

DNA BARCODING: UNO STRUMENTO UTILE ALL'ISPEZIONE DEGLI ALIMENTI

DNA BARCODING: A USEFUL TOOL FOR FOOD INSPECTION

Bernardi C., Colombo F., Balzaretto C., Gagliardi C., Cattaneo P.
Dipartimento di Scienze e Tecnologie Veterinarie per la Sicurezza Alimentare - Università degli Studi di Milano.

SUMMARY

The central concept in species identification by molecular methods is to match the DNA sequence of unknown sample to a reference sequence through DNA sequence similarity searches. The limitation of this process is the lack of authenticated reference sequence databases. For instance, in the GenBank, the biggest database available, the sequences are accepted without any standard protocols or quality control, raising doubts over their suitability for food inspection application. Barcoding may be a tool for species identification because its aim is the production of COI (Cytochrome Oxidase subunit I) reference sequences using standard protocols. Every specimen record will not gain formal barcode status until seven data elements are placed in: species name, voucher data, collector record, identifier of specimen, COI sequence of at least 500 bp, polymerase chain reaction primers used to generate the amplicon and trace files. BOLD (Barcode of Life Data System) employs several tools to identify data anomalies or low quality records. Recently this system has been adopted in Italy too to verify the compliance of prepared fishery products with species declared in label. The authors introduce a discussion about a possible simplification of the Italian official list of seafood, especially for species of the same commercial value.

KEYWORDS

DNA barcoding, species identification, mislabeled seafood

L'identificazione di specie è da sempre un problema con il quale si è confrontato il veterinario ispettore, che dal macello all'esercizio di vendita al dettaglio ha utilizzato le proprie conoscenze per identificare la specie animale. Quando in un alimento gli elementi morfologici disponibili non sono sufficienti ad indirizzare verso una corretta identificazione, si deve ricorrere a tecniche analitiche. Nel tempo sono state sviluppate numerose tecniche, le prime miravano ad evidenziare le differenze intraspecifiche su base proteica (elettroforesi, isoelectrofocusing IEF, elettroforesi capillare CE, high-performance liquid chromatography HPLC) (1, 2, 3). Il limite di questi metodi è legato al ristretto campo di applicazione agli alimenti freschi, congelati o trasformati con metodi blandi che non alterano l'integrità strutturale delle proteine. Al fine di allargare il campo applicativo furono messi a punto altri metodi utili per l'analisi di alimenti trattati con il calore ma non sterilizzati: elettro-

foresi su gel di sodio dodecil solfato policrilamide (SDS-PAGE) e IEF con urea (4). Per le conserve di carne è applicabile il metodo ELISA, che si è dimostrato però poco efficace nel distinguere specie molto vicine ed inoltre richiede lo sviluppo di anticorpi specifici contro proteine specie-specifiche (5). L'impiego di metodi basati sull'analisi del DNA ha evidenziato numerosi vantaggi rispetto ai metodi su base proteica, tra questi una maggiore specificità, sensibilità, robustezza ed affidabilità negli alimenti trasformati, in quanto le molecole di DNA sono più termostabili delle proteine. Diversi metodi per l'analisi del DNA sono stati sviluppati: polimerase chain reaction restriction fragment length polymorphism (PCR-RFLP), forensically informative nucleotide sequencing (FINS), random amplification of polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP) o single-stranded conformational polymorphism (SSCP) (6, 7, 8, 9, 10, 11). Il metodo

PCR-sequencing permette di amplificare un tratto specifico del DNA e la sequenza ottenuta è confrontabile con tutte le sequenze disponibili e catalogate negli anni.

La maggior parte delle identificazioni sono realizzate mediante il confronto tra le sequenze ottenute e quelle disponibili nei database liberi, come la GenBank database realizzato dal National Center for Biotechnology Information (NCBI) e accessibile all'indirizzo: <http://www.ncbi.nlm.nih.gov>. Il database è una preziosa fonte di sequenze di differenti tratti di DNA di moltissime forme viventi; la mole di informazioni presenti, oltre 108 milioni di sequenze, sono anche un suo limite, infatti il database non è privo di errori di identificazione, mancanza di informazioni ed errori terminologici in quanto non è prevista nessuna forma di controllo sulle informazioni inserite. Da questo punto di vista la GenBank non è adatta per la sicura identificazione di specie, requisito fondamentale per la pratica ispettiva e forense. Per rispondere a questa esigenza sono nate diverse iniziative che si pongono l'obiettivo di costituire banche dati di riferimento, dove le sequenze presenti siano prodotte secondo un protocollo che garantisca dei requisiti minimi, che ogni sequenza sia controllata e verificata per minimizzare la presenza di errori. Consortium for the Barcoding of Life, FishTrace Consortium, Fishgen sono alcune delle iniziative nate per offrire un riferimento adatto all'identificazione di specie in campo ispettivo o forense. Il progetto Fishgen (European Commission, Joint Research Center Institute for Systems, Informatics and Safety & Environment Institute) ha come obiettivo di creare un database per l'identificazione dell'origine del pesce, permettendo l'identificazione dei diversi stock ittici con la combinazione di differenti tecniche molecolari (RFLP, DNA sequencing, microsatelliti di DNA e elettroforesi degli alloenzimi), al momento vi sono informazioni solo su 11 specie. Fishtrace è un consorzio finanziato dalla Comunità Europea che si pone l'obiettivo di fornire informazioni a partire dalle sequenze del gene mitocondriale del citocromo b e del gene nucleare della rodopsina, di individuare i polimorfismi utili a identificare le differenti popolazioni di pesci delle acque europee e fornire materiale biologico di riferimento. Fishgen e Fishtrace sono criticabili per due aspetti, sono eccessivamente origine specifici (acque europee) e considerano esclusivamente come obiettivo i pesci (escludono molluschi e crostacei) di conseguenza i dati disponibili sono eccessivamente ridotti rispetto agli scambi commerciali internazionali e al numero di prodotti della pesca commercializzati. I due progetti non rispondono pienamente alle esi-

genze del veterinario ispettore.

Il consorzio del Barcoding of Life si pone l'obiettivo, ambizioso, di raccogliere le sequenze del gene della citocromossidasi subunità I (COI), di tutte le forme di vita animale e se da una parte questo tipo di approccio è criticato dai tassonomisti (12), dall'altra la conoscenza della sequenza di un gene specifico per tutte le specie potenzia le tecniche identificative e permette di focalizzare gli sforzi nella prevenzione delle frodi. Barcoding of Life si adatta a tal punto a questo scopo che i ricercatori USFDA hanno inserito il barcode nella loro Regulatory Fish Encyclopedia (RFE) in quanto risponde alle esigenze ispettive, non si limita ad una specifica categoria animale, od ad un'area specifica. Inoltre Barcoding of Life impone un protocollo che gestisca i soggetti, le analisi e i risultati. Il Barcode of Life Data System (BOLD) rappresenta un ambiente sicuro per conservare questi dati, accessibili, per altro, all'intera comunità scientifica. Essendo l'obiettivo del BOLD quello di raccogliere ed organizzare tutte le sequenze degli esseri viventi, ogni ricercatore interessato è libero di accedere al sito www.boldsystems.org, e, come utente pubblico, può ottenere i dati di progetti pubblici ed usare l'Identification System. All'interno del BOLD System sono disponibili tre unità funzionali: il Management and Analysis System, l'Identification System e l'External Connectivity System. Ogni ricercatore interessato può creare un suo progetto dopo essersi registrato come BOLD user. Egli deve inserire il nome del progetto, un numero identificativo del campione ed un'attribuzione tassonomica (anche solo un riferimento al Phylum d'origine). Inoltre deve essere specificato il nome della specie a cui appartiene il campione, il numero con cui è stato catalogato, autore, data e luogo di reperimento, chi ha eseguito l'identificazione morfologica del campione, la sequenza COI con minimo 500bp, i primer usati per la PCR, e gli elettroferogrammi (trace file). Il BOLD System usa numerosi strumenti per trovare errori ed anomalie. Infatti tutte le sequenze inviate sono trasformate in aminoacidi e sono confrontate con un Hidden Markov Model della proteina codificata dal gene COI, per verificare che essi appartengano effettivamente a quel gene. Dopo aver passato questo primo controllo, le sequenze sono analizzate per individuare possibili codoni di stop e geni anomali, e anche per verificare che non sia presente DNA contaminante (principalmente umano). Se sono stati trovati errori, il mittente viene avvisato e la sequenza viene immediatamente cancellata. Dopo l'invio, i trace file vengono analizzati mediante un sistema a punteggio (PHRED score), che li classifica, a seconda della qualità, in: "failed" (non esiste la

sequenza); “low quality” (punteggio < 30); “medium quality” (punteggio tra 30 e 40); “high quality” (punteggio > 40). I dati validati dal BOLD sono archiviati nelle “pagina del campione”, che raccoglie vari dati riguardo al campione ed una o più sue immagini, e nella “pagina della sequenza”, collegata alla precedente, in cui vi sono la sequenza del Barcoding ed i primer utilizzati per la PCR. In questa pagina, inoltre, possono essere allegati documenti contenenti il protocollo PCR usato. Identification System permette il confronto di una sequenza con quelle contenute nel database; requisito minimo per la sequenza da confrontare è l'inclusione di almeno 300 bp della regione del COI. Il BOLD ha due funzioni di identificazione che agiscono all'interno di ogni progetto. La prima valuta la validità di identificazioni già esistenti, l'altra assegna un'identificazione ai campioni che non hanno una posizione tassonomica. La biblioteca di ricerca degli archivi del barcode è suddivisa in due categorie: una in cui sono raggruppate le specie con almeno tre individui e con una diversità conspecifica genica di massimo il 2%; l'altra categoria raggruppa tutti gli altri archivi, che possono essere consultati, ma non sono considerati validi dal sistema. L'Identification System del BOLD garantisce l'identificazione di una specie solo se la sequenza inviata presenta un'elevata corrispondenza con la sequenza di riferimento, in cui vi è meno dell' 1 % di diversità. Qualora non sia stata ottenuta una corrispondenza elevata con la sequenza di riferimento, cioè, nei casi in cui vi sia una diversità genica minore del 3%, la sequenza inviata viene assegnata ad un Genere. In tutti gli altri casi, il sistema espone i 100 taxa più simili alla sequenza e sintetizza la loro posizione nella gerarchia tassonomica. L'External Connectivity System usa protocolli standard per esporre la funzionalità del sistema ed i dati in esso contenuti a siti internet esterni che monitorano i progressi del barcode.

Questo sistema inizia ad essere adottato anche in Italia; nel 2010 Filonzi e coll.(13) hanno evidenziato numerose non conformità di etichettatura su di un campione di 69 filetti del commercio con sostituzioni di specie pregiate con altre di minor valore economico. Una nostra recente indagine (dati personali non pubblicati) ha confermato questa situazione: l'analisi di un gruppo di 75 preparati di prodotti della pesca (pesci e cefalopodi) campionati in diversi capoluoghi italiani ha rivelato un totale di n. 28 non corrispondenze, di cui 19 differenze anche nel valore commerciale mentre 9 casi erano, a nostro parere, riconducibili ad errori di etichettatura.

In conclusione, senza affermare che il sistema consigliato sia perfetto, il sistema BOLD si pre-

senta come uno strumento utile per l'identificazione di specie ai fini ispettivi e legali, offrendo migliori garanzie sulla qualità dei dati presenti, prevedendo un sistema di controllo che riduca i possibili errori nel database; tuttavia come emerso dai risultati ottenuti non è sempre possibile giungere all'identificazione di tutte le specie per la mancanza delle sequenze in quanto non ancora inserite nella banca dati.

Sulla base delle non corrispondenze di etichettatura verificate nella nostra ultima indagine e anche sulla base della nostra esperienza (14) vorremmo che si aprisse un dibattito sull'opportunità di rivedere il D.M. 31 gennaio 2008 e successive modifiche ed integrazioni allo scopo di semplificarlo e di renderlo maggiormente compatibile con le esigenze del mercato, senza trascurare l'interesse del consumatore. Il decreto negli anni si è arricchito con numerosi inserimenti, raggiungendo ad oggi circa 890 specie, si presenta estremamente dettagliato, approfondito e in grado di evidenziare tra specie simili anche minime differenze non percepite dal consumatore in termini di caratteristiche sensoriali e che non corrispondono a differenze nel valore commerciale, a fronte di notevoli complicazioni per gli operatori del settore. In tale contesto una revisione che semplifichi le denominazioni ufficiali italiane sarebbe auspicabile, allineandosi al percorso intrapreso da altri paesi comunitari. Ad esempio l'Inghilterra, nell'ultimo elenco di specie allegato al regolamento sull'etichettatura del pesce del 2010 (15), prevede la denominazione di “hake” per tutte le specie del genere *Merluccius*, mentre nel nostro attuale D.M. nel genere *Merluccius* sono indicate 12 differenti specie a cui corrispondono 6 differenti denominazioni ufficiali italiane. Analogamente, a fronte di 23 specie del genere *Epinephelus* previste dal D.M. catalogate con 6 denominazioni ufficiali italiane, la legge inglese prevede un unico nome “grouper” per tutte le specie appartenenti ai generi *Epinephelus* e *Mycteroperca*. Per quanto riguarda i cefalopodi la normativa inglese non fa nessuna distinzione all'interno dei generi *Sepia*, *Octopus*, *Loligo* e *Illex* adeguandosi probabilmente al livello medio di conoscenza del consumatore, che difficilmente distingue specie estremamente simili morfologicamente. Inoltre la normativa inglese in certi casi dà facoltà di indicare una specifica specie con un corrispondente nome più esatto, consentendo in questo caso una maggiore precisione a fronte di una maggiore responsabilità dell'operatore. A nostro parere sarebbe auspicabile una semplificazione delle denominazioni ufficiali italiane, che allo stato attuale non rispondono ad un'autentica esigenza di corretta informazione, ma complica a tal punto

l'etichettatura di questi prodotti che spesso è causa di errori involontari che vengono sanzionati, in assenza di vantaggi per il consumatore.

BIBLIOGRAFIA

1. Rehbein H. (1990). Electrophoretic techniques for species identification of fishery products. *Zeitschrift für Lebensmittel Untersuchung und Forschung*, 191, 1-10.
2. Righetti P.G. (1988). Elsevier Biomedical Press. Isoelectrofocusing: theory, methodology and application.
3. Cattaneo P. (1993). Riconoscimento delle specie carnee: una rassegna. *Ingegneria Alimentare*, 9(3), 25-33.
4. Rehbein H., Kuendiger R., Yman I.M., Ferm M., Etienne M., Jerome M., Craig A., Mackie I., Jessen F., Martinez I., Mendes R., Smelt A., Lutén J., Pineiro C., Perez-Martín R. (1999). Species identification of cooked fish by urea isoelectric focusing and sodium dodecylsulfate polycrylamide gel electrophoresis: a collaborative study. *Food chemistry*, 67, 333-339.
5. Carrera E., Garcia T., Cespedes A., Gonzalez I., Sanz B., Hernandez P. E., Martin R. (1997). Immunostick colorimetric ELISA assay for the identification of smoked salmon (*Salmo salar*), trout (*Oncorhynchus mykiss*) and bream (*Brama raii*). *Journal of the Science of Food and Agriculture*, 74, 547-550.
6. Ram, J.L., Ram, M.L. and Baidoun, F.F., (1996). Authentication of canned tuna and bonito by sequence and restriction site analysis of polymerase chain reaction products of mitochondrial DNA. *Journal of Agricultural Food Chemistry*, 44, 2460-2467.
7. Bottero, M. T., Civera, T., Anastasio, A., Turi, R. M., & Rosati, S. (2002). Identification of cows' milk in buffalo cheese by duplex polymerase chain reaction. *Journal of Food Protection*, 65, 362-366.
8. Bartlett S.E., Davidson W.S. (1992). FINS (forensically informative nucleotide sequencing): a procedure for identifying the animal origin of biological specimens. *Biotechniques*, 12(3), 408-411.
9. Colombo, F., Marchisio, E., Pizzini, A., & Cantoni, C. (2002). Identification of the goose species (*Anser anser*) in Italian Mortara salami by DNA sequencing and a polymerase chain reaction with an original primer pair. *Meat Science*, 61, 291-294.
10. Colombo, F., Vincava, R., Sacchelli, D., Colombo, M., & Camisasca, S. (1998). Differentiation of the ostrich (*Struthio camelus*) and emu (*Dromaius novaehollandiae*) species by polymerase chain reaction. I. PCR on the muscular tissue of *Struthio camelus*. *Industria Alimentari*, 37, 618-619.
11. Maldini M., Nonnis Marzano F., González Fortesa G., Papa R. and Gandolfi G. (2006). Fish and seafood traceability based on AFLP markers: Elaboration of a species database. *Aquaculture*, 261(2), 487-494.
12. DeSalle R., Egan M.G. and Siddall M. (2005). The unholy trinity: taxonomy, species delimitation and DNA barcoding. *Philosophical Transactions of the Royal society*, 360, 1905-1916.
13. Filonzi L., Chiesa S., Vaghi M., Nonnis Marzano F. (2010). Molecular barcoding reveals mislabelling of commercial fish products in Italy. *Food Research International*, 43, 1383-1388.
14. Cattaneo P., Balzaretto C., Bernardi C., Colombo F., Marzano M.A. (2007). Principali problematiche legate ai prodotti della pesca nella ristorazione collettiva. *Industria Alimentari*, 470, 625-632.
15. The Fish Labelling (England) Regulations 2010.