers targeting different genomic sequences were designed. The first set of primers (Engra JEF 5′-biot- GCAGCCTTCCTTACCTTAAACA-3′ and Engra JER 5′-GTAGGAGGTTTGTGGCGAGAG-3′ for amplification, Engra JES 5′-AGTCACTTGGGTAAGAATC-3′ for sequencing), targeting the NADH dehydrogenase subunit II gene (ND2), allows discrimination of *Engraulis japonicus* from *Engraulis encrasicolus* on the basis of the pyrosequences obtained. The second set (PleuFbio 5′-ATCGCAAACGATGCTTTAG-3′ and PleuRseq1 5′-GGAARAGAAAGTGGAAKGC-3′ for amplification, PleuRseq2 5′-GGAAGAGRAAGTGGAATGC-3′ for sequencing), targeting a portion of the *cytb* gene of the *Pleuronectiformes*, discriminates species of *Hippoglossoides elassodon*, *Reinhardtius hippoglossoides*, *Limanda limanda*, *Limanda aspera*, *Pleuronectes quadrituberculatus*, *Pleuronectes platessa*, *Platichtys flesus*, and *Lepidopsetta bilineata*.

For each set, one of the amplification primer had to be biotinylated according to the pyrosequencing chemistry to allow the purification and selection of the specific fragment in the subsequent procedures.

Amplification of DNA was performed by using a PCR in a 50 μL reaction volume with 5 μL of DNA, 1.5 U AmpliTaq Gold (Roche Diagnostics, Mannheim, Germany), 2 mM $MgCl_2$, 1 mMdNTPs, and 0.5 μM of each specific primer (Eurofins MWG-Operon, Ebersberg, Germany). PCR was performed with the following thermal conditions: 10 min at 95 °C, 40 cycles consisting in 30 s at 95 °C, 30 s at 58 °C, 55 °C, 52 °C respectively for ND2, 16S and cyt B amplification, 30 s at 72 °C, with a final extension of 5 min at 72 °C. The amplified products, 289 bp for 16S, 520bp for cyt B and 291bp for ND2, were analyzed on a 2% agarose gel (Sigma-Aldrich, St. Louis, MO) in TAE buffer with 0.1 μL/mL GelRed (Biotium, Hayward, CA) and visualized under UV illumination.

Pyrosequencing Reaction and Sequence Library. The biotinylated PCR products were purified and prepared for the pyrosequencing reaction according to the manufacturer's instructions (Biotage, Uppsala, Sweden). Briefly, 20 μL of the biotinylated PCR product was immobilized onto 4 μLof Streptavidine-Sepharose (GE Healthcare, Uppsala, Sweden) beads in 40 μL of binding buffer with 20 μL of DEPC water at room temperature for 30 min. Single-stranded DNA was prepared with the PyroMark Vacuum Prep Workstation (Biotage, Uppsala, Sweden). The single-stranded, biotinylated DNA products were hybridized to sequencing primer (final concentrations of 0.5 μM) in a 96-well plate . The reaction was performed in 40 μL of annealing buffer in a sample plate heated at 80 °C for 4 min and then cooled to room temperature for 10 min before being placed into the pyrosequencing instrument. Pyrosequencing was performed using PyroGold reagents according to manufacturer's recommendations. The sequences originated by the analysis were automatically compared with a sequences library containing all the reference sequences of the species of interest in this study. The sequences library was generated, and the analysis was performed using the IdentiFire software (Biotage, Uppsala, Sweden).

For some of the fish species included in this project, sequences of the target genes were not available in the public databases. Those sequences were generated using the classical sequencing protocol (Sanger), added to the library (IdentiFire software), and deposited in a public database (see following section).

This software will assign a score value to each tested sequence, indicating the quality and the affordability of the result. This score is not only based on the sequence homology, but also on the position of any possible mismatch, the quality of the pyrogram, and the intensity of the peaks. Thus, the same number of mismatches may give different scores according to the positions of the polymorphisms along the pyrosequence. Mismatches closer to the 5′-end of the pyrosequences can affect the score value, which can be decreased more than a mismatch close to the 3′-end.

RESULTS

Preliminary Identification. The samples analyzed in this study originated from species belonging to *Clupeiformes* (n = 57) and *Pleuronectiformes* (n = 59) (Table 1).

Sequences obtained by the Sanger method were used as the gold standard for pyrosequences identification. For some of the species analyzed in the present study, the sequence of 16S gene was not available in the public databases, and therefore the newly Sanger-generated 16S sequences were submitted to Genbank [accession numbers: KC461225 (*Alosa falla nilotica*), KC461223, KC461224 (*Alosa fallax lacustris*), and KC461222 (*Sardinella fimbriata*)]. Similarly, *cytb* PCR products from samples of *Lepidopsetta polyxystra* (identified by COI gene Sanger sequence DNA analysis) were submitted to GenBank (Acc numbers: KF007183, KF007184, KF007185).

Pyrosequencing Identification: Experimental Design and Proposed Workflow. In order to correctly and unambiguously identify all the 25 fish species targeted in this study, the entire procedure was divided into two phases (herein, phase 1 and phase 2). Phase 1 was based on the use of one PCR reaction targeting the 16S rRNA gene, followed by the pyrosequencing analysis to identify and differentiate 20 out of the 25 targeted species. For those sequences not uniquely identified in phase 1, two other PCR amplifications followed by pyrosequencing were performed (phase 2), specifically targeting the group of *Clupeiformes* (two species) and the *Pleuronectiformes* (four species). The proposed workflow is schematically illustrated in Figure 1.

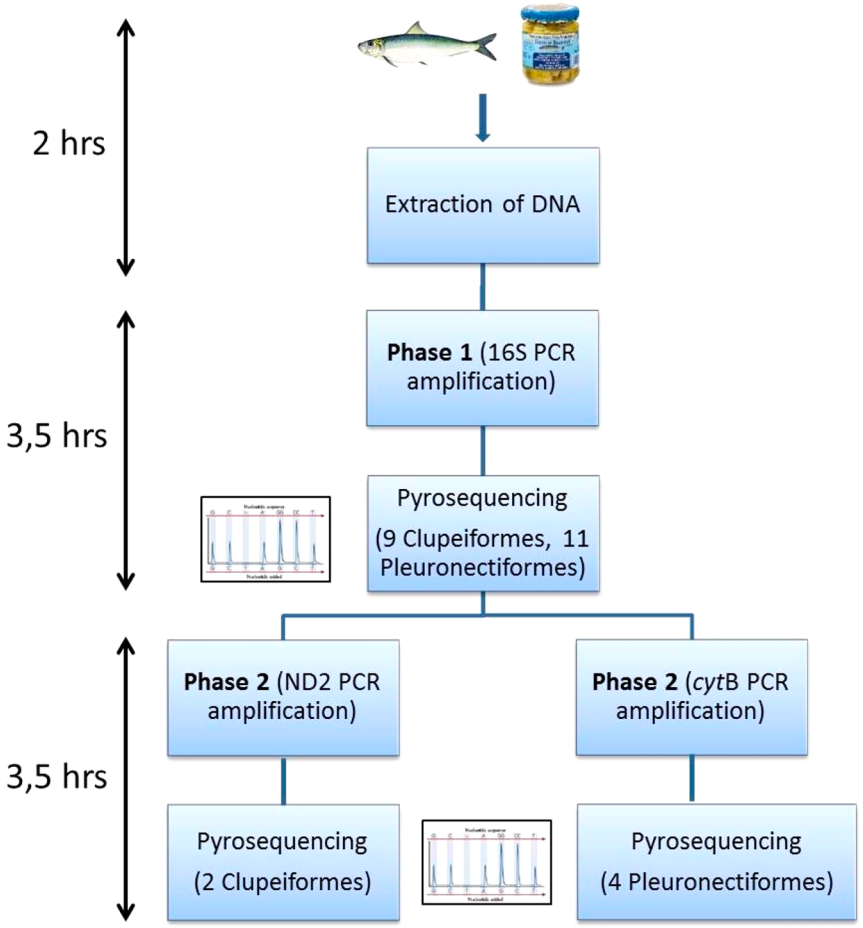

Figure 1. Schematic representation of pyrosequencing workflow. The number of fish species identified throughout the two phases is indicated as well as the analysis turn-around time. Overall, 20 fish species can be identified with this procedure in 3.5 h. The identification of the 25 species targeted in this study is achieved in a maximum time frame of 7 h. It can be reduced if phase 1 and 2 are performed simultaneously.

Phase 1 – Targeting the 16S rRNA Gene. The DNA isolated from all the fresh and processed samples collected in this study was amplified using the PCR targeting the 16S rRNA gene.

The pyrosequencing reaction was able to produce the short sequences, and the subsequent comparison of the query sequences was carried out with the reference sequences in the library database (based on the Identifire software). The results demonstrated that for 20/25 species the 16S rRNA sequences identified correctly and uniquely with a score range in the Identify software between 66.8 and 100 (*Coilia spp, Sardinops sagax, Sprattus sprattus, Sardinella aurita, Sardinella fimbriata, Sardina pilchardus, Clupea harengus, Sardinella jusseu, Alosa spp., Atheresthes stomias, Reinhardtius hippoglossoides, Paralitchthys patagonicus, Eopsetta jordani, Solea solea, Solea senegalensis, Solea lascaris, Synaptura lusitanica, Microchirus azevia, Psettodes belcheri, Arnoglossus spp.*) (Table 1). Some samples were ambiguously identified because of the similarity of their sequences with more than one reference sequences included in the database, as was the case for two samples identified in the *Clupeiformes* group as *Engraulis encrasicolus* and *Engraulis japonicus.* As illustrated in Figure S-1 (Supporting Information), the same score with three different reference sequences, *Engraulis australis*, *Engraulis encrasicolus*, and *Engraulis japonicus* was obtained. Similarly, a unique and precise identification of *Limanda limanda*, *Pleuronectes platessa*, *Platichthys flesus*, and *Lepidopsetta polyxystra* in the *Pleuronctiformes* group was not possible in phase 1 (Figure S-2, Supporting Information).

Phase 2 – Targeting NADH and cyt b Gene. The samples assigned to *Clupeiformes* group but ambiguously identified in phase 1 were tested with a PCR assay targeting the NADH dehydrogenase subunit II gene (ND2). The subsequent pyrosequencing procedure enabled the discrimination between genetic sequences belonging to *Engraulis encrasicolus* and *Engraulis japonicus* and the correct identification of these species with scores ranging from 77.8 to 100 (Figure 2).

The samples assigned to the *Pleuronectiformes* group but ambiguously identified in phase 1 were tested with a distinct PCR targeting the cytochrome *b* (*cyt* b) gene. The sequences obtained allowed the correct and unambiguous identification of four of the previously misclassified species (*Limanda limanda, Pleuronectes platessa, Platichtys flesus, Lepidopsetta polyxystra*), with scores ranging from 79.4 (in the case of *Platichtys flesus*) to 100 (in the case of *Pleuronectes platessa*) (Figure S-3, Supporting Information).

■ DISCUSSION

The identification of a species substitution using morphological detection or other classical systems (IEF) can fail when processed fish products are considered, since the diagnostic

dx.doi.org/10.1021/jf403545m | J. Agric. Food Chem. XXXX, XXX, XXX−XXX

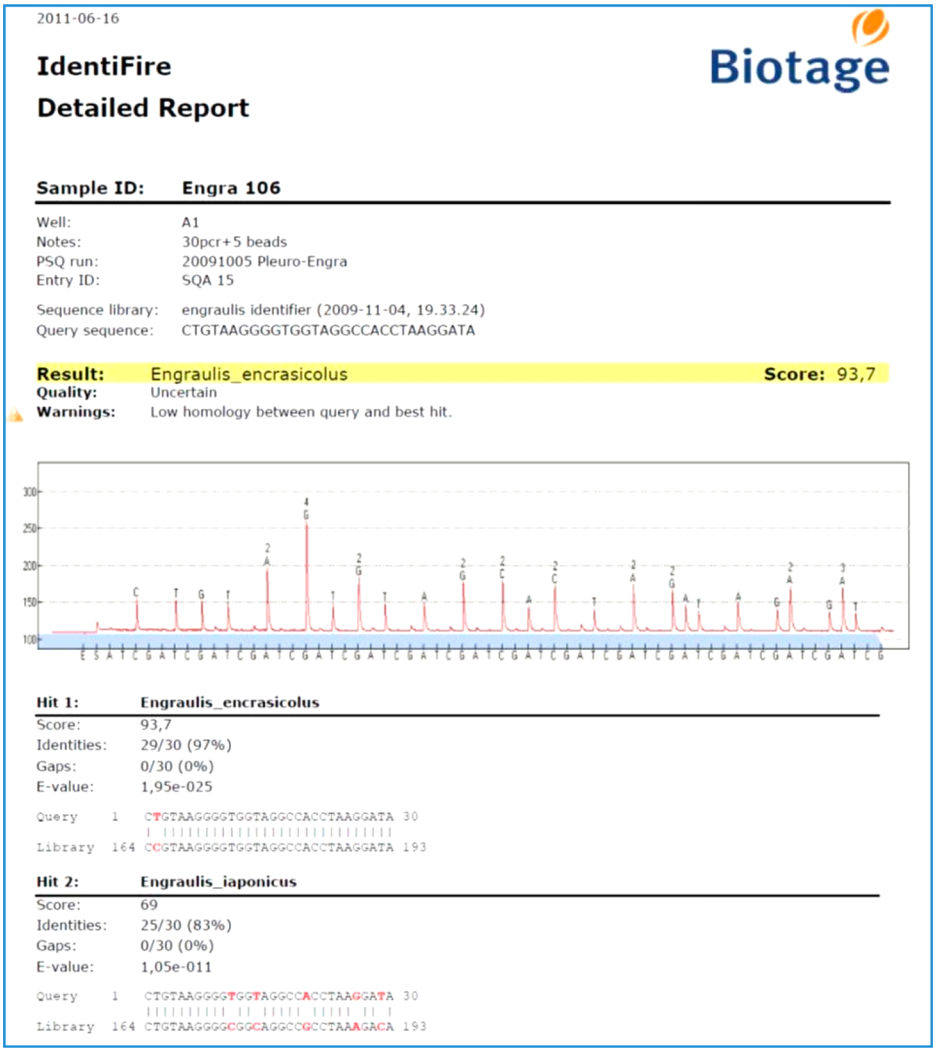

Figure 2. Example of IdentiFire results of the phase 2 discriminating *Engraulis encrasicolus* from *Engraulis japonicus*. The different score assigned indicates a level of polymorphism that enables correct identification.

target can be damaged. In those cases, traditional methods do not always provide satisfactory results.[8]

The selection of the proper DNA target is a crucial point in the DNA-based fish species identification methods. The DNA target should be present in large amounts to ensure it will always be detectable and show a sufficient number of polymorphisms between the species to be unequivocally assigned to a given species.[13] Several genes or genomic sequences satisfy these criteria, but the mtDNA has been widely used because of its particular features: it is present in several copies into the cell (up to 1000 times more than in the nuclear DNA), and it shows more polymorphisms than nuclear DNA because of its faster evolution.[13] Besides this, mtDNA is more stable because of its circular structure that may constitute a crucial advantage in the case of processed seafood, where thermal treatment could reduce the integrity of DNA originating fragments from less than 100 bp up to about 500 bp.[2] Among the mtDNA, some genes have been widely used for fish species identification, especially cytochrome oxidase c subunit 1 (COI), cytochrome *b* (*cytb*), and 16S rRNA.[1,2,11,13,38−44] The 650 base pairs sequence of COI is the main barcode for animal species identification and can discriminate all the commercially available fish species.[45,46] The pyrosequencing technique provides the sequence of a conserved region containing a short stretch of nucleotides (30−40) with a number of polymorphisms sufficient enough to enable differentiation of the species under investigation.

The testing protocol developed in this study is based on three different PCR reactions followed by pyrosequencing analysis and provides a rapid, sensitive, specific, and robust system to identify fish species based on the nucleotide sequence of a specific locus.[34]

The simple workflow schematically illustrated in Figure 1 allowed for an efficient and accurate identification of 25 species (11 *Clupeiformes* and 15 *Pleuronectiformes*), and the analysis detected some of the most important commercial frauds related to these groups of fish that could have a great impact on the market.[47]

The precise and unambiguous discrimination of *Engraulis encrasicolus* (commercially, a high-value species) is relevant in the framework of the official control of commercial fish frauds.[48] *E. encrasicolus* can also be substituted by *Sprattus sprattus*,[49] and this fraud can be clearly identified by the pyrosequencing protocol as well. The lack of available samples belonging to *E.* japonicus in this study did not allow for testing of this species in the laboratory; nevertheless, the *in silico* analysis confirmed the clear DNA discrimination between *E.*

dx.doi.org/10.1021/jf403545m | J. Agric. Food Chem. XXXX, XXX, XXX−XXX

japonicus and *E. encrasicolus* on the basis of the *cyt* b available sequences from public databases.

In *Pleuronectiformes*, *Pleuronectes platessa* and *Solea solea* are the two species that are most widely substituted by various low-value species such as members of the genus *Limanda* and *Lepidopsetta*. The results presented showed that *P. platessa* and *S. solea* can be unambiguously differentiated. It is clear that the identification of species by using sequence data is strongly affected by the availability of reference sequences in public databases.

A major point of interest is that the method developed in this study has proven to be highly efficient on processed fish products as well. As a matter of fact, the DNA from all these samples was amplified and correctly assigned to a given species, indicating that the method developed is not affected by the various processing methods.

As for other barcoding techniques, such as mini-barcoding,[50] one of the advantages of the pyrosequencing protocol is that the relatively small size of the amplified PCR products, less than 520 bp, coupled with the high efficiency and accuracy of this sequencing method for short sequence fragments allows detection of high quality genetic sequences, even in samples where DNA may be partially damaged or degraded by the processing, providing a robust identification system.[2]

It was possible to identify the specimens at the species level in most of the analyzed samples; however, in some cases the identification limited to the Genus level (e.g., *Arnoglossus sp*, *Coilia sp*, and *Alosa sp*) was still considered sufficient for commercial purposes. In those cases, all the members of the genus presented comparable commercial values and/or the same commercial denomination. The analysis in phase 1 can be considered as a primary screening test to identify most of the species of interest being capable to identify 20 species in 3.5 h (Figure 1). Phase 2 analysis acts as a fine-tuned identification step, but both can be performed simultaneously to reduce the turn-around time, if necessary (Figure 1). To make the entire process as fast as possible (i.e., about 3.5 h), all three PCR reactions (phase 1 and 2) can be performed at the same time, and the resulting PCR products can then be analyzed in a single run on the pyrosequencing platform. Furthermore, depending on the species to be investigated (and the correspondent fraud), only one specific PCR could be performed. Indeed, in most cases of mislabeling fraud, only one of the PCR presented was sufficient to identify the species involved. According to this, the system presented can be considered as a modular system, and either phase can be used independently. This reduces the total analysis time to that of a single PCR step.

The classical sequencing system (Sanger method) is a very useful tool to identify fish species, but it still is rather expensive and time-consuming; therefore, its use is not recommended for routine analyses to date.[51] On the basis of the commercial value of the reagents and platform, the cost per sample of PyroMark ID pyrosequencing analysis ranges from 5 to 6 euro (from 480 to 576 euro/96 plate). On the other hand, the cost per sample of the Sanger sequencing system is of about 11 or 12 euro (1056 to 1152 euro/96 plate). Interestingly, the pyrosequencing pipeline is more time-efficient and cheaper than Sanger sequencing. Indeed, pyrosequencing requires only few protocol steps, and the sequence detection is performed in real-time. The additional mandatory steps to be performed in the Sanger protocol, consisting of big-dye terminator sequencing reaction, postreaction purification, and final electrophoresis, prolong the total time of the analysis and increase costs.[50] Conversely, the total cost of the pyrosequencing analysis has decreased since it does not need terminator enzymes and time-consuming purification steps.[52−54]

The newest upgrade to PyroMark platform makes the read-length up to 140 nucleotides (instead of the current 30−40) longer (www.qiagen.com). This improvement could extend the application field of the pyrosequencing technique to a wider target range. In addition, the cost of the new upgraded platform is still lower than the Sanger sequencing platform commercially available making pyrosequencing even cheaper and more powerful.

In conclusion, the analysis of the 30-bp fragment following the sequencing primer enabled the discrimination of 25 fish species in 3.5−7 h, indicating that the fish species identification system based on pyrosequencing analysis is a powerful tool to discriminate the most important cases of commercial fraud regarding *Clupeiformes* and *Pleuronectiformes* groups.

■ ASSOCIATED CONTENT

Supporting Information

IdentiFire analysis figures. Figure S-1: IdentiFire results of phase 1 on species belonging of the Clupeiformes group. The maximum score was assigned to more than one reference sequence, indicating that for some samples the phase 1 was not sufficient for the correct and unambiguous identification. Figure S-2: Example of IdentiFire results of phase 1 on species belonging to the Pleuronectiformes group. The maximum score was assigned to more than one reference sequence, indicating that for some samples the phase 1 was not sufficient for a correct and unambiguous identification. Figure S-3: IdentiFire results of the phase 2 discriminating *Pleuronectes platessa* from its major substitutions. The different score assigned indicates a level of polymorphism which enables correct identification. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Tel. +390498084381; fax +390498084360; e-mail cdebattisti@izsvenezie.it.

Present Address

⊥(C.M.) Sacco Srl, Via Manzoni 29/A, 22071 Cadorago (Co), Italy.

Funding

This work was supported by the RC VE 3-09 (Italian Ministry of Health) and partially by the Italian Ministry of Agriculture (INRAN, no. E12900E1).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We are grateful to Francesca Ellero and Samantha Kasloff (Research & Innovation Department, IZSVe, Italy) for revising the manuscript.

■ REFERENCES

(1) Comi, G.; Iacumin, L.; Rantsiou, K.; Cantoni, C.; Cocolin, L. Molecular methods for the differentiation of species used in production of cod-fish can detect commercial frauds. *Food Control* **2005**, *16*, 37−42.

(2) Rasmussen, R. S.; Morrissey, M. T. DNA-Based Methods for the Identification of Commercial Fish and Seafood Species. *Compr. Rev. Food Sci. Food Saf.* **2008**, *7*, 280−295.
(3) Hanner, R.; Becker, S.; Ivanova, N. V.; Steinke, D. FISH-BOL and seafood identification: Geographically dispersed case studies reveal systemic market substitution across Canada. *Mitochondrial DNA* **2011**, *22*, 106−122.
(4) Marko, P. B.; Lee, S. C.; Rice, A. M.; Gramling, J. M.; Fitzhenry, T. M.; McAlister, J. S.; Harper, G. R.; Moran, A. L. Fisheries: mislabelling of a depleted reef fish. *Nature* **2004**, *430*, 309−310.
(5) Pepe, T.; Trotta, M.; Di Marco, I.; Anastasio, A.; Bautista, J. M.; Cortesi, M. L. Fish species identification in surimi-based products. *J. Agric. Food Chem.* **2007**, *55*, 3681−3685.
(6) Garcia-Vazquez, E.; Perez, J.; Martinez, J. L.; Pardinas, A. F.; Lopez, B.; Karaiskou, N.; Casa, M. F.; Machado-Schiaffino, G.; Triantafyllidis, A. High level of mislabeling in Spanish and Greek hake markets suggests the fraudulent introduction of African species. *J. Agric. Food Chem.* **2011**, *59*, 475−480.
(7) Di Pinto, A.; Di Pinto, P.; Terio, V.; Bozzo, G.; Bonerba, E.; Ceci, E.; Tantillo, G. DNA barcoding for detecting market substitution in salted cod fillets and battered cod chunks. *Food Chem.* **2013**, *141*, 1757−1762.
(8) Yancy, H. F.; Zemlak, T. S.; Mason, J. A.; Washington, J. D.; Tenge, B. J.; Nguyen, N. L.; Barnett, J. D.; Savary, W. E.; Hill, W. E.; Moore, M. M.; Fry, F. S.; Randolph, S. C.; Rogers, P. L.; Hebert, P. D. Potential use of DNA barcodes in regulatory science: applications of the Regulatory Fish Encyclopedia. *J. Food Prot.* **2008**, *71*, 210−217.
(9) Civera, T. Species identification and safety of fish products. *Vet. Res. Commun.* **2003**, *27* (Suppl 1), 481−489.
(10) Sanchez, A.; Quinteiro, J.; Rey-Mendez, M.; Perez-Martin, R. I.; Sotelo, C. G. Identification of European Hake species (*Merluccius merluccius*) using real-time PCR. *J. Agric. Food Chem.* **2009**, *57*, 3397−3403.
(11) Kochzius, M.; Seidel, C.; Antoniou, A.; Botla, S. K.; Campo, D.; Cariani, A.; Vazquez, E. G.; Hauschild, J.; Hervet, C.; Hjorleifsdottir, S.; Hreggvidsson, G.; Kappel, K.; Landi, M.; Magoulas, A.; Marteinsson, V.; Nolte, M.; Planes, S.; Tinti, F.; Turan, C.; Venugopal, M. N.; Weber, H.; Blohm, D. Identifying Fishes through DNA Barcodes and Microarrays. *PLoS One* **2010**, *5*, e12620.
(12) Rehbein, H. Electrophoretic techniques for species identification of fishery products. *Z. Lebensmittel-Untersuchung Forschung A* **1990**, *191*, 1−10.
(13) Teletchea, F. Molecular identification methods of fish species: reassessment and possible applications. *Rev. Fish Biol. Fisheries* **2009**, *19*, 265−293.
(14) Calo-Mata, P.; Sotelo, C. G.; Pérez-Martín, R. I.; Rehbein, H.; Hold, G. L.; Russell, V. J.; Pryde, S.; Quinteiro, J.; Rey-Méndez, M.; Rosa, C.; Santos, A. T. Identification of gadoid fish species using DNA-based techniques. *Eur. Food Res. Technol.* **2003**, *217*, 259−264.
(15) Horstkotte, B.; Rehbein, H. Fish Species Identification by Means of Restriction Fragment Length Polymorphism and High-Performance Liquid Chromatography. *J. Food Sci.* **2003**, *68*, 2658−2666.
(16) Liu, Z. J.; Cordes, J. F. DNA marker technologies and their applications in aquaculture genetics. *Aquaculture* **2004**, *238*, 1−37.
(17) Hellberg, R. S.; Morrissey, M. T. Advances in DNA-based techniques for the detection of seafood species substitution on the commercial market. *J. Lab. Autom.* **2011**, *16*, 308−321.
(18) Perez, J.; Garcia-Vazquez, E. Genetic identification of nine hake species for detection of commercial fraud. *J. Food Prot.* **2004**, *67*, 2792−2796.
(19) Perez, M.; Presa, P. Validation of a tRNA-Glu-cytochrome b key for the molecular identification of 12 hake species (Merluccius spp.) and Atlantic Cod (*Gadus morhua*) using PCR-RFLPs, FINS, and BLAST. *J. Agric. Food Chem.* **2008**, *56*, 10865−10871.
(20) Aranishi, F.; Okimoto, T.; Izumi, S. Identification of gadoid species (Pisces, Gadidae) by PCR-RFLP analysis. *J. Appl. Genet.* **2005**, *46*, 69−73.
(21) Weder, J.; Rehbein, H.; Kaiser, K. On the specificity of tuna-directed primers in PCR-SSCP analysis of fish and meat. *Eur. Food Res. Technol.* **2001**, *213*, 139−144.
(22) Casper, R. M.; Jarman, S. N.; Deagle, B. E.; Gales, N. J.; Hindell, M. A. Detecting prey from DNA in predator scats: A comparison with morphological analysis, using Arctocephalus seals fed a known diet. *J. Exp. Mar. Biol. Ecol.* **2007**, *347*, 144−154.
(23) Lakra, W. S.; Goswami, M.; Gopalakrishnan, A. Molecular identification and phylogenetic relationships of seven Indian Sciaenids (Pisces: Perciformes, Sciaenidae) based on 16S rRNA and cytochrome c oxidase subunit I mitochondrial genes. *Mol. Biol. Rep.* **2009**, *36*, 831−839.
(24) Maldini, M.; Nonnis Marzano, F.; Fortes, G. G.; Papa, R.; Gandolfi, G. Fish and seafood traceability based on AFLP markers: Elaboration of a species database. *Aquaculture* **2006**, *261*, 487−494.
(25) Wong, E.H.-.; Hanner, R. H. DNA barcoding detects market substitution in North American seafood. *Food Res. Int.* **2008**, *41*, 828−837.
(26) Keskin, E.; Atar, H. H. Genetic structuring of European anchovy (*Engraulis encrasicolus*) populations through mitochondrial DNA sequences. *Mitochondrial DNA* **2012**, *23*, 62−69.
(27) Allander, T.; Emerson, S. U.; Engle, R. E.; Purcell, R. H.; Bukh, J. A virus discovery method incorporating DNase treatment and its application to the identification of two bovine parvovirus species. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 11609−11614.
(28) Pace, N. R. A molecular view of microbial diversity and the biosphere. *Science* **1997**, *276*, 734−740.
(29) Hamels, S.; Gala, J. L.; Dufour, S.; Vannuffel, P.; Zammatteo, N.; Remacle, J. Consensus PCR and microarray for diagnosis of the genus Staphylococcus, species, and methicillin resistance. *BioTechniques* **2001**, *31*, 1364−6, 1368, 1370−2.
(30) Nanney, D. L. Genes and phenes in Tetrahymena. *Bioscience* **1982**, *32*, 783−788.
(31) Balitzki-Korte, B.; Anslinger, K.; Bartsch, C.; Rolf, B. Species identification by means of pyrosequencing the mitochondrial 12S rRNA gene. *Int. J. Legal Med.* **2005**, *119*, 291−294.
(32) Ronaghi, M. Pyrosequencing sheds light on DNA sequencing. *Genome Res.* **2001**, *11*, 3−11.
(33) Woolfe, M.; Primrose, S. Food forensics: using DNA technology to combat misdescription and fraud. *Trends Biotechnol.* **2004**, *22*, 222−226.
(34) Keller, I.; Taverna, A.; Seehausen, O. A pyrosequencing assay for the rapid discrimination of mitochondrial lineages in the Salmo trutta species complex. *Mol. Ecol. Resour.* **2011**, *11*, 196−199.
(35) Fischer W.; Bauchot, M. L.; Schneider, M. *Fiches FAO d'identification des espèces pour les besoins de la pêche. Méditerranée et mer Noire. Zone de Pêche 37 (rev. 1)*; Commission des Communautés Européennes and FAO: Rome, 1987; Vol. *II*, pp 891−1421.
(36) Baldwin, C. C.; Weight, L. A.; Smith, D. G.; Mounts, J. H. Reconciling Genetic Lineages with Species in Western Atlantic Coryphopterus (Teleostei: Gobiidae). *Smithsonian contributions to the marine science* **2009**, 111−138.
(37) Tamura, K.; Dudley, J.; Nei, M.; Kumar, S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* **2007**, *24*, 1596−1599.
(38) Kartavtsev, Y. P.; Sharina, S. N.; Goto, T.; Chichvarkhin, A. Y.; Balanov, A. A.; Vinnikov, K. A.; Ivankov, V. N.; Hanzawa, N. Cytochrome oxidase 1 gene sequence analysis in six flatfish species (Teleostei, Pleuronectidae) of Far East Russia with inferences in phylogeny and taxonomy. *Mitochondrial DNA* **2008**, *19*, 479−489.
(39) Aquino, L. M.; Tango, J. M.; Canoy, R. J.; Fontanilla, I. K.; Basiao, Z. U.; Ong, P. S.; Quilang, J. P. DNA barcoding of fishes of Laguna de Bay, Philippines. *Mitochondrial DNA* **2011**, *22*, 143−153.
(40) Ribeiro, A. O.; Caires, R. A.; Mariguela, T. C.; Pereira, L. H.; Hanner, R.; Oliveira, C. DNA barcodes identify marine fishes of Sao Paulo State, Brazil. *Mol. Ecol. Resour.* **2012**, *12*, 1012−1020.
(41) Kochzius, M. Trends in Fishery Genetics. In *The Future of Fisheries Science in North America, 1*; Beamish, R. J., Rothschild, B. J.,

Eds.; Springer Netherlands: AK Houten, Netherlands, 2009; Vol. *31*, pp 453−493.
(42) Espineira, M.; Gonzalez-Lavin, N.; Vieites, J. M.; Santaclara, F. J. Authentication of anglerfish species (Lophius spp) by means of polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) and forensically informative nucleotide sequencing (FINS) methodologies. *J. Agric. Food Chem.* **2008**, *56*, 10594−10599.
(43) Espineira, M.; Gonzalez-Lavin, N.; Vieites, J. M.; Santaclara, F. J. Development of a method for the genetic identification of flatfish species on the basis of mitochondrial DNA sequences. *J. Agric. Food Chem.* **2008**, *56*, 8954−8961.
(44) Rasmussen, R. S.; Morrissey, M. T.; Hebert, P. D. DNA barcoding of commercially important salmon and trout species (Oncorhynchus and Salmo) from North America. *J. Agric. Food Chem.* **2009**, *57*, 8379−8385.
(45) Hebert, P. D.; Ratnasingham, S.; deWaard, J. R. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proc. Biol. Sci.* **2003**, *270* (Suppl 1), S96−9.
(46) Hebert, P. D.; Cywinska, A.; Ball, S. L.; deWaard, J. R. Biological identifications through DNA barcodes. *Proc. Biol. Sci.* **2003**, *270*, 313−321.
(47) Lahsen, A. *World Inventory of Fisheries. Trade Implications of Fish Species Identification. Issues Fact Sheets*; Food and Agriculture Organization of the United Nations, Fisheries and Aquaculture Department: Rome, Italy, 2005; http://www.fao.org/fishery/topic/14807/en.
(48) Santaclara, F.; Cabado, A.; Vieites, J. Development of a method for genetic identification of four species of anchovies: *E. encrasicolus*, *E. anchoita*, *E. ringensand*, *E. japonicus*. *Eur. Food Res. Technol.* **2006**, *223*, 609−614.
(49) Sebastio, P.; Zanelli, P.; Neri, T. M. Identification of anchovy (*Engraulis encrasicholus* L.) and gilt sardine (*Sardinella aurita*) by polymerase chain reaction, sequence of their mitochondrial cytochrome b gene, and restriction analysis of polymerase chain reaction products in semipreserves. *J. Agric. Food Chem.* **2001**, *49*, 1194−1199.
(50) Shokralla, S.; Zhou, X.; Janzen, D. H.; Hallwachs, W.; Landry, J. F.; Jacobus, L. M.; Hajibabaei, M. Pyrosequencing for mini-barcoding of fresh and old museum specimens. *PLoS One* **2011**, *6*, e21252.
(51) Durand, J. D.; Diatta, M. A.; Diop, K.; Trape, S. Multiplex 16S rRNA haplotype-specific PCR, a rapid and convenient method for fish species identification: an application to West African Clupeiform larvae. *Mol. Ecol. Resour.* **2010**, *10*, 568−572.
(52) Pizzuto, M. S.; De Battisti, C.; Marciano, S.; Capua, I.; Cattoli, G. Pyrosequencing analysis for a rapid classification of fowl adenovirus species. *Avian Pathol.* **2010**, *39*, 391−398.
(53) Dudley, D. M.; Chin, E. N.; Bimber, B. N.; Sanabani, S. S.; Tarosso, L. F.; Costa, P. R.; Sauer, M. M.; Kallas, E. G.; O'Connor, D. H. Low-Cost Ultra-Wide Genotyping Using Roche/454 Pyrosequencing for Surveillance of HIV Drug Resistance. *PLoS One* **2012**, *7*, e36494.
(54) Timbs, A. T.; Rugless, M. J.; Gallienne, A. E.; Haywood, A. M.; Henderson, S. J.; Old, J. M. Prenatal diagnosis of hemoglobinopathies by pyrosequencing: a more sensitive and rapid approach to fetal genotyping. *Hemoglobin* **2012**, *36*, 144−150.