(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 392 610**

(21) Número de solicitud: 201130534

(51) Int. CI.:

C12Q 1/68 (2006.01)

(12) SOLICITUD DE PATENTE A1

(22) Fecha de presentación:

04.04.2011

(43) Fecha de publicación de la solicitud:

12.12.2012

(43) Fecha de publicación del folleto de la solicitud:

12.12.2012

(71) Solicitantes:

UNIVERSIDAD DEL PAIS VASCO - EUSKAL HERRIKO UNIBERTSITATEA (50.0%)
Bº Sarriena, s/n
48940 Leioa, Bizkaia, ES y
FUNDACION AZTI/AZTI FUNDAZIOA (50.0%)

(72) Inventor/es:

Estonba Rekalde, Andone;
Manzano Basabe, Carmen;
Zarraonaindia Martínez, Iratxe ;
Pardo González, Miguel Ángel;
Iriondo Orensanz , Mikel;
Escuredo García de Galdeano , Pedro Ramón y
Irigoyen Larrazabal , Xabier

(74) Agente/Representante:

CARPINTERO LÓPEZ, Mario

(54) Título: **MÉTODO DE IDENTIFICACIÓN GENÉTICA DE LA ANCHOA EUROPEA Y DE SU ORIGEN GEOGRÁFICO O POBLACIÓN.**

(57) Resumen:

Método de identificación genética de la anchoa europea y de su origen geográfico o población. La invención pertenece al campo del análisis genético de la anchoa europea o Engraulis encrasicolus. Más concretamente, la invención consiste en unos marcadores moleculares del tipo SNP (Single Nucleotide Polymorphisms) y un método genético que utiliza dicho panel de marcadores SNP para identificar la anchoa europea (Engraulis encrasicolus) y su origen geográfico o población en productos derivados de la pesca. Asimismo, la invención consiste en un kit que comprende las sondas necesarias para la realización del método de la invención, y adicionalmente los cebadores necesarios para la realización de dicho método.

ES 2 392 610 A1

DESCRIPCIÓN

MÉTODO DE IDENTIFICACIÓN GENÉTICA DE LA ANCHOA EUROPEA Y DE SU ORIGEN GEOGRÁFICO O POBLACIÓN

CAMPO DE LA INVENCION

La invención pertenece al campo del análisis genético de la anchoa europea o *Engraulis encrasicolus*. Más concretamente, la invención consiste en unos marcadores moleculares del tipo SNP (*Single Nucleotide Polymorphisms*) y un método genético que utiliza dicho panel de marcadores SNP para identificar la anchoa europea (*Engraulis encrasicolus*) y su origen geográfico o población en productos derivados de la pesca. Asimismo, la invención consiste en un kit que comprende las sondas necesarias para la realización del método de la invención, y adicionalmente los cebadores necesarios para la realización de dicho método.

ANTECEDENTES DE LA INVENCIÓN

La anchoa europea (*E. encrasicolus*) es un pequeño pez pelágico que vive en las aguas del Atlántico Oriental y del Mediterráneo, hasta 150 metros de profundidad. En los últimos años se ha registrado un notable incremento en la demanda de pescados azules, apoyados por fuertes campañas publicitarias. Datos conjuntos de consumo de anchoa y sardina en España, las colocan como las segundas más consumidas, tras las merluzas, con más de 125.000 toneladas/año. Económicamente hablando, además de la gran cantidad de anchoa que es consumida fresca, la comercialización de semiconservas de anchoas tiene una gran importancia en España, con un consumo cercano a 15.000 toneladas/año, así como en otros países europeos como Italia y Francia.

Las características sensoriales de firmeza y sabor ligeramente amargo de la anchoa europea hacen que esta especie sea muy valorada culinariamente en comparación con otras especies de anchoas. Más aún, esta especie de anchoa, y en particular la proveniente del Cantábrico, está especialmente bien considerada por la industria conservera debido a su firmeza en el proceso de fileteado y especialmente al intenso sabor que caracteriza el producto final. Estas características hacen que su valor en el mercado sea muy alto y por lo tanto muy apreciada por los consumidores.

La creciente actividad pesquera registrada en Europa durante los últimos años está ocasionando una considerable merma en la captura de los peces comercialmente más importantes, incluida la anchoa. Sin embargo, el consumo de productos derivados de la anchoa, como las semiconservas, salazones u otros, no decrece sino todo lo contrario. Esta escasez de la anchoa europea, y en particular de la anchoa del cantábrico, está ocasionando una creciente importación de especies de anchoas procedentes de otros países. De hecho, las principales especies del género *Engraulis* que se aprovechan en otros países son, además de la anchoa europea, la anchoa japonesa, *E. japonicus,* la de Australia y Nueva Zelanda, *E. australis,* en África del Sur, *E. capensis;* en la costa pacífica de Panamá, *E. clarki,* en las costas de Argentina, Uruguay y sur de Brasil, *E. anchoita* y la presente principalmente en las costas chilenas y peruanas, *E. ringens*. Sin embargo, de entre todas estas especies, las procedentes de Sudamérica son principalmente las que más rápida y ampliamente se están introduciendo en el mercado español y por extensión en el europeo.

El Reglamento (CE) nº 2065/2001 de la Comisión de 22 de octubre de 2001 establece las disposiciones de aplicación del Reglamento (CE) nº 104/2000 del Consejo en lo relativo a la información del consumidor en el sector de los productos de la pesca y de la acuicultura. Dicho reglamento incluye la obligatoriedad de informar en la etiqueta sobre el nombre científico, el método de producción (acuicultura o pesca extractiva) y la zona de captura del producto de origen. La Unión Europea está haciendo un gran esfuerzo en asegurar que se está cumpliendo la ley en lo que respecta a la correcta comercialización de los productos derivados de la pesca, para evitar fraudes en el etiquetado de las muestras y sobre todo como medida de protección de denominaciones de origen de productos con un alto estándar de calidad no siempre asociado únicamente a una especie específica sino también a un área o región geográfica concreta dentro de nuestras fronteras.

En cuanto a la identificación de especies de anchoa existen varios métodos descritos en la bibliografía (Sebastio *et al.,* 2001; Jérôme *et al.,* 2003; Ratnasingham y Hebert, 2007). Sin embargo, todavía no se han descrito métodos que permitan diferenciar poblaciones de especies de anchoas y por lo tanto tampoco su origen geográfico. Hasta el momento, se han realizado trabajos con marcadores moleculares del tipo de los alozimas (Spanakis *et al.,* 1989; Bembo *et al.,* 1996 a, b; Tudela *et al.,* 1999) y los polimorfismos en la longitud de fragmentos de restricción del ADN nuclear y mitocondrial (Bembo *et al.,* 1995; Magoulas *et al.,* 1996, 2006; Borsa 2002; Grant, 2005; Borsa *et al.,* 2004), siendo ampliamente utilizados en estudios genéticos de la anchoa, pero con un poder limitado para detectar diferenciación poblacional en esta especie. Recientemente, la aplicación de marcadores de ADN

microsatélites, altamente variables, ha probado ser informativa para la detección de subpoblaciones en especies de anchoas pero sin obtener resultados concluyentes (Chiu *et al.,* 2002; Liu *et al.,* 2004; Yu *et al.,* 2002, 2005; Landi *et al.,* 2005; Zarraonaindia *et al.,* 2009). De hecho, para poder obtener resultados con un alto poder de asignación a una población de anchoa determinada, habría que aumentar el número de marcadores microsatélites. Sin embargo, este tipo de marcadores no son apropiados para determinar el origen geográfico de un producto procesado como puede ser, por ejemplo, una semiconserva o una pasta de anchoa, ya que el procesado degrada el ADN dificultando y/o impidiendo la amplificación de la región de interés.

La alternativa a los microsatélites son, sin ningún lugar a duda, los polimorfismos de un único nucleótido, comúnmente denominados como SNP. Los SNP representan la clase más abundante de polimorfismos en cualquier organismo, son adaptables a la automatización y revelan polimorfismos no detectados por otro tipo de marcadores y métodos. Además, la tasa de mutación de este tipo de marcador es significativamente menor a la de los microsatélites (Landegren *et al.,* 1998; Nielsen 2000). Esta estabilidad es ventajosa en estudios evolutivos, de genética de poblaciones y en estudios de pedigrís y asignación individual (Rengmark *et al*., 2005). Por otro lado, y dada la naturaleza dialélica de este tipo de marcador, la información que aporta cada SNP es menor a la de los *loci* microsatélites, pero el contenido informativo puede aumentarse mediante la inclusión de un mayor número de SNP (Glaubitz *et al*., 2003). En este sentido, Rengmark *et al.* (2006), realizaron un estudio comparativo del poder y eficiencia de microsatélites y SNP para estudios de poblaciones, análisis de parentesco y la asignación individual en salmón. En él concluyeron que los SNP mostraban resultados semejantes a los microsatélites y que dado que posibilitan un genotipado de alto rendimiento podrían llegar a ser los marcadores de elección en un futuro próximo.

Además, el uso de marcadores SNP posibilita el análisis de muestras muy degradadas (Römpler *et al.,* 2006), lo que los hace el marcador ideal para realizar la identificación del origen geográfico o poblaciones de anchoa europea, ya que algunos productos derivados de la pesca, como por ejemplo sucede con el caso de las semiconservas, presentan el ADN muy degradado debido al procesamiento al que se ven sometidos. Así, los fragmentos de ADN que se pueden amplificar a partir de este tipo de muestras son muy pequeños dificultando, e incluso imposibilitando, la utilización de ciertos marcadores genéticos como los microsatélites. Por el contrario, cuando se utilizan marcadores del tipo SNP, se pueden genotipar fragmentos de ADN muy pequeños (menos de 100 pb), por lo que son excelentes marcadores genéticos para muestras muy degradadas.

Hasta la fecha, no tenemos constancia de ningún método genético que utilice un panel de SNP específico para la identificación del origen geográfico o poblaciones de la anchoa europea (*E. encrasicolus*) en productos derivados de la pesca, ya sea fresco, congelado y mucho menos procesado. Por lo tanto, la invención aquí presentada es del todo novedosa y original.

OBJETO DE LA INVENCIÓN

La invención consiste en un panel de 18 marcadores de tipo SNP que permiten, en primer lugar, la identificación de la especie de anchoa europea (*Engraulis encrasicolus)* frente al resto de especies de anchoas pertenecientes al género *Engraulis* y en segundo lugar, la identificación del origen geográfico o poblaciones de la anchoa europea, en particular la población del Atlántico Este, del Golfo de Vizcaya o del Mediterráneo. Además, la invención consiste en el método para llevar a cabo dicha identificación y los cebadores y sondas necesarios para realizarlo. Asimismo, la invención consiste en un kit que comprende las sondas necesarias para la realización del método de la invención, y adicionalmente los cebadores necesarios para la realización del método de la invención.

DESCRIPCIÓN DETALLADA DE LA INVENCIÓN

La invención consta de un conjunto de secuencias de ADN SEQ ID NO 01 a SEQ ID NO 15 caracterizadas por que comprenden los siguientes SNP: SNP1 y SNP2 (SEQ ID NO 01), SNP3 (SEQ ID NO 02), SNP4 y SNP14 (SEQ ID NO 03), SNP5 (SEQ ID NO 04), SNP6 (SEQ ID NO 05), SNP7 (SEQ ID NO 06), SNP15 (SEQ ID NO 07), SNP16 (SEQ ID NO 08), SNP17 (SEQ ID NO 09), SNP18 (SEQ ID NO 10), SNP8 y SNP9 (SEQ ID NO 11), SNP10 (SEQ ID NO 12), SNP11 (SEQ ID NO 13), SNP12 (SEQ ID NO 14) y SNP13 (SEQ ID NO 15) (TABLA 1). Asimismo, es objeto de la invención el uso de dichos SNP como marcadores para identificar la especie *E. encrasicolus* y su origen geográfico o población.

Las secuencias SEQ ID NO 1 y SEQ ID NO 9 codifican fragmentos del gen mitocondrial citocromo b, las secuencias SEQ ID NO 7, SEQ ID NO 8, SEQ ID NO 10 y SEQ ID NO 11 codifican fragmentos del gen ribosómico mitocondrial 12S, del gen catepsina L, de la rodopsina y de la creatina-quinasa, respectivamente. El resto de secuencias se corresponden con clones que forman parte de una librería genómica parcial (realizada en la UPV/EHU). Las secuencias de estos clones no muestran homologías con genes que figuran en la base de datos NCBI, por lo que no se les atribuye una función concreta. Los SNP se

localizan en dichas secuencias en las posiciones especificadas en la TABLA 1. Los distintos marcadores, SNP1 a SNP18, se identificaron cómo se indica a continuación:

Los marcadores SNP1 y SNP2 se encontraron tras alinear secuencias de anchoas pertenecientes al género *Engraulis* publicadas en el *National Center for Biotechnology Information* (NCBI; http://www.ncbi.nlm.nih.gov). Estos SNP fueron obtenidos tras amplificar y secuenciar un fragmento del gen mitocondrial citocromo b de 307 pares de bases (pb), de diferentes muestras de especies de anchoas, con los oligos L14841 (SEQ ID NO 107) y H15149 (SEQ ID NO 108), tal y como se describe en Kocher *et al.* (1989). Tras este análisis se obtuvo una secuencia consenso de 157 pb de la especie *E. encrasicolus* denominada CYTb-157 (SEQ ID NO 01, donde los SNP1 y 2 se encuentran en las posiciones 90 y 96, respectivamente (posiciones 468 y 474 del gen mitocondrial citocromo b, respectivamente).

El marcador SNP3 se obtuvo tras amplificar un fragmento denominado clon-369 (SEQ ID NO 02) con los cebadores clon369-F (SEQ ID NO 16) y clon369-R (SEQ ID NO 17). Dicho SNP3 se encuentra en la posición 220 de la SEQ ID NO 02.

Los marcadores SNP4 y SNP14 se obtuvieron tras amplificar un fragmento denominado clon-286 (SEQ ID NO 03) con los cebadores clon286-F (SEQ ID NO 18) y clon286-R (SEQ ID NO 19). Los SNP4 y SNP14 se encuentran en las posiciones 531 y 449, respectivamente.

El marcador SNP5 se obtuvo tras amplificar un fragmento denominado clon-491 (SEQ ID NO 04) con los cebadores clon491-F (SEQ ID NO 20) y clon491-R (SEQ ID NO 21). Dicho SNP5 se encuentra en la posición 206 de la SEQ ID NO 04.

El marcador SNP6 se obtuvo tras amplificar un fragmento denominado clon-M3EK (SEQ ID NO 05) con los cebadores clonM3EK-F (SEQ ID NO 22) y clonM3EK-R (SEQ ID NO 23). Dicho SNP6 se localiza en la posición 166 de la SEQ ID NO 05.

El marcador SNP7 se obtuvo tras amplificar un fragmento denominado clon-289 (SEQ ID NO 06) con los cebadores clon289-F (SEQ ID NO 24) y clon289-R (SEQ ID NO 25). El SNP7 ocupa la posición 494 de la SEQ ID NO 06.

Los marcadores SNP8 y SNP9 se obtuvieron tras amplificar un fragmento denominado MT-12S (SEQ ID NO 07) del gen ribosómico mitocondrial 12S, con los cebadores MT-12S-F (SEQ ID NO 34) y MT-12S-R (SEQ ID NO 35). Los SNP8 y SNP9 están en las posiciones 102 y 198 de la SEQ ID NO 07.

Tabla 1.- Descripción de los 18 marcadores genéticos de la invención.

A nivel de especie *Engraulis* (2 SNP)					
Gen	**Polinucleótido con SNP**	**SNP**			***Engraulis encrasicolus***
		Nº	**base**	**posición**	
CYTb-157	SEQ ID NO 01	1	G/A	90	G
		2	T/C	96	T
A nivel de población *Engraulis encrasicolus* (11 SNP, fases i) y ii) etapa c))					
Gen/clon	**Polinucleótido con SNP**	**SNP**			**Atlántico Este vs Vizcaya/Mediterráneo**
		Nº	**base**	**posición**	
clon-369	SEQ ID NO 02	3	A/C	220	
clon-286	SEQ ID NO 03	4	G/C	531	
clon-491	SEQ ID NO 04	5	A/G	206	
clon-M3EK	SEQ ID NO 05	6	C/T	166	
clon-289	SEQ ID NO 06	7	G/A	494	
MT-12S	SEQ ID NO 07	8	G/A	102	
MT-12S	SEQ ID NO 07	9	G/A	198	
CATL-7	SEQ ID NO 08	10	G/A	72	
CYTb	SEQ ID NO 09	11	C/T	516	
Rodopsina	SEQ ID NO 10	12	G/A	127	
CK6-1	SEQ ID NO 11	13	C/T	276	
A nivel de población *Engraulis encrasicolus* (9 SNP, fases iii) y iv) etapa c))					
Gen/clon	**Polinucleótido con SNP**	**SNP**			**Vizcaya vs Mediterráneo**
		Nº	**base**	**posición**	
clon-289	SEQ ID NO 6	7	G/A	494	
CYTb	SEQ ID NO 9	11	C/T	516	
Rodopsina	SEQ ID NO 10	12	G/A	127	
CK6-1	SEQ ID NO 11	13	C/T	276	
clon-286	SEQ ID NO 3	14	T/C	449	
clon-470	SEQ ID NO 12	15	A/T	403	
clon-357	SEQ ID NO 13	16	G/C	506	
clon-M4	SEQ ID NO 14	17	C/T	352	
clon-483	SEQ ID NO 15	18	C/T	165	

El marcador SNP10 se obtuvo tras amplificar un fragmento denominado CATL-7 (SEQ ID NO 08) del gen catepsina L, con los cebadores CATL7-F (SEQ ID NO 36) y CATL7-R (SEQ ID NO 37). Dicho SNP10 se encuentra en la posición 72 de la SEQ ID NO 08.

El marcador SNP11 se obtuvo tras amplificar un fragmento denominado CYTb (SEQ ID NO 09) del gen mitocondrial citocromo b, con los cebadores CytBI-7F (SEQ ID NO 105) y Trucytb-R (SEQ ID NO 106) previamente descritos en Jérôme *et al.* 2003b y siguiendo las indicaciones de la referencia Jérôme *et al.* 2008. El SNP11 está en la posición 516 de la SEQ ID NO 09.

El marcador SNP12 se obtuvo tras alinear las siguientes secuencias del NCBI (http://www.ncbi.nlm.nih.gov) 159462259, 159462261, 161089181 y 161089183, que codifican la rodopsina de *Engraulis encrasicolus*. La secuencia consenso se denominó Rodopsina (SEQ ID NO 10) y el SNP12 se localiza en su posición 127.

El marcador SNP13 se obtuvo tras alinear las siguientes secuencias del NCBI (http://www.ncbi.nlm.nih.gov) AY486347.1 y AY486346.1, que codifican la creatin-quinasa de *Engraulis encrasicolus*, siendo la secuencia consenso CK6-1 (SEQ ID NO 11). El SNP13 se encuentra en la posición 276 de la SEQ ID NO 11.

El marcador SNP15 se obtuvo tras amplificar un fragmento denominado clon-470 (SEQ ID NO 12) con los cebadores clon470-F (SEQ ID NO 26) y clon470-R (SEQ ID NO 27). Dicho SNP15 se encuentra en la posición 403 de la SEQ ID NO 12.

El marcador SNP16 se obtuvo tras amplificar un fragmento denominado clon-357 (SEQ ID NO 13) con los cebadores clon357-F (SEQ ID NO 28) y clon357-R (SEQ ID NO 29). El SNP16 se localiza en la posición 506 de la SEQ ID NO 13.

El marcador SNP17 se obtuvo tras amplificar un fragmento denominado clon-M4 (SEQ ID NO 14) con los cebadores clonM4-F (SEQ ID NO 30) y clonM4-R (SEQ ID NO 31). Dicho SNP17 está en la posición 352 de la SEQ ID NO 14.

Por último, el marcador SNP18 se obtuvo tras amplificar un fragmento denominado clon-483 (SEQ ID NO 15) con los cebadores clon483-F (SEQ ID NO 32) y clon483-R (SEQ ID NO 33). El SNP18 se localiza en la posición 165 de la SEQ ID NO 15.

Otro aspecto de la invención consiste en el uso de los SNP: SNP1 a SNP18, comprendidos en las secuencias SEQ ID NO 01 a SEQ ID NO 15, como marcadores para identificar la especie *E. encrasicolus* y su origen geográfico o población.

A este respecto, la invención comprende un método de identificación genética del origen geográfico y poblaciones de la anchoa europea (*Engraulis encrasicolus*) que comprende:

a) Extracción del ADN de la muestra de un producto derivado de la anchoa

b) Identificación de la especie *Engraulis encrasicolus* frente al resto de especies de anchoas pertenecientes al género *Engraulis* que comprende la amplificación y genotipado de un fragmento que contenga el SNP1 y el SNP2 (SEQ ID NO 01), por diferentes metodologías, donde la presencia en las posiciones 468 y 474 del gen mitocondrial citocromo b de una G y una T (posiciones 90 y 96 de la SEQ ID NO 01), respectivamente, es indicativa de la especie *E. encrasicolus*

c) Identificación del origen geográfico o poblaciones de la anchoa europea (*E. encrasicolus*) que comprende:

   i. amplificación y genotipado de los siguientes 11 SNP: SNP3 a SNP13, comprendidos en las secuencias SEQ ID NO 2 a SEQ ID NO 13

   ii. análisis de los resultados obtenidos en la fase i) y asignación del genotipo problema a una población de referencia: Atlántico Este o Mediterráneo/Golfo de Vizcaya

   iii. amplificación y genotipado de los siguientes 9 SNP: SNP7 y SNP11 a SNP18

   iv. análisis de los resultados obtenidos en la fase iii) y asignación de la muestra problema a una población de referencia: Mediterráneo o Golfo de Vizcaya.

La etapa a) del método de la invención comprende la extracción del ADN de la muestra a analizar. Dicha muestra de ADN se puede obtener de pescado fresco, congelado, salazones, semiconservas, procesado o de una mezcla de los mismos. Cualquier método de extracción comúnmente utilizado por un experto medio en la materia es apropiado en el contexto de la presente invención, como por ejemplo los kits comerciales *Wizard DNA Clean Up* (Promega), *Nucleospin* (Clontech), *GenomicPrep* (Amersham Pharmacia Biotech) tal y como se describe en Chapela *et al.* (2007).

El ADN extraído debe estar en cantidad y calidad suficientes para el análisis. Los métodos para calcular la cantidad y la calidad del ADN extraído son conocidos por el experto en la materia. Así, la cantidad y calidad del ADN de las muestras se puede calcular, por ejemplo, utilizando el ratio absorbancia a 260nm/absorbancia a 280nm (A260/A280) (Chapela *et al.* 2007).

La etapa b) del método de la invención comprende la identificación de la especie *E. encrasicolus* frente a otras especies de anchoas pertenecientes al género *Engraulis.* Para ello, se examinan dos marcadores genéticos del tipo SNP situados en las posiciones 468 y 474 del gen mitocondrial citocromo b; SNP1 y SNP2, respectivamente. La amplificación y detección de los SNP1 y SNP2, por diferentes metodologías, permite la identificación de la especie *E. encrasicolus* de otras especies de anchoas pertenecientes al género *Engraulis,* como por ejemplo, *E. ringens*, *E. anchoita*, *E. japonicus* y *E. mordax.* La presencia del genotipo G en la posición 468 del gen mitocondrial citocromo b (posición 90 en la SEQ ID NO 01) y T en la posición 474 del gen mitocondrial citocromo b (posición 96 en la SEQ ID NO 01) es indicativa de la especie *E. encrasicolus.* Una vez identificada la especie *E. encrasicolus* se realiza la etapa c) del método de la invención para identificar el origen geográfico o población de la anchoa europea. En caso de que la muestra problema no sea identificada como perteneciente a la especie *E. encrasicolus* no se continuará con la etapa c) y se concluirá que la muestra problema no se corresponde con la anchoa europea.

En una realización particular de la etapa b) del método de la invención el genotipado de los SNP1 y SNP2 se lleva a cabo con la sonda SNP1,2-FAM (SEQ ID NO 72).

En otra realización particular de la etapa b) del método de la invención, la amplificación del fragmento que contiene los SNP1 y SNP2 se lleva a cabo con los cebadores SNP1,2-F (SEQ ID NO 38) y SNP1,2-R (SEQ ID NO 39).

La etapa c) del método de la invención comprende examinar, en cuatro fases (i-iv), un total de 16 marcadores genéticos del tipo SNP situados en diferentes posiciones del genoma. La amplificación y genotipado por diferentes metodologías, comúnmente utilizadas por el experto en la materia (entre otras, la reacción en cadena de la polimerasa (PCR), la PCR a tiempo real o la tecnología *Taqman Open Array*) de los 16 marcadores genéticos SNP3 a SNP18 mencionados, permiten identificar el origen geográfico o poblaciones de la anchoa europea.

En la fase i) de la etapa c) del método de la invención se amplifican y genotipan los siguientes 11 SNP: SNP3 a SNP13, comprendidos en las secuencias SEQ ID NO 2 a SEQ ID NO 13 (TABLA 1).

En una realización particular de la fase i) de la etapa c) del método de la invención, el genotipado de los SNP3 a SNP13 se lleva a cabo con las siguientes sondas; SNP3 con las sondas SNP3-VIC (SEQ ID NO 73) y SNP3-FAM (SEQ ID NO 74); SNP4 con las sondas

SNP4-VIC (SEQ ID NO 75) y SNP4-FAM (SEQ ID NO 76); SNP5 con las sondas SNP5-VIC (SEQ ID NO 83) y SNP5-FAM (SEQ ID NO 84); SNP6 con las sondas SNP6-VIC (SEQ ID NO 89) y SNP6-FAM (SEQ ID NO 90); SNP7 con las sondas SNP7-VIC (SEQ ID NO 93) y SNP7-FAM (SEQ ID NO 94); SNP8 con las sondas SNP8-VIC (SEQ ID NO 77) y SNP8-FAM (SEQ ID NO 78); SNP9 con las sondas SNP9-VIC (SEQ ID NO 79) y SNP9-FAM (SEQ ID NO 80); SNP10 con las sondas SNP10-VIC (SEQ ID NO 91) y SNP10-FAM (SEQ ID NO 92); SNP11 con las sondas SNP11-VIC (SEQ ID NO 81) y SNP11-FAM (SEQ ID NO 82); SNP12 con las sondas SNP12-VIC (SEQ ID NO 85) y SNP12-FAM (SEQ ID NO 86); SNP13 con las sondas SNP13-VIC (SEQ ID NO 87) y SNP13-FAM (SEQ ID NO 88).

En otra realización particular de la fase i) de la etapa c) del método de la invención, la amplificación de los SNP3 a SNP13 se lleva a cabo con los cebadores: SNP3-F (SEQ ID NO 40) y SNP3-R (SEQ ID NO 41) para el SNP3; SNP4-F (SEQ ID NO 42) y SNP4-R (SEQ ID NO 43) para el SNP4; SNP5-F (SEQ ID NO 50) y SNP5-R (SEQ ID NO 51) para el SNP5; SNP6-F (SEQ ID NO 56) y SNP6-R (SEQ ID NO 57) para el SNP6; SNP7-F (SEQ ID NO 60) y SNP7-R (SEQ ID NO 61) para el SNP7; SNP8-F (SEQ ID NO 44) y SNP8-R (SEQ ID NO 45) para el SNP8; SNP9-F (SEQ ID NO 46) y SNP9-R (SEQ ID NO 47) para el SNP9; SNP10-F (SEQ ID NO 58) y SNP10-R (SEQ ID NO 59) para el SNP10; SNP11-F (SEQ ID NO 48) y SNP11-R (SEQ ID NO 49) para el SNP11; SNP12-F (SEQ ID NO 52) y SNP12-R (SEQ ID NO 53) para el SNP12; SNP13-F (SEQ ID NO 54) y SNP13-R (SEQ ID NO 55) para el SNP13.

En la siguiente fase del método de la invención, fase ii) de la etapa c), se lleva a cabo el análisis de los resultados obtenidos en la fase i), mediante programas informáticos de asignación de individuos a una población. El procedimiento a seguir para la obtención de porcentajes de asignación correcta de individuos a su población de origen se basa en la comparación entre dos tipos de paneles: “panel de referencia” y “panel de asignación”. En el “panel de referencia” se incluyen individuos tomados al azar que pertenecen a cada una de las siguientes poblaciones: Atlántico Este, Mediterráneo o Golfo de Vizcaya (referida como Vizcaya de aquí en adelante). Los paneles de referencia incluyen de 50-100 individuos en cada uno de los grupos, siendo lo más importante que estén equilibrados, es decir, que el número mínimo de individuos lo marca aquella población que tenga el número más bajo de individuos en cada caso particular.

Existe también un “panel de referencia” que incluye dos poblaciones, la del Mediterráneo y la del Golfo de Vizcaya (referido como Vizcaya/Mediterráneo) para el que se han tomado al

azar el mismo número de individuos de cada una de estas poblaciones. Los individuos del "panel de asignación" (individuos problema) son cotejados, en base a su genotipo, con los individuos del "panel de referencia". Dependiendo de la etapa del método de la invención que se esté llevando a cabo, en el "panel de referencia" se incluyen unas poblaciones u otras.

Con objeto de minimizar el error a la hora de asignar un origen a un individuo problema, se determina un *score* umbral de probabilidad de asignación (*score* umbral, de aquí en adelante). Dicho *score* umbral se define como el valor mínimo que debe alcanzar el *score* de probabilidad de asignación para que el programa de asignación considere una muestra como asignada a alguna de las poblaciones de referencia. Asimismo, se define el *score* de probabilidad de asignación (*score,* de aquí en adelante) como la probabilidad de pertenencia a cada una de las poblaciones de referencia calculada a partir del resultado de los SNP analizados. Así, cada individuo problema tiene dos *scores* de asignación (o probabilidad de pertenencia), uno para cada grupo poblacional incluido en el "panel de referencia", siendo la suma de ambos 100%. Por último, destacar que si la probabilidad de pertenencia a la población es inferior al umbral (*score* inferior al *score* umbral), la muestra se considerará como no asignada.

Se define por "asignación incorrecta" el hecho de que perteneciendo un individuo problema a una población de referencia se haya clasificado incorrectamente como perteneciente a otra población. A mayor *score* umbral, en general, menor error de asignación incorrecta, pero también menor porcentaje de individuos asignados, ya que los individuos que no superen el valor de *score* umbral no serán asignados a ninguna de las poblaciones de referencia. De esta manera, el *score* umbral puede ser distinto dependiendo del error de asignación incorrecta que se quiera asumir, así como de la proporción de individuos que se quiera asignar. Tras la realización de distintos análisis, se han elegido una serie de *scores* umbrales para cada una de las fases de análisis (fases ii) y iv)) de la etapa c) del método de la invención, siendo los más preferidos los óptimos en cuanto a asignación correcta y número de individuos asignados.

En la fase ii) de la etapa c) del método de la invención, el "panel de referencia" consta de la población Atlántico Este *versus* la población Vizcaya/Mediterráneo. El *score* umbral elegido es 70%. En una realización preferente dicho score umbral es 85% y en una realización aún más preferente es 95%. Al utilizar un *score* umbral de 95%, una asignación con un s*core* mayor o igual a 95 % a una de las poblaciones de referencia (Atlántico Este o Vizcaya/Mediterráneo) es indicativa de pertenencia a dicha población. Utilizando dicho

score umbral (95%), el porcentaje de individuos asignados a la población Atlántico Este es del 81,5%, con un porcentaje de asignación correcta del 98,4%. En cuanto a la población Vizcaya/Mediterráneo, el porcentaje de individuos asignados es del 80%, con un porcentaje de asignación correcta del 99%.

Cuando el genotipo problema se asigna a la población del Mediterráneo/Golfo de Vizcaya se lleva a cabo la siguiente fase iii) del método de la invención. En caso contrario, no se realiza la fase siguiente y se concluye que el genotipo problema no pertenece ni a la población del Mediterráneo ni a la población del Golfo de Vizcaya.

La fase iii) de la etapa c) del método de la invención se realiza a aquellos individuos problema que hayan sido identificados en la fase ii) dentro de la población Vizcaya/Mediterráneo. En esta fase iii) se amplifican y genotipan los siguientes 9 SNP: SNP7, SNP11 a SNP18, comprendidos en las SEQ ID NO 6, SEQ ID NO 9 a SEQ ID NO 11, SEQ ID NO 3 y SEQ ID NO 12 a SEQ ID NO 15, respectivamente (TABLA 1).

En una realización particular de la fase iii) de la etapa c) del método de la invención, el genotipado de los SNP7 y SNP11 a SNP18 se lleva a cabo con las siguientes sondas: SNP7 con las sondas SNP7-VIC (SEQ ID NO 93) y SNP7-FAM (SEQ ID NO 94); SNP11 con las sondas SNP11-VIC (SEQ ID NO 81) y SNP11-FAM (SEQ ID NO 82); SNP12 con las sondas SNP12-VIC (SEQ ID NO 85) y SNP12-FAM (SEQ ID NO 86); SNP13 con las sondas SNP13-VIC (SEQ ID NO 87) y SNP13-FAM (SEQ ID NO 88); SNP14 con las sondas SNP14-VIC (SEQ ID NO 99) y SNP14-FAM (SEQ ID NO 100); SNP15 con las sondas SNP15-VIC (SEQ ID NO 95) y SNP15-FAM (SEQ ID NO 96); SNP16 con las sondas SNP16-VIC (SEQ ID NO 97) y SNP16-FAM (SEQ ID NO 98); SNP17 con las sondas SNP17-VIC (SEQ ID NO 103) y SNP17-FAM (SEQ ID NO 104); SNP18 con las sondas SNP18-VIC (SEQ ID NO 101) y SNP18-FAM (SEQ ID NO 102).

En otra realización particular de la fase iii) de la etapa c) del método de la invención, se utilizan los siguientes cebadores: SNP7-F (SEQ ID NO 60) y SNP7-R (SEQ ID NO 61) para el SNP7; SNP11-F (SEQ ID NO 48) y SNP11-R (SEQ ID NO 49) para el SNP11; SNP12-F (SEQ ID NO 52) y SNP12-R (SEQ ID NO 53) para el SNP12; SNP13-F (SEQ ID NO 54) y SNP13-R (SEQ ID NO 55) para el SNP13; SNP14-F (SEQ ID NO 66) y SNP14-R (SEQ ID NO 67) para el SNP14; SNP15-F (SEQ ID NO 62) y SNP15-R (SEQ ID NO 63) para el SNP15; SNP16-F (SEQ ID NO 64) y SNP16-R (SEQ ID NO 65) para el SNP16; SNP17-F (SEQ ID NO 70) y SNP17-R (SEQ ID NO 71) para el SNP17; SNP18-F (SEQ ID NO 68) y SNP18-R (SEQ ID NO 69) para el SNP18.

Por último, en la fase iv) de la etapa c) del método de la invención el objetivo es discriminar el origen de la muestra problema entre el Golfo de Vizcaya y el Mediterráneo, por lo tanto el "panel de referencia" es la población Vizcaya *versus* la población Mediterráneo. En esta fase, el *score* umbral es 70%. En una realización preferente dicho *score* umbral es 80% y en una realización aún más preferente es 85%. En este último caso, una asignación con un *score* mayor o igual a 85% a una población de referencia dada (Mediterráneo o Vizcaya) es indicativa de pertenencia a dicha población. Con este *score* umbral de 85%, el porcentaje de individuos asignados a la población del Golfo de Vizcaya es del 53%, con un porcentaje de asignación correcta del 94,4%. En cuanto a la población del Mediterráneo, el porcentaje de individuos asignados es del 56,5%, con un porcentaje de asignación correcta del 96,7%.

Como alternativa para minimizar el error de asignación, pero en detrimento del número de individuos asignados, el *score* umbral seleccionado en la fase iv) puede ser de 99%, en cuyo caso el porcentaje de asignación correcta es del 99,99%, sin embargo el porcentaje de asignados desciende hasta un 12,5% para la población del Golfo de Vizcaya y un 15,7% para la población del Mediterráneo.

Cualquier algoritmo implementado en un programa informático de análisis poblacional puede ser utilizado para analizar los resultados de genotipado del panel de marcadores incluidos en la invención (fases de análisis ii) y iv) de la etapa c) del método de la invención), así como los programas informáticos de asignación que implementan estos algoritmos. Los métodos más usados en asignación son los basados en frecuencias, métodos Bayesianos, Distancias y *Clustering* (Hauser *et al.,* 2006). Los siguientes programas son un ejemplo y por lo tanto no excluyen otros programas que pudieran aparecer en el futuro: WHICHRUN 3.2, GENECLASS 1.0.02, ASSIGNMENT CALCULATOR, G-STAR, ARLEQUIN2.00, IMMANC, STRUCTURE, Excel spreadsheet, GMA, BAYES y BAPS (Banks y Eichert, 2000; Cornuet *et al.,* 1996; Paetkau *et al.*, 1995; Siefismund, 1995; Schneider *et al.,* 2000; Rannala y Mountain, 1997; Pritchard *et al.,* 2000; Bernatchez y Duchesne, 2000; Kalinowski, 2003; Pella y Masuda, 2001; Corander *et al.,* 2003).

En una realización preferida del método de la invención, la amplificación y detección de los SNP1 a SNP18 (etapas b) y c)), se lleva a cabo utilizando los cebadores SEQ ID NO 38 a SEQ ID NO 71 y las sondas SEQ ID NO 72 a SEQ ID NO 104, tal y como se indica a continuación y como se especifica en la TABLA 2. Para la amplificación y genotipado del

SNP1 y del SNP2 se utilizan los cebadores SNP1,2-F (SEQ ID NO 38) y SNP1,2-R (SEQ ID NO 39) y las sonda SNP1,2-FAM (SEQ ID NO 72); para el SNP3 los cebadores SNP3-F (SEQ ID NO 40) y SNP3-R (SEQ ID NO 41) y las sondas SNP3-VIC (SEQ ID NO 73) y SNP3-FAM (SEQ ID NO 74); para el SNP4 los cebadores SNP4-F (SEQ ID NO 42) y SNP4-R (SEQ ID NO 43) y las sondas SNP4-VIC (SEQ ID NO 75) y SNP4-FAM (SEQ ID NO 76); para el SNP5 los cebadores SNP5-F (SEQ ID NO 50) y SNP5-R (SEQ ID NO 51) y las sondas SNP5-VIC (SEQ ID NO 83) y SNP5-FAM (SEQ ID NO 84); para el SNP6 los cebadores SNP6-F (SEQ ID NO 56) y SNP6-R (SEQ ID NO 57) y las sondas SNP6-VIC (SEQ ID NO 89) y SNP6-FAM (SEQ ID NO 90); para el SNP7 los cebadores SNP7-F (SEQ ID NO 60) y SNP7-R (SEQ ID NO 61) y las sondas SNP7-VIC (SEQ ID NO 93) y SNP7-FAM (SEQ ID NO 94); para el SNP8 los cebadores SNP8-F (SEQ ID NO 44) y SNP8-R (SEQ ID NO 45) y las sondas SNP8-VIC (SEQ ID NO 77) y SNP8-FAM (SEQ ID NO 78); para el SNP9 los cebadores SNP9-F (SEQ ID NO 46) y SNP9-R (SEQ ID NO 47) y las sondas SNP9-VIC (SEQ ID NO 79) y SNP9-FAM (SEQ ID NO 80); para el SNP10 los cebadores SNP10-F (SEQ ID NO 58) y SNP10-R (SEQ ID NO 59) y las sondas SNP10-VIC (SEQ ID NO 91) y SNP10-FAM (SEQ ID NO 92); para el SNP11 los cebadores SNP11-F (SEQ ID NO 48) y SNP11-R (SEQ ID NO 49) y las sondas SNP11-VIC (SEQ ID NO 81) y SNP11-FAM (SEQ ID NO 82); para el SNP12 los cebadores SNP12-F (SEQ ID NO 52) y SNP12-R (SEQ ID NO 53) y las sondas SNP12-VIC (SEQ ID NO 85) y SNP12-FAM (SEQ ID NO 86), para el SNP13 los cebadores SNP13-F (SEQ ID NO 54) y SNP13-R (SEQ ID NO 55) y las sondas SNP13-VIC (SEQ ID NO 87) y SNP13-FAM (SEQ ID NO 88); para el SNP14 los cebadores SNP14-F (SEQ ID NO 66) y SNP14-R (SEQ ID NO 67) y las sondas SNP14-VIC (SEQ ID NO 99) y SNP14-FAM (SEQ ID NO 100); para el SNP15 los cebadores SNP15-F (SEQ ID NO 62) y SNP15-R (SEQ ID NO 63) y las sondas SNP15-VIC (SEQ ID NO 95) y SNP15-FAM (SEQ ID NO 96); para el SNP16 los cebadores SNP16-F (SEQ ID NO 64) y SNP16-R (SEQ ID NO 65) y las sondas SNP16-VIC (SEQ ID NO 97) y SNP16-FAM (SEQ ID NO 98); para el SNP17 los cebadores SNP17-F (SEQ ID NO 70) y SNP17-R (SEQ ID NO 71) y las sondas SNP17-VIC (SEQ ID NO 103) y SNP17-FAM (SEQ ID NO 104); para el SNP18 los cebadores SNP18-F (SEQ ID NO 68) y SNP18-R (SEQ ID NO 69) y las sondas SNP18-VIC (SEQ ID NO 101) y SNP18-FAM (SEQ ID NO 102).

TABLA 2.- Cebadores y sondas que permiten amplificar y genotipar los marcadores genéticos SNP. F indica el cebador *forward* y R indica el cebador *reverse*.

SNP	Cebadores	Sondas
1, 2	SNP1,2-F (SEQ ID NO 38)	SNP1,2-FAM (SEQ ID NO 72)
	SNP1,2-R (SEQ ID NO 39)	
3	SNP3-F (SEQ ID NO 40)	SNP3-VIC (SEQ ID NO 73)
	SNP3-R (SEQ ID NO 41)	SNP3-FAM (SEQ ID NO 74)
4	SNP4-F (SEQ ID NO 42)	SNP4-VIC (SEQ ID NO 75)
	SNP4-R (SEQ ID NO 43)	SNP4-FAM (SEQ ID NO 76)
5	SNP5-F (SEQ ID NO 50)	SNP5-VIC (SEQ ID NO 83)
	SNP5-R (SEQ ID NO 51)	SNP5-FAM (SEQ ID NO 84)
6	SNP6-F (SEQ ID NO 56)	SNP6-VIC (SEQ ID NO 89)
	SNP6-R (SEQ ID NO 57)	SNP6-FAM (SEQ ID NO 90)
7	SNP7-F (SEQ ID NO 60)	SNP7-VIC (SEQ ID NO 93)
	SNP7-R (SEQ ID NO 61)	SNP7-FAM (SEQ ID NO 94)
8	SNP8-F (SEQ ID NO 44)	SNP8-VIC (SEQ ID NO 77)
	SNP8-R (SEQ ID NO 45)	SNP8-FAM (SEQ ID NO 78)
9	SNP9-F (SEQ ID NO 46)	SNP9-VIC (SEQ ID NO 79)
	SNP9-R (SEQ ID NO 47)	SNP9-FAM (SEQ ID NO 80)
10	SNP10-F (SEQ ID NO 58)	SNP10-VIC (SEQ ID NO 91)
	SNP10-R (SEQ ID NO 59)	SNP10-FAM (SEQ ID NO 92)
11	SNP11-F (SEQ ID NO 48)	SNP11-VIC (SEQ ID NO 81)
	SNP11-R (SEQ ID NO 49)	SNP11-FAM (SEQ ID NO 82)
12	SNP12-F (SEQ ID NO 52)	SNP12-VIC (SEQ ID NO 85)
	SNP12-R (SEQ ID NO 53)	SNP12-FAM (SEQ ID NO 86)
13	SNP13-F (SEQ ID NO 54)	SNP13-VIC (SEQ ID NO 87)
	SNP13-R (SEQ ID NO 55)	SNP13-FAM (SEQ ID NO 88)
14	SNP14-F (SEQ ID NO 66)	SNP14-VIC (SEQ ID NO 99)
	SNP14-R (SEQ ID NO 67)	SNP14-FAM (SEQ ID NO 100)
15	SNP15-F (SEQ ID NO 62)	SNP15-VIC (SEQ ID NO 95)
	SNP15-R (SEQ ID NO 63)	SNP15-FAM (SEQ ID NO 96)
16	SNP16-F (SEQ ID NO 64)	SNP16-VIC (SEQ ID NO 97)

	SNP16-R (SEQ ID NO 65)	SNP16-FAM (SEQ ID NO 98)
17	SNP17-F (SEQ ID NO 70)	SNP17-VIC (SEQ ID NO 103)
	SNP17-R (SEQ ID NO 71)	SNP17-FAM (SEQ ID NO 104)
18	SNP18-F (SEQ ID NO 68)	SNP18-VIC (SEQ ID NO 101)
	SNP18-R (SEQ ID NO 69)	SNP18-FAM (SEQ ID NO 102)

El método de la invención se puede llevar a cabo partiendo de muestras de ADN obtenidas de pescado fresco, congelado, salazones, semiconservas, procesado o de una mezcla de los mismos.

Por último, también es objeto de la presente invención un kit para llevar a cabo el método de la invención que comprende el set de sondas SEQ ID NO 72 a SEQ ID NO 104. Adicionalmente, el kit puede comprender un set de cebadores SEQ ID NO 16 a SEQ ID NO 71 y/o *microarrays* con dichos cebadores y las sondas SEQ ID NO 72 a SEQ ID NO 104.

A continuación se describen una serie de ejemplos que ilustran la presente invención, pero no limitan su alcance en modo alguno.

EJEMPLOS

EJEMPLO 1.- Autentificar la anchoa bajo la denominación del "Cantábrico" en una semiconserva.

El objetivo de este ejemplo es el de identificar la muestra problema como una anchoa europea (*E. encrasicolus*) perteneciente a la población del Golfo de Vizcaya. Al tratarse de una muestra en semiconserva, y por lo tanto muy procesada, se debe utilizar una metodología que permita genotipar a partir de fragmentos de ADN muy pequeños (< 150 pb).

Etapa a) Extracción del ADN de la muestra de tejido en semiconserva

En primer lugar se elimina el exceso de aceite del filete de anchoa en semiconserva mediante secado con papel de filtro. Tras este paso se incuba a temperatura ambiente en agitación durante varias horas en agua destilada con el objetivo de eliminar el exceso de sal de la muestra. Durante la incubación se realizan varios cambios de agua destilada.

Para digerir el tejido se toman entre 0,2 y 0,3 g de tejido muscular en un tubo eppendorf de 1,5 ml. Al tubo se le añaden:

- 0,45 ml de tampón de extracción [SDS 1% (p/v), 150 mM NaCl, 2 mM EDTA, 10mM Tris-HCl pH 8,0], suplementado con 50 µl de una solución de tiocianato de guanidio 5 M.
- 40 µl de una solución de proteinasa k (600 U ml^{-1}) (Bioline).

La mezcla se incuba a 56ºC durante al menos 1 h. Si el tejido no se ha digerido completamente se puede dejar más tiempo. Tras el tiempo de incubación se elimina el material no digerido mediante centrifugación (Centrífuga refrigerada 5417 C/R, Eppendorf) a 13.000 rpm durante 5 min a temperatura ambiente. La centrifugación se repite tantas veces como sea necesario hasta conseguir un sobrenadante limpio de restos orgánicos en suspensión. El sobrenadante resultante se trata con un kit comercial de extracción de ADN (*Wizard-DNA Clean-Up*, Promega) siguiendo las especificaciones del fabricante, con la finalidad de extraer el ADN. El ADN se alícuota y almacena en agua estéril a -20ºC hasta su uso posterior.

El ADN extraído debe estar en cantidad y calidad suficientes para el análisis. La concentración de ADN de las muestras se calcula mediante espectrofotometría a una longitud de onda de 260 nm. En cuanto a la calidad del ADN, aunque se trate de muestras altamente procesadas en las cuales el ADN está muy degradado, esto no supone un problema ya que los tamaños de los fragmentos que se amplifican en el sistema *Real Time PCR* y *Taqman Open Array* presentan la suficiente longitud (<150 pb) como para encontrarse abundantemente incluso en las muestras más degradadas.

Etapa b) Identificación de la especie E. encrasicolus frente a otras anchoas del género Engraulis

La amplificación y genotipado del fragmento específico del gen mitocondrial citocromo b (CYTb-157, SEQ ID NO 1) que contiene los dos marcadores SNP1 y SNP2, se lleva a cabo en un volumen de reacción de 25 µl que contiene 2,25 µl de una solución 300 nM del cebador SNP1,2-F (SEQ ID NO 38), 2,25 µl de una solución 300 nM del cebador SNP1,2-R (SEQ ID NO 39), 2,5 µl de una solución 2,50 µM de la sonda SNP1,2-FAM (SEQ ID NO 72), 12,5 µl de la solución TaqMan™ Universal Master Mix (Applied Biosystems), 10 ng de ADN extraído de la muestra problema y finalmente agua destilada milliQ estéril para ajustar el volumen final de la reacción a 25 µl. De cada muestra se realizan duplicados. Además, en

el análisis se introduce un control negativo también por duplicado (*non template control*, ntc) que consiste en sustituir el ADN por agua, y controles positivos (duplicados) con 10 ng de ADN de la especie *E. encrasicolus*. El programa de amplificación en el equipo de *Real Time PCR* es como sigue: un ciclo inicial a 50ºC durante 2 min seguido de 10 min a 95ºC, para posteriormente hacer 40 ciclos de 95ºC durante 15 s y 60ºC durante 1 min. Al finalizar la reacción, sólo se detecta el amplificado en aquellos genotipos que tengan una G en la posición 468 (SNP1) y una T en la posición 474 (SNP2) del gen mitocondrial citocromo b. Esto se debe a que la sonda está diseñada para unirse específicamente a ese genotipo, que es exclusivo de *E. encrasicolus*. La detección se realiza observando los valores de Ct o la fluorescencia FAM de la sonda TaqMan, concluyéndose que si hay fluorescencia esa muestra pertenece a la especie *E. encrasicolus*.

En el supuesto de que no haya amplificación, la muestra problema se vuelve a analizar dopándola con 40 ng de ADN de la especie *E. encrasicolus* para evaluar el efecto negativo de inhibidores de la amplificación de PCR (control de inhibición). De este modo, si el resultado del control de inhibición es positivo indica que la muestra, efectivamente no es la especie *E. encrasicolus*. En caso contrario se repite la extracción de ADN para intentar obtener un ADN de mayor calidad y exento de inhibidores de la reacción de amplificación.

Etapa c) Identificación del origen geográfico o poblaciones de la anchoa europea

Para poder amplificar los 14 fragmentos del genoma y genotipar los 16 marcadores genéticos SNP3 a SNP18 (fases i) y iii) de la etapa c)) se utiliza la tecnología de genotipado *Taqman Open Array* (*Applied Biosystems*), implementada en la unidad de secuenciación y genotipado del servicio de genómica (SGIKER) de la Universidad del País Vasco (UPV/EHU). Los cebadores y sondas (FAM y VIC) específicos para la amplificación de cada marcador se encuentran incorporados en los *arrays* preparados por la casa comercial (*Applied Biosystems*). Se cargan las muestras en los *arrays* (22 ng/µl de cada muestra) añadiendo al menos dos controles negativos (agua en lugar de ADN) y dos controles positivos (dos muestras por duplicado), introduciendo el *array* cargado en un termociclador *GeneAmp PCR System* 9700 (*Applied Biosystems*). El protocolo de amplificación seguido es el indicado por la casa comercial y se encuentra especificado en la guía *TaqMan Open Array™ Genotyping System User Guide*. Una vez realizada la amplificación, el genotipado de las muestras se obtiene mediante el uso del *Software Autocaller v1.1* (*Applied Biosystems*). En el caso específico en el que el análisis se realiza usando varios SNP

mitocondriales (SNP 8, 9 y 11), la información de estos marcadores se usa de forma haplotípica.

Para la amplificación y genotipado de los distintos SNP realizados en estas fases i) y iii) del método de la invención se utilizan los siguientes cebadores y sondas (TABLA 2). La amplificación y genotipado del SNP3 se lleva a cabo con los cebadores SNP3-F (SEQ ID NO 40) y SNP3-R (SEQ ID NO 41) y las sondas SNP3-VIC (SEQ ID NO 73) y SNP3-FAM (SEQ ID NO 74). La del SNP4 con los cebadores SNP4-F (SEQ ID NO 42) y SNP4-R (SEQ ID NO 43) y las sondas SNP4-VIC (SEQ ID NO 75) y SNP4-FAM (SEQ ID NO 76). La del SNP5 con los cebadores SNP5-F (SEQ ID NO 50) y SNP5-R (SEQ ID NO 51) y las sondas SNP5-VIC (SEQ ID NO 83) y SNP5-FAM (SEQ ID NO 84). La del SNP6 con los cebadores SNP6-F (SEQ ID NO 56) y SNP6-R (SEQ ID NO 57) y las sondas SNP6-VIC (SEQ ID NO 89) y SNP6-FAM (SEQ ID NO 90). La del SNP7 con los cebadores SNP7-F (SEQ ID NO 60) y SNP7-R (SEQ ID NO 61) y las sondas SNP7-VIC (SEQ ID NO 93) y SNP7-FAM (SEQ ID NO 94). La del SNP8 con los cebadores SNP8-F (SEQ ID NO 44) y SNP8-R (SEQ ID NO 45) y las sondas SNP8-VIC (SEQ ID NO 77) y SNP8-FAM (SEQ ID NO 78). La del SNP9 con los cebadores SNP9-F (SEQ ID NO 46) y SNP9-R (SEQ ID NO 47) y las sondas SNP9-VIC (SEQ ID NO 79) y SNP9-FAM (SEQ ID NO 80). La del SNP10 con los cebadores SNP10-F (SEQ ID NO 58) y SNP10-R (SEQ ID NO 59) y las sondas SNP10-VIC (SEQ ID NO 91) y SNP10-FAM (SEQ ID NO 92). La del SNP11 con los cebadores SNP11-F (SEQ ID NO 48) y SNP11-R (SEQ ID NO 49) y las sondas SNP11-VIC (SEQ ID NO 81) y SNP11-FAM (SEQ ID NO 82). La del SNP12 con los cebadores SNP12-F (SEQ ID NO 52) y SNP12-R (SEQ ID NO 53) y las sondas SNP12-VIC (SEQ ID NO 85) y SNP12-FAM (SEQ ID NO 86). La del SNP13 con los cebadores SNP13-F (SEQ ID NO 54) y SNP13-R (SEQ ID NO 55) y las sondas SNP13-VIC (SEQ ID NO 87) y SNP13-FAM (SEQ ID NO 88). La del SNP14 con los cebadores SNP14-F (SEQ ID NO 66) y SNP14-R (SEQ ID NO 67) y las sondas SNP14-VIC (SEQ ID NO 99) y SNP14-FAM (SEQ ID NO 100). La del SNP15 con los cebadores SNP15-F (SEQ ID NO 62) y SNP15-R (SEQ ID NO 63) y las sondas SNP15-VIC (SEQ ID NO 95) y SNP15-FAM (SEQ ID NO 96). La del SNP16 con los cebadores SNP16-F (SEQ ID NO 64) y SNP16-R (SEQ ID NO 65) y las sondas SNP16-VIC (SEQ ID NO 97) y SNP16-FAM (SEQ ID NO 98). La del SNP17 con los cebadores SNP17-F (SEQ ID NO 70) y SNP17-R (SEQ ID NO 71) y las sondas SNP17-VIC (SEQ ID NO 103) y SNP17-FAM (SEQ ID NO 104). La del SNP18 con los cebadores SNP18-F (SEQ ID NO 68) y SNP18-R (SEQ ID NO 69) y las sondas SNP18-VIC (SEQ ID NO 101) y SNP18-FAM (SEQ ID NO 102).

A continuación se realiza un análisis de resultados mediante programas informáticos de asignación de individuos a una población (fases ii) y iv) de la etapa c)). Se utiliza el programa informático GENECLASS v2.0 que implementa un método bayesiano (Rannala y Mountain, 1997).

Para llevar a cabo la fase ii) se introduce el genotipo de la muestra problema obtenido para los 11 marcadores genéticos: SNP3 a SNP13, como "panel de asignación". Como "panel de referencia" se toma la población Atlántico Este frente a la población Vizcaya/Mediterráneo (100 individuos tomados al azar para cada panel de referencia) y se fija un *score* umbral de 95%. De este modo, si el genotipo problema se asigna a la población Atlántico Este con un *score* mayor o igual de 95 % se concluye que la muestra pertenece a dicha población. Si por el contrario el genotipo problema se asigna a la población Vizcaya/Mediterráneo con un *score* mayor o igual de 95 %, se concluye que la muestra pertenece a la población Vizcaya/Mediterráneo y se continúa el análisis como se describe a continuación para poder diferenciar entre población del Mediterráneo y del Golfo de Vizcaya. Los individuos que no superen el *score* umbral de 95% quedan sin ser asignados.

Para llevar a cabo la fase iv) se introduce el genotipo de la muestra problema obtenido para los siguientes 9 marcadores genéticos SNP7 y SNP11 a SNP18 como "panel de asignación". Como "panel de referencia" se toma la población Vizcaya frente a la población Mediterráneo (50 individuos tomados al azar para cada panel de referencia) y como *score* umbral 85%. De este modo, si el genotipo problema se asigna a la población de referencia Vizcaya con un *score* mayor o igual de 85 % se concluye que la muestra pertenece a dicha población. Si por el contrario el genotipo problema se asigna a la población Mediterráneo con un *score* mayor o igual de 85 %, se concluye que la muestra pertenece a la población Mediterráneo. Los individuos que no superen el *score* umbral de 85% quedan sin ser asignados.

Ejemplo 2: Autentificar la anchoa bajo la denominación del "Cantábrico" en una anchoa fresca

El objetivo de este ejemplo es el de identificar la muestra problema como una anchoa europea (*E. encrasicolus*) perteneciente a la población del Golfo de Vizcaya. Al tratarse de una muestra fresca, y por lo tanto no procesada, se puede utilizar la secuenciación como metodología de genotipado.

Etapa a) Extracción del ADN de la muestra de tejido

Para digerido el tejido se toman entre 0,2 y 0,3 g de tejido muscular en un tubo eppendorf de 1,5 ml. Al tubo se le añaden:

- 0,45 ml de tampón de extracción [SDS 1% (p/v), 150 mM NaCl, 2 mM EDTA, 10 mM Tris-HCl pH 8,0], suplementado con 50 µl de una solución de tiocianato de guanidio 5 M.
- 40 µl de una solución de proteinasa k (600 U ml^{-1}) (*Bioline*).

La mezcla se incuba a 56°C durante al menos 1 h. Si el tejido no se ha digerido completamente se puede dejar más tiempo. Tras el tiempo de incubación se elimina el material no disuelto mediante centrifugación (Centrífuga refrigerada 5417 C/R, Eppendorf) a 13.000 rpm durante 5 min a temperatura ambiente. La centrifugación se repite tantas veces como sea necesario hasta conseguir un sobrenadante limpio de restos orgánicos en suspensión. El sobrenadante resultante se trata con un kit comercial de extracción de ADN (*Wizard-DNA Clean-Up*, Promega) siguiendo las especificaciones del fabricante, con la finalidad de aislar el ADN. El ADN se alicuota y almacena en agua estéril a -20°C hasta su uso posterior.

El ADN extraído debe estar en cantidad y calidad suficientes para el análisis. La concentración de ADN de las muestras se calcula mediante espectrofotometría a una longitud de onda de 260 nm.

Etapa b) Identificación de la especie de anchoa europea

Se amplifica un fragmento del gen mitocondrial citocromo b de 307 pares de bases (pb), tal y como se describe en Kocher (1989). Antes de purificar el producto amplificado se realiza una comprobación de tamaño mediante electroforesis en gel de agarosa al 2%. Si el tamaño estimado es de 307 pb, el fragmento se purifica con un kit comercial *GFX-PCR-DNA and Gel Band Purification Kit* (Amersham Biosciences AB) siguiendo las especificaciones del fabricante, con el objetivo de eliminar restos de cebadores y enzimas que puedan afectar al proceso de secuenciación. El amplicón se eluye en agua bidestilada estéril.

La secuenciación se lleva a cabo directamente sobre el amplicón previamente purificado, en la unidad de secuenciación y genotipado del servicio de genómica (SGIKER) de la Universidad del País Vasco (UPV/EHU) con un *ABI PRISM 3130XL DNA Analyzer*

utilizando el kit *ABI Prism BigDye Terminator Cycle Sequencing Ready Reaction Kit version 3.0* (Applied Biosystems). Se envían a la unidad de secuenciación dos tubos con 10 µl del amplicón junto con cada uno de los cebadores L14841 (SEQ ID NO 109) y H15149 (SEQ ID NO 108) a 3.2 µM, con el objeto de poder secuenciar ambas hebras. La secuenciación se lleva a cabo a la temperatura de amplificación de los fragmentos, que es de 50ºC.

Específicamente se analizan los nucleótidos en las posiciones 468 y 474 del gen mitocondrial citocromo b. Si ambos nucleótidos se corresponden con los nucleótidos G y T, respectivamente, se concluye que la muestra problema es la especie *Engraulis encrasicolus* y se sigue con el análisis.

Etapa c) Identificación del origen geográfico o poblaciones de la anchoa europea

El siguiente paso es identificar la población de la muestra problema previamente identificada como anchoa europea. Para ello se amplifican 14 fragmentos del ADN en donde se sitúan los 16 marcadores genéticos del tipo SNP ((pasos i) y iii) de la fase c)) tal y como se detalla a continuación:

El fragmento clon-369 (SEQ ID NO 02) se amplifica con los cebadores clon369-F (SEQ ID NO 16) y clon369-R (SEQ ID NO 17).

El fragmento clon-286 (SEQ ID NO 03) se amplifica con los cebadores clon286-F (SEQ ID NO 18) y clon286-R (SEQ ID NO 19).

El fragmento clon-491 (SEQ ID NO 04) se amplifica con los cebadores clon491-F (SEQ ID NO 20) y clon491-R (SEQ ID NO 21).

El fragmento clon-M3EK (SEQ ID NO 05) se amplifica con los cebadores clonM3EK-F (SEQ ID NO 22) y clonM3EK-R (SEQ ID NO 23).

El fragmento clon-289 (SEQ ID NO 06) se amplifica con los cebadores clon289-F (SEQ ID NO 24) y clon289-R (SEQ ID NO 25).

El fragmento clon-470 (SEQ ID NO 12) se amplifica con los cebadores clon470-F (SEQ ID NO 26) y clon470-R (SEQ ID NO 27).

El fragmento clon-357 (SEQ ID NO 13) se amplifica con los cebadores clon357-F (SEQ ID NO 28) y clon357-R (SEQ ID NO 29).

El fragmento clon-M4 (SEQ ID NO 14) se amplifica con los cebadores clonM4-F (SEQ ID NO 30) y clonM4-R (SEQ ID NO 31).

El fragmento clon-483 (SEQ ID NO 15) se amplifica con los cebadores clon483-F (SEQ ID NO 32) y clon483-R (SEQ ID NO 33).

En todos estos casos, la reacción de amplificación se lleva a cabo como se describe a continuación: 0,8 mM dNTPs, 2,5 mM $MgCl_2$ y 0,7 µM de cada cebador y 25-50 ng de ADN molde en 1X PCR Buffer II (Applied Biosystems). La reacción comienza añadiendo a la mezcla de reacción 1.5 U del enzima *AmpliTaq Gold* (Applied Biosystems). La reacción comienza con un ciclo de activación del enzima a 95°C durante 10 min, seguido de 20 ciclos [30 s a 95 °C, 45 s a 64 °C (-0,5 °C por ciclo) y 30 s a 72 °C], de 22 ciclos [30 s a 95 °C, 30 s a 48 °C y 30 s a 72 °C] y finalmente un ciclo final de extensión del fragmento amplificado durante 10 min a 72 °C.

El fragmento MT-12S (SEQ ID NO 07) se amplifica con los cebadores MT-12S-F (SEQ ID NO 34) y MT-12S-R (SEQ ID NO 35). La reacción se lleva a cabo como se describe a continuación: 1 mM dNTPs, 2,5 mM $MgCl_2$ y 1 µM de cada cebador y 25-50 ng de ADN molde en 1X *GoTaq Flexi Buffer* (Promega). La reacción comienza añadiendo a la mezcla de reacción 1,5 U del enzima *GoTaq* (Promega). La reacción comienza con un ciclo de activación del enzima a 95°C durante 3 min, seguido de 35 ciclos [60 s a 95 °C, 30 s a 54 °C y 30 s a 72 °C] y un ciclo final de extensión del fragmento amplificado durante 5 min a 72 °C.

El fragmento CATL-7 (SEQ ID NO 08) se amplifica con los cebadores CATL-7-F (SEQ ID NO 36) y CATL-7-R (SEQ ID NO 37). La reacción se lleva a cabo como se describe a continuación: 0,8 mM dNTPs, 2,5 mM $MgCl_2$ y 1 µM de cada cebador y 25-50 ng de ADN molde en 1X PCR Buffer II (Applied Biosystems). La reacción comienza añadiendo a la mezcla de reacción 1,5 U del enzima AmpliTaq Gold (Applied Biosystems). La reacción comienza con un ciclo de activación del enzima a 95°C durante 10 min, seguido de 35 ciclos [60 s a 95 °C, 60 s a 60 °C y 80 s a 72 °C] y un ciclo final de extensión del fragmento amplificado durante 10 min a 72 °C.

El fragmento CYTb (SEQ ID NO 09) se amplifica con los cebadores CytBI-7F (SEQ ID NO 105) y Trucytb-R (SEQ ID NO 106) previamente descritos en Jérôme *et al.* 2003b y siguiendo las indicaciones de la referencia Jérôme *et al.* 2008.

Por último, el genotipado de los SNP de los genes Rodopsina (SEQ ID NO 14) y CK6-1 (SEQ ID NO 15) se realiza con los cebadores SNP12-F (SEQ ID NO 52) y SNP-12-R (SEQ ID NO 53), y SNP13-F (SEQ ID NO 54) y SNP13-R (SEQ ID NO 55), respectivamente, siguiendo las condiciones de amplificación especificadas en la guía *TaqMan OpenArray™ Genotyping System User Guide*.

La comprobación del tamaño de los fragmentos amplificados se realiza mediante electroforesis en gel de agarosa. Una vez comprobado su tamaño se procede a purificarlos con un kit comercial *GFX-PCR-DNA and Gel Band Purification Kit* (Amersham Biosciences AB) siguiendo las especificaciones del fabricante, con el objetivo de eliminar restos de cebadores y enzimas que puedan afectar al proceso de secuenciación. El amplicón se eluye en agua bidestilada estéril y la secuenciación se lleva a cabo directamente sobre el amplicón previamente purificado, en la unidad de secuenciación y genotipado del servicio de genómica (SGIKER) de la Universidad del País Vasco (UPV/EHU) con un *ABI PRISM 3130XL DNA Analyzer* utilizando el kit *ABI Prism BigDye Terminator Cycle Sequencing Ready Reaction Kit version 3.0* (Applied Biosystems). Se envían a la unidad de secuenciación dos tubos con 10 µl del amplicón junto con cada uno de los cebadores a 3.2 µM con el objeto de poder secuenciar ambas hebras. La secuenciación se lleva a cabo a la temperatura de *annealing* especificada para cada pareja de cebadores. En cada una de las secuencias obtenidas se interrogan los nucleótidos correspondientes a los 16 SNP de la invención: SNP3-SNP18.

A continuación se realiza el análisis de resultados mediante programas informáticos de asignación de individuos a una población (fases ii) y iv) de la etapa c)). Para ello se utiliza el programa informático GENECLASS v2.0 que implementa un método bayesiano (Rannala y Mountain, 1997).

En la fase ii) el *score* umbral es 95% y como “panel de referencia” se toma el panel de referencia de la población Atlántico Este frente al resto (Vizcaya/Mediterráneo). Se introduce el genotipo obtenido con los siguientes 11 marcadores genéticos: SNP3 a SNP13, como “panel de asignación”. Si el genotipo problema se asigna a la población del Atlántico Este, con un *score* mayor o igual del 95 %, se concluye que la muestra no pertenece al Golfo de Vizcaya. En cambio, si el genotipo problema se asigna a la población Vizcaya/Mediterráneo de referencia con un *score* mayor o igual del 95 %, se continúa el análisis como se describe a continuación para poder diferenciar entre población del Mediterráneo y del Golfo de Vizcaya. Los individuos que no superen el *score* umbral de 95% quedan sin ser asignados.

En la fase iv) de la etapa c) se fija un valor de *score* umbral del 85% y como “panel de referencia” se toma el panel de referencia de la población Vizcaya frente a la población Mediterráneo. Se introduce el genotipo de los siguientes 9 marcadores genéticos: SNP7 y SNP11 a SNP18 como “panel de asignación”. De este modo, si el genotipo problema se

asigna a la población Vizcaya, con un *score* mayor o igual del 85 %, se concluye que la muestra pertenece a la población del Golfo de Vizcaya. Si el genotipo problema se asigna a la población Mediterráneo con un *score* mayor o igual del 85% se concluye que pertenece a la población del Mediterráneo. Los individuos que no superen el *score* umbral de 85% quedan sin ser asignados.

BIBLIOGRAFÍA

Banks, M.A. y Eichert, W. WHICHRUN (version 3.2): a computer program for population assignment of individuals based on multilocus genotype data. *Journal of Heredity*, 2000, 91:87-89

Bembo, D.G., Carvalho, G.R., Cingolani, N., Arneri E., Giannetti G., Pitcher, T.J. Allozymic and morphometric evidence for twostocks of the European anchovy *Engraulis encrasicolus* in Adriatic waters. *Marine Biology,* 1996a, 126: 529-538

Bembo, D.G., Carvalho, G.R., Cingolani, N., Pitcher, T.J. Electrophoretic analysis of stock structure in northern Mediterranean anchovies, *Engraulis encrasicolus. Journal of Marine Science*, 1996b, 53: 115-128

Bembo, D.G., Carvalho, Snow, M., Cingolani, N., Pitcher, T.J. Stock discrimination among European anchovies, *Engraulis encrasicolus,* by means of PCR-amplified mitochondrial DNA analysis. *Fishery Bulletin,* 1995, 94: 31-40

Bernatchez, L. y Duchesne, P. Individual-based genotype analysis in studies of parentage and population assignment: how many loci, how many alleles?. *Canadian Journal of Fisheries and Aquatic Science*, 2000, 57: 1–12.

Borsa, P. Allozyme, mitichondrial-DNA, and morphometric variability indicate cryptic species of anchovy (*Engraulis encrasicolus). Biological Journal of the Linnean Society,* 2002, 75: 261-269

Borsa, P., Collet, A., Durand, J.D. Nuclear-DNA markers confirm the presence of two anchovy species in the Mediterranean. *C. R. Biologies*, 2004, 327: 1113-1123

Chapela M.J., Sotelo C.G., Pérez-Martín R.I., Pardo M.A., Pérez-Villareal B., Gilardi P. y Riese J. Comparison of DNA extraction methods from muscle of canned tuna for species identification. *Food Control*, 2007, 18: 1211-1215.

Chiu, T.-S.; Lee, Y.-J.; Huang, S.W.; Yu, H.-T. Polymorphic microsatellite markers for stock identification in Japanese anchovy (*Engraulis japonica*). *Molecular Ecology Notes*, 2002, 2: 49-50.

Corander, J., Waldmann, P., y Sillanpää, M. J. Bayesian analysis of genetic differentiation between populations. *Genetics*, 2003, 163, 367-374.

Cornuet, J.M., Piry, S., Luikart, G., Estoup, A. y Solignac, M. Comparison of methods employing multilocus genotypes to select or exclude populations as origins of individuals. *Genetics*, 1999, 153, 1989–2000.

Glaubitz, J.C., Rhodes, E., DeWoody, A. Prospects for inferring pairwise relationships with single nucleotide polymorphisms. *Mol. Ecol.,* 2003, 12: 1039-1047

Grant, W.S. A second look at mitochondrial DNA variability in European anchovy (*Engraulis encrasicolus)*: assesing models of population structure and the Black Sea isolation hypothesis. *Genetica*, 2005, 125: 293-309

Hauser, L., Seamons, T. R., Dauer, M., Naish, K. A., y Quinn,T. P. An empirical verification of population assignment methods by marking and parentage data: hatchery and wild steelhead (*Oncorhynchus mykiss*). *Molecular Ecology*, 2006, 15, 3157–3173

Jérôme, M., Lemaire, C., Verrez-Bagnis, V. y Etienne, M. Direct sequencing method for species identification of canned sardine and sardine-type products. *J. Agric. Food Chem.* 2003a, 51, 7326-7332

Jérôme, M.; Lemaire, C.; Bautista, J. M.; Fleurence, J.; Etienne, M. Molecular phylogeny and species identification of sardines. J. Agric. Food Chem. 2003b, 51, 43–50.

Jérôme, M., Martinsohn JM., Ortega D., Carreau P., Verrez-Bagnis, V. y Mouchel O. Toward Fish and Seafood Traceability: Anchovy Species Determination in Fish Products by Molecular Markers and Support through a Public Domain Database. *J. Agric. Food Chem.* 2008, 56, 3460-3469.

Kalinowski, ST. Genetic Mixture Analysis 1.0. Department of Ecology, Montana State University, Bozeman, 2003, MT 59717.

Kocher, T.D., et al., Dynamics of mitochondrial DNA evolution in animals: Amplification and sequencing with conserved primers. *Proc Natl Acad Sci U S A*, 1989. 86: p. 6196-6200.

Landegren, U., Nilsson, M., Kwok, P.Y. Reading bits of genetic information: methods for single-nucleotide polymorphism analysis. *Genome Res.,* 1998, 8: 769-776

Landi, M., Garoia, F., Piccinetti, C., Tinti, F. Isolation of polymorphic microsatellite loci for the European anchovy, Engraulis encrasicolus. *Molecular ecology notes*, 2005, 5: 266-268.

Liu, Z.J., Cordes, J.F. DNA marker technologies and their applications in aquaculture genetic. *Aquaculture,* 2004, 238: 1-7

Magoulas, A., Castilho, R., Caetano, S., Marcato, S., y Patarnello, T. 2006. Mitochondrial DNA reveals a mosaic pattern of phylogeographical structure in Atlantic and Mediterranean populations of anchovy (Engraulis encrasicolus). *Molecular Phylogenetics and Evolution*, 39: 734–746

Magoulas, A., Tsimenides, N., Zouros, E. Mitochondrial DNA phylogeny and the reconstruction of the population history of a species: the case of the European anchovy (*Engraulis encrasicolus). Mol. Biol. Evol.*, 1996, 13(1): 178-190

Nielsen, R. Estimation of population parameters and recombination rates from single nucleotide polymorphism. *Genetics,* 2000, 154: 931-942

Paetkau, D., Calvert, W., Stirling, I. y Strobeck, C. Microsatellite analysis of population structure in Canadian polar bears. *Molecular Ecology*, 1995, 4: 347–354.

Pella, J., y Masuda, M. Bayesian methods for analysis of stock mixtures from genetic characters. *Fishery Bulletin*, 2001, 99: 151–167.

Pritchard, J. K., Stephens, M. y Donnelly, P. Inference of population structure using multilocus genotype data. *Genetic*, 2000, 155: 945–959.

Rannala, B. y Mountain, J.L. Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences USA*, 1997, 94: 9197–9201.

Ratnasingham, S. & Hebert, P. D. N. BOLD: The Barcode of Life Data System (www.barcodinglife.org). Molecular Ecology Notes, 2007, 7: 355–364

Rengmark, A.H., Slettan, A., Skaala, O., Lie, O., Lingaas, F. Genetic variability in wild and farmed Atlantic salmon (*Salmo salar)* strains estimated by SNP and microsatellites. *Aquaculture*, 2006, 253: 229-237

Römpler, H., Rohland, N., Lalueza- Fox, C., Willerslev, E., Kuznetsova, T., Rabeder, G., Bertranpetit, J., Schöneberg, T., y Hofreiter, M. Nuclear gene indicates coat-color polymorphism in mammoths. Science, 2006, 313-62

Schneider, S., Kueffer, J.-M., Roessli, D. y Excoffier, L. Arlequin, Version 2.0. Genetics and Biometry Lab, Dept. of Anthropology, University of Geneva, Switzerland. 2000

Sebastio, P, Zanelli, P. y Tauro, M. N. Identification of anchovy (*Engraulis encrasicolus* 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

Spanakis, E., Tsimenides, N., Zouros, E. Genetic differences between populations of sardine, *Sardina pilchardus*, and anchovy, *Engraulis encrasicolus,* in the Aegean and Ionian seas. *J. Fish. Biol.*, 1989, 35: 417-437

Siegismund, H. R. G-STAT, Version 3.1. Genetical statistical programs for the analysis of population data. Botanical Institute, 1995, University of Copenhagen, Denmark.

Tudela, S., Gracia-Marin, J.L., Pla, C. Genetic structure of the European anchovy, *Engraulis encrasicolus*, in the north-west Mediterranean. *Journal of Experimental Marine Biology and Ecology,* 1999, 234: 95-109

Yu, H.T., Lee, Y.J., Huang, S.W., Chiu, T.S Genetic analysis of the populations of japanese anchovy (Engraulidae: *Engraulis japonicus*) using microsatellite DNA. *Marine Biotechnology*, 2002, 4: 471-479

Yu, Z. N., Kong, X. Y., Guo, T. H., Jiang, Y. Y., Zhuang, Z. M., y Jin, X. S. Mitochondrial DNA sequence variation of Japanese anchovy Engraulis japonicus from the Yellow Sea and East China Sea. Fisheries Science, 2005, 71: 299–307.

Zarraonaindia, I., Pardo, M. A., Iriondo, M., Manzano, C., y Estonba, A. Microsatellite variability in European anchovy (*Engraulis encrasicolus*) calls for further investigation of its genetic structure and biogeography. – ICES Journal of Marine Science, 2009, 66: 2176-2182.

REIVINDICACIONES

1) Conjunto de las secuencias de ADN SEQ ID NO 1 a SEQ ID NO 15 caracterizadas por que comprenden los siguientes SNP: SNP1 y SNP2 (SEQ ID NO 1), SNP3 (SEQ ID NO 2), SNP4 y SNP14 (SEQ ID NO 3), SNP5 (SEQ ID NO 4), SNP6 (SEQ ID NO 5), SNP7 (SEQ ID NO 6), SNP15 (SEQ ID NO 7), SNP16 (SEQ ID NO 8), SNP17 (SEQ ID NO 9), SNP18 (SEQ ID NO 10), SNP8 y SNP9 (SEQ ID NO 11), SNP10 (SEQ ID NO 12), SNP11 (SEQ ID NO 13), SNP12 (SEQ ID NO 14) y SNP13 (SEQ ID NO 15).

2) Uso de las secuencias de acuerdo a la reivindicación 1 y de los SNP incluidos en ellas como marcadores para identificar la especie *Engraulis encrasicolus* y su origen geográfico o población.

3) Método de identificación genética del origen geográfico y poblaciones de la anchoa europea (*Engraulis encrasicolus*) que comprende:

a) Extracción del ADN de la muestra de un producto derivado de la anchoa
b) Identificación de la especie *Engraulis encrasicolus* frente al resto de especies de anchoas pertenecientes al género *Engraulis* que comprende la amplificación y genotipado de un fragmento que contenga el SNP1 y el SNP2 (SEQ ID NO 01), donde la presencia en las posiciones 468 y 474 del gen mitocondrial citocromo b de una G y una T, respectivamente, es indicativa de la especie *E. encrasicolus*
c) Identificación del origen geográfico o poblaciones de la anchoa europea (*E. encrasicolus*) que comprende:
    i. amplificación y genotipado de los siguientes 11 SNP según la reivindicación 1: SNP3 a SNP13,
    ii. análisis de los resultados obtenidos en la fase i) y asignación del genotipo problema con una población de referencia: población Atlántico Este o población Vizcaya/Mediterráneo
    iii. amplificación y genotipado de los siguientes 9 SNP según la reivindicación 1: SNP7 y SNP11 a SNP18,
    iv. análisis de los resultados obtenidos en la fase iii) y asignación de la muestra problema con una población de referencia: población Mediterráneo o población Vizcaya.

4) Método según la reivindicación 3, caracterizado por que el genotipado de los SNP1 y SNP2 de la etapa b) del método de la invención se lleva a cabo con la sonda SNP1,2-FAM (SEQ ID NO 72).

5) Método según las reivindicaciones 3 ó 4, caracterizado por que la amplificación del SNP1 y del SNP2 de la etapa b) del método de la invención se lleva a cabo con los oligonucleótidos SNP1,2-F (SEQ ID NO 38) y SNP1,2-R (SEQ ID NO 39).

6) Método según cualquiera de las reivindicaciones 3 a 5, donde el genotipado de los SNP3 a SNP13 de la fase i) de la etapa c) se lleva a cabo con las siguientes sondas; SNP3 con las sondas SNP3-VIC (SEQ ID NO 73) y SNP3-FAM (SEQ ID NO 74); SNP4 con las sondas SNP4-VIC (SEQ ID NO 75) y SNP4-FAM (SEQ ID NO 76); SNP5 con las sondas SNP5-VIC (SEQ ID NO 83) y SNP5-FAM (SEQ ID NO 84); SNP6 con las sondas SNP6-VIC (SEQ ID NO 89) y SNP6-FAM (SEQ ID NO 90); SNP7 con las sondas SNP7-VIC (SEQ ID NO 93) y SNP7-FAM (SEQ ID NO 94); SNP8 con las sondas SNP8-VIC (SEQ ID NO 77) y SNP8-FAM (SEQ ID NO 78); SNP9 con las sondas SNP9-VIC (SEQ ID NO 79) y SNP9-FAM (SEQ ID NO 80); SNP10 con las sondas SNP10-VIC (SEQ ID NO 91) y SNP10-FAM (SEQ ID NO 92); SNP11 con las sondas SNP11-VIC (SEQ ID NO 81) y SNP11-FAM (SEQ ID NO 82); SNP12 con las sondas SNP12-VIC (SEQ ID NO 85) y SNP12-FAM (SEQ ID NO 86); SNP13 con las sondas SNP13-VIC (SEQ ID NO 87) y SNP13-FAM (SEQ ID NO 88).

7) Método según cualquiera de las reivindicaciones 3 a 6, donde el genotipado de los SNP7 y SNP11 a SNP18 de la fase iii) de la etapa c) se lleva a cabo con las siguientes sondas: SNP7 con las sondas SNP7-VIC (SEQ ID NO 93) y SNP7-FAM (SEQ ID NO 94); SNP11 con las sondas SNP11-VIC (SEQ ID NO 81) y SNP11-FAM (SEQ ID NO 82); SNP12 con las sondas SNP12-VIC (SEQ ID NO 85) y SNP12-FAM (SEQ ID NO 86); SNP13 con las sondas SNP13-VIC (SEQ ID NO 87) y SNP13-FAM (SEQ ID NO 88); SNP14 con las sondas SNP14-VIC (SEQ ID NO 99) y SNP14-FAM (SEQ ID NO 100); SNP15 con las sondas SNP15-VIC (SEQ ID NO 95) y SNP15-FAM (SEQ ID NO 96); SNP16 con las sondas SNP16-VIC (SEQ ID NO 97) y SNP16-FAM (SEQ ID NO 98); SNP17 con las sondas SNP17-VIC (SEQ ID NO 103) y SNP17-FAM (SEQ ID NO 104); SNP18 con las sondas SNP18-VIC (SEQ ID NO 101) y SNP18-FAM (SEQ ID NO 102).

8) Método según cualquiera de las reivindicaciones 3 a 7, caracterizado por que los oligonucleótidos utilizados en las fases i) y iii) de la etapa c) del método de la invención son los siguientes: SNP3-F (SEQ ID NO 40) y SNP3-R (SEQ ID NO 41) para el SNP3; SNP4-F (SEQ ID NO 42) y SNP4-R (SEQ ID NO 43) para el SNP4; SNP5-F (SEQ ID NO 50) y SNP5-R (SEQ ID NO 51) para el SNP5; SNP6-F (SEQ ID NO 56) y SNP6-R (SEQ ID NO 57) para el SNP6; SNP7-F (SEQ ID NO 60) y SNP7-R (SEQ ID NO 61) para el SNP7; SNP8-F (SEQ ID NO 44) y SNP8-R (SEQ ID NO 45) para el SNP8; SNP9-F (SEQ ID NO 46) y SNP9-R (SEQ ID NO 47) para el SNP9; SNP10-F (SEQ ID NO 58) y SNP10-R (SEQ ID NO 59) para el SNP10; SNP11-F (SEQ ID NO 48) y SNP11-R (SEQ ID NO 49) para el SNP11; SNP12-F (SEQ ID NO 52) y SNP12-R (SEQ ID NO 53) para el SNP12; SNP13-F (SEQ ID NO 54) y SNP13-R (SEQ ID NO 55) para el SNP13; SNP14-F (SEQ ID NO 66) y SNP14-R (SEQ ID NO 67) para el SNP14; SNP15-F (SEQ ID NO 62) y SNP15-R (SEQ ID NO 63) para el SNP15; SNP16-F (SEQ ID NO 64) y SNP16-R (SEQ ID NO 65) para el SNP16; SNP17-F (SEQ ID NO 70) y SNP17-R (SEQ ID NO 71) para el SNP17; SNP18-F (SEQ ID NO 68) y SNP18-R (SEQ ID NO 69) para el SNP18.

9) Método según cualquiera de las reivindicaciones 3 a 8, caracterizado por que los oligonucleótidos y las sondas utilizados en las etapas b) y c) del método de la invención son los siguientes: SNP1,2-F (SEQ ID NO 38), SNP1,2-R (SEQ ID NO 39) y SNP1,2-FAM (SEQ ID NO 72) para los SNP1 y SNP2; SNP3-F (SEQ ID NO 40), SNP3-R (SEQ ID NO 41), SNP3-VIC (SEQ ID NO 73) y SNP3-FAM (SEQ ID NO 74) para el SNP3; SNP4-F (SEQ ID NO 42), SNP4-R (SEQ ID NO 43), SNP4-VIC (SEQ ID NO 75) y SNP4-FAM (SEQ ID NO 76) para el SNP4; SNP5-F (SEQ ID NO 50), SNP5-R (SEQ ID NO 51), SNP5-VIC (SEQ ID NO 83) y SNP5-FAM (SEQ ID NO 84) para el SNP5; SNP6-F (SEQ ID NO 56), SNP6-R (SEQ ID NO 57), SNP6-VIC (SEQ ID NO 89) y SNP6-FAM (SEQ ID NO 90) para el SNP6; SNP7-F (SEQ ID NO 60), SNP7-R (SEQ ID NO 61), SNP7-VIC (SEQ ID NO 93) y SNP7-FAM (SEQ ID NO 94) para el SNP7; SNP8-F (SEQ ID NO 44), SNP8-R (SEQ ID NO 45), SNP8-VIC (SEQ ID NO 77) y SNP8-FAM (SEQ ID NO 78) para el SNP8; SNP9-F (SEQ ID NO 46), SNP9-R (SEQ ID NO 47), SNP9-VIC (SEQ ID NO 79) y SNP9-FAM (SEQ ID NO 80) para el SNP9; SNP10-F (SEQ ID NO 58), SNP10-R (SEQ ID NO 59), SNP10-VIC (SEQ ID NO 91) y SNP10-FAM (SEQ ID NO 92) para el SNP10; SNP11-F (SEQ ID NO 48), SNP11-R (SEQ ID NO 49), SNP11-VIC (SEQ ID NO 81) y SNP11-FAM (SEQ ID NO 82) para el SNP11; SNP12-F (SEQ ID NO 52), SNP12-R (SEQ ID NO 53), SNP12-VIC (SEQ ID NO 85) y SNP12-FAM (SEQ ID NO 86) para el SNP12; SNP13-F (SEQ ID NO 54), SNP13-R (SEQ ID NO 55), SNP13-VIC (SEQ ID NO 87) y SNP13-FAM (SEQ ID NO 88) para el

SNP13; SNP14-F (SEQ ID NO 66), SNP14-R (SEQ ID NO 67), SNP14-VIC (SEQ ID NO 99) y SNP14-FAM (SEQ ID NO 100) para el SNP14; SNP15-F (SEQ ID NO 62), SNP15-R (SEQ ID NO 63), SNP15-VIC (SEQ ID NO 95) y SNP15-FAM (SEQ ID NO 96) para el SNP15; SNP16-F (SEQ ID NO 64), SNP16-R (SEQ ID NO 65), SNP16-VIC (SEQ ID NO 97) y SNP16-FAM (SEQ ID NO 98) para el SNP16; SNP17-F (SEQ ID NO 70), SNP17-R (SEQ ID NO 71), SNP17-VIC (SEQ ID NO 103) y SNP17-FAM (SEQ ID NO 104) para el SNP17; SNP18-F (SEQ ID NO 68), SNP18-R (SEQ ID NO 69), SNP18-VIC (SEQ ID NO 101) y SNP18-FAM (SEQ ID NO 102) para el SNP18.

10) Método según cualquiera de las reivindicaciones 3 a 9 caracterizado por que la fase ii) de la etapa c) se lleva a cabo eligiendo un *score* umbral de 70% y donde la asignación del genotipo problema con una de las poblaciones de referencia Atlántico Este o Mediterráneo/Vizcaya con un *score* mayor o igual a 70% es indicativa de pertenencia a dicha población.

11) Método según cualquiera de las reivindicaciones 3 a 9 caracterizado por que la fase ii) de la etapa c) se lleva a cabo eligiendo un *score* umbral de 85% y donde la asignación del genotipo problema con una de las poblaciones de referencia Atlántico Este o Mediterráneo/Vizcaya con un *score* mayor o igual a 85% es indicativa de pertenencia a dicha población.

12) Método según cualquiera de las reivindicaciones 3 a 9 caracterizado por que la fase ii) de la etapa c) se lleva a cabo eligiendo un *score* umbral de 95% y donde la asignación del genotipo problema con una de las poblaciones de referencia Atlántico Este o Mediterráneo/Vizcaya con un *score* mayor o igual a 95% es indicativa de pertenencia a dicha población.

13) Método según cualquiera de las reivindicaciones 3 a 12 caracterizado por que la fase iv) de la etapa c) se lleva a cabo eligiendo un *score* umbral de 70% y donde la asignación del genotipo problema con una de las poblaciones de referencia Mediterráneo o Vizcaya con un *score* mayor o igual a 70% es indicativa de pertenencia a dicha población.

14) Método según cualquiera de las reivindicaciones 3 a 12 caracterizado por que la fase iv) de la etapa c) se lleva a cabo eligiendo un *score* umbral de 80% y donde la asignación del genotipo problema con una de las poblaciones de referencia Mediterráneo o Vizcaya con un *score* mayor o igual a 80% es indicativa de pertenencia a dicha población.

15) Método según cualquiera de las reivindicaciones 3 a 12 caracterizado por que la fase iv) de la etapa c) se lleva a cabo eligiendo un *score* umbral de 85% y donde la asignación del genotipo problema con una de las poblaciones de referencia Mediterráneo o Vizcaya con un *score* mayor o igual a 85% es indicativa de pertenencia a dicha población.

16) Método según cualquiera de las reivindicaciones 3 a 15 caracterizado por que la muestra de ADN se obtiene de pescado fresco, congelado, salazones, semiconservas, procesado o de una mezcla de los mismos.

17) Kit para llevar a cabo el método según cualquiera de las reivindicaciones 3 a 16 caracterizado por que comprende el set de sondas SEQ ID NO 72 a SEQ ID NO 104.

18) Kit según la reivindicación 17 caracterizado por que además comprende un set de cebadores SEQ ID NO 16 a SEQ ID NO 71 y/o *microarrays* con estos cebadores y las sondas SEQ ID NO 72 a SEQ ID NO 104.

Listado de Secuencias

<110> Universidad del País Vasco/Euskal Herriko Unibertsitatea
Fundación AZTI/AZTI Fundazioa

<120> Método de identificación genética de la anchoa europea y de su origen geográfico o población.

<130> 305/10

<160> 108

<170> PatentIn version 3.5

<210> 1
<211> 161
<212> DNA
<213> Engraulis sp.

<400> 1
gccttcgttg ggtacgtact accctgagga caaatgtcct tctgaggagc aactgtcatt 60

actaacctta tgtctgcagt tccttacgtg ggcactgaat trgtscartg aatttgaggr 120

ggcttttcgg tcgacaacgc gactcttact cggttcttcg c 161

<210> 2
<211> 478
<212> DNA
<213> Engraulis sp.

<400> 2
tggaaagaga aggtggcatc ttttttgca cacaaatcat cacaagccac tgctaatcag 60

cattcataaa tcagattcaa cctagtctaa tttgtctcat caaagacact taactgtcaa 120

gtcaaactat atcagaatat gcagttagaa tgtacaacac ttgctcaaca ggtaactctt 180

gcccatcaat cttgcagtct gtttggcatg gctgtcgttm accatttccc atgtagccta 240

ccattacatg cttggcatgt gatcatttcc attagccttg aggctggaca gtttacattt 300

cagacttcac ttatgatttg atttgtacac aaagcagcac atttgagcgt tcaaaatgtg 360

cttagcccaa taccacggcc ttgactggat ataaaaacga attagcgtcc aagagattta 420

gggagatggg catgaaaaat tattctccag ggagtgagaa aaaaaacgag cgagttgg 478

<210> 3
<211> 680
<212> DNA
<213> Engraulis sp.

<400> 3
gtgtgtttct ttttttaaca cgcaggccat tgttagaagt taatgcttga aagcagcaag 60

gcaatttagc acgctttttt gctacgcttt gctttgcttt gattgaaatt tgcttgatgg 120

cttttcgcgg cttttgaatt gctacaacac acgctggcat gcgacgcttc acaagtgtgt 180

gtatttttt tttttaaacc atgtgcatgc ttggctgtgc attggtgagt ggtgagtatg 240

ttggatgaga gtgagtatgt cgatcgtctc tttttctctc tgaacaacat taatgctgac 300

acggcgaagg tggtggtgaa tgtcttgtat gtcccccagg ttctggggcc cacccgctgg 360

tcgacccgct ggccgccgga ccgcacctgg cccgattccc ctacccgccc ggagccatcc 420

ccaacgccct gctgggacag ccccacccyg agcatgagat gctgcggcac cctgtgttcg 480

gtaagcaccc accggactgg acccacacat ctcactgtac aggaggacac sttttagata 540

ttacaaatgt tgtgttgggt catggtctcc attggaccaa tacacagcac taatcaccca 600

cacttcgatt cacaggacac actttacacc tttcaaatgc tgcactccat cataatctct 660

gactacaaac tttgacacat 680

<210> 4
<211> 538
<212> DNA
<213> Engraulis sp.

<400> 4
agcatgtgtg ttttccccca tattttaatg ttttgccacc catcagaagt acagtatttc 60

cattcccaaa attgtgtgat aggaaaataa ttgacagttc tatcactatt tgtccttttt 120

tcccatcagg agaattcagc tatgggctgg tggtctggtc tggtgattgt catcttttca 180

gtcccgttta gtgatctaca gttcgrgggc ataaggccat cgcagcaata tgagccctca 240

tcatcctgtg ctgtctcctt taaatcccct cactctccat acatcagaaa gaaggactac 300

tggaggctta caggctgatt gaccaatggg aagcttacat agaaaactga aaaaactaaa 360

tcatccatcc aggtcaccgt agtcaagact agtggaaagt tagtggaagt tagttagata 420

gttaaagaag ttaaggttat cttccggaga acattgcatg gagacccagc acacagcctc 480

tctatccgcc tttgactgtg aggcaatggt aagtcatcat agtagagagt agctggga 538

<210> 5
<211> 606
<212> DNA
<213> Engraulis sp.

<400> 5
gcaatcgacc agacgtgatg tggcgaggat gttatacata cagtatgatg cgtgtggtct 60

gagctctcag ctggttcaca agtggttgag ctacagtgta gcttgcatca ggttgagcat 120

gcaggtggca ctgatgacaa gaatgaatgt tggcaaacac tcactrgcct tagaattgca 180

tgtgtagtct gaggcagccg ttgaaaggga gagcaatcgg ccagctgagg ctgatgcttt 240

tatccaaagc gacttagtta ttttacaggg tattggttat tgtccctgga gcaatgttgg 300

gtttggtgtc ttgctcaagg agtaaggata ggtaagcatg ggattcgaac ctgcaaccca 360

cagatcttaa ggccacctcc ctaatcatta agccacggct gcacaggcat atgatgtgtg 420

ttgtccaagc tgagctagct agttagcaag tgcctgagct agttcacatc acgtagaagc 480

aggtcagtgg tttctgcaga gcaccatgat ggttttaaag gctatgctcg gcagaatcct 540

gaaatacatg aaaccctctg aaaattaata gcttgtgcat ttctttgtta gacattcttt 600

ataaat 606

<210> 6
<211> 675
<212> DNA

<213> Engraulis sp.

<400> 6
ctgtatggtt atcagagctt aaacaatgaa atgtgcacta cttatatggc acataggtta 60

actgtttagc aaaaactgtg ttatatgaat gggaaatatg tcaagtcaac aagaatgtgt 120

taacacatat gcacaatctg tcttttgctt tgttgaaata gtgatcatga cgactttgtg 180

tgagtcagtt taaaaatttg tgttaaagcg atcaagaaaa actgtaatat tgctagattt 240

tccatcaata aaaatgtatt gaactatggt agctgttagg tgttgttaac ctgtggattt 300

cccatttta tgcaatgtcc ctttcgtcag tggcatccct tcagtaccct ccacgacccc 360

cctgggggtc ctgcccccca gtttgggaac cactgctgta cagcacagcg cactagtcca 420

catttctatt gtttacggtg caaagccaat gggcctgtca agccacaggt gcagtcaaac 480

tcaaattgtg aaaragaacc gacagcatac agtacggctg atgtgtattt gttttcctaa 540

atgccctcgc aggactgagg gaaagcctgg tggttttcct gtcacagtgt gacaaataag 600

caaaaaggag ggggagaaaa atgcccggga gggcttggac aggattcatg ggcctgcact 660

gctgccacaa agaaa 675

<210> 7
<211> 505
<212> DNA
<213> Engraulis sp.

<400> 7
cttaattgat ggaacatact taattgatgg aacattatca gcggctcact tcttgttttt 60

gttataagat ttttttaggg ctttttgtgc ctttctagtg ataggacagg gagagaggga 120

caggaaatga gtgtgggtca gggaaggacc ccgggtcgcc ggcgtagtag ccttgttccc 180

tattgttaga ccacgatagg tcctcatttg tttggttatg aatatgacat tctagtgtac 240

tccaactccc attgttgtgt gcacactgca cacaacggaa ttgcatgtat gccccaccca 300

tgcaagggga cagcccccaa tagcgcccaa agggagcagt gcggcggcag ggaccgtgct 360

cagggcacct cagtcatgga ggatgatggg gtagagcact ggwtaatgat ttccccccac 420

caacttggtg gttcaggagt tgaacctgca acgcttgggc tacaagtctg acgccctaat 480

cgcttaccca tgactgctcg tcccc 505

<210> 8
<211> 658
<212> DNA
<213> Engraulis sp.

<400> 8
aggctttgtg taggtggcta ttctaataat taaaaataac ccatttccct ttattcttca 60

gaaagatagg cttattgtat atgtaaaatg tggggtacat ggggtgtaac tactatttta 120

aaagcttcca taacaaatgg aatagggaag ccatcactgt ctcgcagggc ttgatgtgat 180

gtcagctggg ggccttcgtg gggttgatga tgcatggtga ctctccaaga ccacgatcat 240

gccatactta tcagcgacga ccagagagta catatgtaac catcagccta ccatgaggga 300

gtcactggtt tggtcctgtt ccaattaggt gggtttcatg ggtgtgcacc tctgatagga 360

cagctgatgt ggtttgcaag tcactcacgt ccgtctccct gttgattagg acatgaaacc 420

atcttgagga cagcacagaa caatgtataa acaacacaca gccactggtg actctttgaa 480

aggtaatgtt ctcttactaa tatgasactt cttcaatgta tgtgcttgta ttgtgagcta 540

aagacttaaa acatcaacta tgcaccagac tatctttatt cgttgtcaac agagtattaa 600

acgcataagg ctgctttact gatgatgctg cagtactgtt tgttacagta taataatg 658

<210> 9
<211> 568
<212> DNA
<213> Engraulis sp.

<400> 9
gccagttgtt ttttaagata gattgaggat atcacaaagt gaacaactaa tggtacggtt 60

gacattagga aaccctgaga taaacactta gcgttgcacc aaaaactttc ttcgttccag 120

taaacaagtg atttgtcttc ggtggtaatt agactattca gcacatttag ctctgagagg 180

tcttaattgg agagggacgt gcagaactgg aagcttggct ttgcattctg atggtcgggt 240

tagcctctgc ggtgaccctt cagagtccga gtcgcatatt caagagctgc tgaaaagaat 300

ggcatcttaa tggatttgaa ttggttgcgg cccaacaaaa cgggagccat tyagtgcacc 360

ttttaaaacg ttctcaagcc tgaacaagga aatgctgaat gtatggggtg taatgtcatg 420

cactatagga gaagtaaagg cgttgctatc ttttgttata aaggttggag caggcttcac 480

ccaacaagtt tttccgcttt tgaaagtgct gaccaaaata ttgtaaaact gaaaaatgag 540

aaaaggtcgc attttctcat ttttcagt 568

<210> 10
<211> 530
<212> DNA
<213> Engraulis sp.

<400> 10
caccaagtga cccagtgtta atgcaggtgg ccccagaaat atggataagc taatctcaat 60

gctccctcat gacctgcaac atcatcaaga gaaacacagt gtgtggtggt atacaacaac 120

atgaacatca gcaagagata cagataggca tagttattgt aaaayagcct ctcctctgct 180

atcctctggc ctggatcgag atacacactg tgccagccag cgttcctcca cagctctcca 240

aaggtcagca tgtgtcggct ctgctcctgc ctgctcagct gataactcat gagaatataa 300

ttgccacatg catcagaccc cacagcggag ggggtccggc cattccgata aacgattcct 360

tttatttttc atttcttaat ccagtcgctg gaaaacaaaa gagaaaatgt taaaaaatga 420

ctgacagagg ggggaattga attaggcaca tttgatagac cttcctttgg ctcgggccga 480

aaccctatga agatttataa gactttggga cttatgcatc acgagtctcg 530

<210> 11
<211> 695
<212> DNA
<213> Engraulis sp.

<220>
<221> misc_feature
<222> (50)..(50)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (54)..(54)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (122)..(122)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (130)..(130)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (134)..(134)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (158)..(158)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (234)..(234)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (238)..(238)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (392)..(392)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (402)..(402)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (451)..(451)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (472)..(472)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (513)..(513)
<223> n is a, c, g, or t

<220>
<221> misc_feature

<222> (533)..(534)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (598)..(599)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (605)..(605)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (615)..(615)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (632)..(632)
<223> n is a, c, g, or t

<400> 11
tacgagagac cctagttgat tgaagcggcg taaagagtgg ttatggaatn tttnacccta 60

aagcagaaaa cctctcaaac tgttatacgc acccagaggt traaacccct tacacgaaag 120

tnactttatn ttcngcctac cagaacccac gaaagctngg atacaaactg ggattagata 180

ccccactatg cctagccrta aactttaatg ctaatgtaca actagcatcc gccngggnac 240

tacgagcacc agcttaaaac ccaaaggact tggcggtgcc tcagaccccc ctagaggagc 300

ctgttctaga accgataacc cccgttcaac ctcaccactc cttgcccttt ccgcctatat 360

accaccgtcg ccagcttacc ctgtgaagga anagtaagcg anatggagac tctccaaaac 420

gtcaggtcga ggtgtagcac acggagtggg nagaaatggg ctacattgcc tnatttaggc 480

tactaacgga aagtcgcctg aaaagactat ttnaaggtgg atttagcagt aannggggaa 540

tagagtgccc ctttgaagcc ggctctgagg cgcgcacaca ccgcccgtca ctctcccnna 600

caacnctaat tcacnagtaa tttaacccta cncatgtaaa aaggggaggc aagtcgtaac 660

atggtaagtg taccggaagg tgcacttgga ataac 695

<210> 12
<211> 370
<212> DNA
<213> Engraulis sp.

<220>
<221> misc_feature
<222> (305)..(305)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (345)..(345)
<223> n is a, c, g, or t

<400> 12
tctatggctt tgagggtgag gatgtcgatg gcaagaagta ctggatcgtc aagaacaggt 60

aatggcaatt crctttcgtg ccattgaatt gcaacacagt gcatttattg ttgtagttgc 120

tattgtattt attgcatatg gacaccccaa gttgacaatt tgtgcgtctt tttgttcttg 180

tccccttgca gctggagtga gaagtggggt gatgagggtt acatctacat ggctaaggac 240

cgcaaaaacc actgtggcat cgccaccgcc gccagctacc cactcatgta gacggaccac 300

tgaangcacc gacactttta gaccacattt cgttttgaaa ggacngctcc tccgtctcac 360

tggctcgctc 370

<210> 13
<211> 574
<212> DNA
<213> Engraulis sp.

<220>
<221> misc_feature
<222> (60)..(60)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (111)..(111)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (120)..(120)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (126)..(126)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (147)..(147)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (156)..(156)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (165)..(165)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (168)..(168)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (171)..(171)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (225)..(225)
<223> n is a, c, g, or t

<220>

<221> misc_feature
<222> (246)..(246)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (249)..(249)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (279)..(279)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (282)..(282)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (288)..(288)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (300)..(300)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (303)..(303)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (318)..(318)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (327)..(327)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (339)..(339)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (390)..(390)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (399)..(399)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (534)..(535)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (543)..(543)
<223> n is a, c, g, or t

<400> 13
gcaggtgtta ctatccttca ccttttattt ttacacgaaa cagggtctaa caacccagcn 60

gggctcaatt ctgacgccga taaaatcgca ttccacccat acttctctta naaagatctn 120

ttaggnttcg cagtaatgct gctggcncta acctcnctgg cattnttntc ncctaactta 180

ttaggagacc ctgataattt taccccagct aatccccttg tcacnccacc tcatattaag 240

cctgantgnt acttcctctt tgcttacgct attctccgnt cnatcccnaa caaactaggn 300

ggngtacttg ccctcctntt ttctatncta gttcttatng tcgtacctat cttgcacacc 360

tctaaacaac gaggaattac tttccgaccn attacgcant tcctattctg aacccttgtt 420

gccgacgtca tcattcttac atgaatcgga ggtatgccag tcgaacaccc attcattatt 480

attggtcaag tagcatcgct actctacttc tctatyttcc ttgtactagc tccnntggca 540

ggntgattag agaacaaagc cctaaactga aact 574

<210> 14
<211> 460
<212> DNA
<213> Engraulis sp.

<220>
<221> misc_feature
<222> (10)..(10)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (41)..(41)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (166)..(166)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (197)..(197)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (217)..(217)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (311)..(311)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (314)..(314)
<223> n is a, c, g, or t

<400> 14
ccacgccatn gggggtgttg ccttctcctg gatcatggcc nctgcttgtg ccgtgcctcc 60

cctggtcggc tggtcccgtt acatccccga gggcatgcag tgctcgtgcg gaatcgacta 120

ctacacrcgt gccgagggct tcaacaacga gtcctttgtc atctanatgt tcgtggtgca 180

cttcatgtgc cccttcntca tcatcacctt ctgctanggt aaccttgtgt gcgcggtcaa 240

ggcagctgcc gcggcgcaac aggagtccga gaccacccag agggccgagc gagaggtcac 300

acgcatggtc ntcntcatgt tcatcgcctt cctggtatgc tgggtgccct acgccagtgt 360

agcctggttc atcttctgca atcagggcag cgagttcggg cccgtcttca tgaccatgcc 420

cgccttcttt gccaagagct ccgccatcta caacccccctc 460

<210> 15
<211> 393
<212> DNA
<213> Engraulis sp.

<220>
<221> misc_feature
<222> (39)..(39)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (67)..(68)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (129)..(129)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (174)..(174)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (177)..(177)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (214)..(214)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (224)..(224)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (227)..(227)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (258)..(259)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (323)..(323)
<223> n is a, c, g, or t

<220>

<221> misc_feature
<222> (380)..(380)
<223> n is a, c, g, or t

<400> 15
tacaatgatg ttatattttg ataatgctag taagcaaana ggtgatgtgt atgccaagtg 60

cagatannat gctattttaa tgatattata atgacattat aatggtataa taatgaccgt 120

tttacagtnt cctccattgg tttgtctgat ttcttgaaag tgattgtaat cttntantca 180

tgctgttaaa ttccaaggct tccatgacga gtcnccattg attnaangtg gctgtgctat 240

tctaaatatt gaacaatnnc ttttaagtgg gtatayagtt agcaatttta ggaaggtata 300

ttgaaattcc caacctttc tgntgggtat actgcatgat ccacacatcg gattgttcac 360

agctccagag aagagccaan ccaaaggaaa tca 393

<210> 16
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador clon369-F

<400> 16
gcacacaaat catcacaagc 20

<210> 17
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador clon369-R

<400> 17
tgcccatctc cctaaatctc 20

<210> 18
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador clon286-F

<400> 18
gtgcattggt gagtggtgag 20

<210> 19
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador clon286-R

<400> 19
tgtgggtgat tagtgctgtg 20

<210> 20

<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador clon491-F

<400> 20
gccacccatc agaagtacag 20

<210> 21
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador clon491-R

<400> 21
cattgcctca cagtcaaagg 20

<210> 22
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador clonM3EK-F

<400> 22
aatcgaccag acgtgatgtg 20

<210> 23
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador clonM3EK-R

<400> 23
cagagggttt catgtatttc agg 23

<210> 24
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador clon289-F

<400> 24
gcacaatctg tcttttgctt tg 22

<210> 25
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador clon289-R

<400> 25
accaggcttt ccctcagtc 19

<210> 26
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador clon470-F

<400> 26
atcagcggct cacttcttgt 20

<210> 27
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador clon470-R

<400> 27
gcgattaggg cgtcagactt 20

<210> 28
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador clon357-F

<400> 28
tggggtacat ggggtgtaac 20

<210> 29
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador clon357-R

<400> 29
gcagcatcat cagtaaagca g 21

<210> 30
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador clonM4-F

<400> 30
atggtacggt tgacattagg 20

<210> 31
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador clonM4-R

<400> 31
ccttgttcag gcttgagaac 20

<210> 32
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador clon483-F

<400> 32
gccccagaaa tatggataag c 21

<210> 33
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador clon483-R

<400> 33
ccgagccaaa ggaaggtc 18

<210> 34
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador MT-12S-F

<400> 34
cggtaaaact cgtgccag 18

<210> 35
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador MT-12S-R

<400> 35
gcaccttccg gtacactta 19

<210> 36
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador CATL7-F

<400> 36
ctatggcttt gagggtgagg 20

<210> 37
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador CATL7-R

<400> 37
gagcgagcca gtgagacg 18

<210> 38
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador SNP1,2-F

<400> 38
tgtcattact aaccttatgt ctgcagtt 28

<210> 39
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador SNP1,2-R

<400> 39
ccgaaaagcc ycctcaaatt c 21

<210> 40
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador SNP3-F

<400> 40
gcccatcaat cttgcagtct gt 22

<210> 41
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador SNP3-R

<400> 41
ggaaatgatc acatgccaag ca 22

<210> 42
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador SNP4-F

<400> 42
ggacccacac atctcactgt ac 22

<210> 43

<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador SNP4-R

<400> 43
gaccatgacc caacacaaca tttg 24

<210> 44
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador SNP8-F

<400> 44
accctaaagc agaaaacctc tcaaa 25

<210> 45
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador SNP8-R

<400> 45
agctttcgtg ggttctggta 20

<210> 46
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador SNP9-F

<400> 46
ctaccagaac ccacgaaagc t 21

<210> 47
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador SNP9-R

<400> 47
ggttttaagc tggtgctcgt agt 23

<210> 48
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador SNP11-F

<400> 48
acccattcat tattattggt caagtagca 29

<210> 49
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador SNP11-R

<400> 49
agtttcagtt tagggctttg ttctct 26

<210> 50
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador SNP5-F

<400> 50
tggtggtctg gtctggtga 19

<210> 51
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador SNP5-R

<400> 51
ggagacagca caggatgatg ag 22

<210> 52
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador SNP12-F

<400> 52
gctcgtgcgg aatcgacta 19

<210> 53
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador SNP12-R

<400> 53
acatgaagtg caccacgaac at 22

<210> 54
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador SNP13-F

<400> 54
tgctgttaaa ttccaaggct tccat 25

<210> 55
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador SNP13-R

<400> 55
cagaaaaggt tgggaatttc aatatacct 29

<210> 56
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador SNP6-F

<400> 56
tcagctggcc gattgct 17

<210> 57
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador SNP6-R

<400> 57
ggttcacaag tggttgagct aca 23

<210> 58
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador SNP10-F

<400> 58
tggatcgtca agaacaggta atgg 24

<210> 59
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador SNP10-R

<400> 59
gcactgtgtt gcaattca 18

<210> 60
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador SNP7-F

<400> 60
aagccacagg tgcagtca 18

<210> 61
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador SNP7-R

<400> 61
cacatcagcc gtactgtatg ct 22

<210> 62
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador SNP16-F

<400> 62
cacctcagtc atggaggatg atg 23

<210> 63
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador SNP16-R

<400> 63
ttcaactcct gaaccaccaa gtt 23

<210> 64
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador SNP15-F

<400> 64
cacacagcca ctggtgact 19

<210> 65
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador SNP15-R

<400> 65
agtctttagc tcacaataca agcacat 27

<210> 66

<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador SNP14-F

<400> 66
caacgccctg ctggga 16

<210> 67
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador SNP14-R

<400> 67
ggtgccgcag catctc 16

<210> 68
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador SNP18-F

<400> 68
gctgctgaaa agaatggcat ct 22

<210> 69
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador SNP18-R

<400> 69
cttgttcagg cttgagaacg 20

<210> 70
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador SNP17-F

<400> 70
caacaacatg aacatcagca agagat 26

<210> 71
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador SNP17-R

<400> 71
gcacagtgtg tatctcgatc ca 22

<210> 72
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Sonda SNP1,2-FAM

<400> 72
ccttacatgg gcactga 17

<210> 73
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Sonda SNP3-VIC

<400> 73
tggctgtcgt taaccat 17

<210> 74
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Sonda SNP3-FAM

<400> 74
atggctgtcg ttcaccat 18

<210> 75
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Sonda SNP4-VIC

<400> 75
aggaggacac gttttaga 18

<210> 76
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Sonda SNP4-FAM

<400> 76
aggaggacac cttttaga 18

<210> 77
<211> 14
<212> DNA
<213> Artificial Sequence

<220>
<223> Sonda SNP8-VIC

<400> 77
agaggttgaa accc 14

<210> 78
<211> 14
<212> DNA
<213> Artificial Sequence

<220>
<223> Sonda SNP8-FAM

<400> 78
agaggttaaa accc 14

<210> 79
<211> 15
<212> DNA
<213> Artificial Sequence

<220>
<223> Sonda SNP9-VIC

<400> 79
cctagccgta aactt 15

<210> 80
<211> 15
<212> DNA
<213> Artificial Sequence

<220>
<223> Sonda SNP9-FAM

<400> 80
cctagccata aactt 15

<210> 81
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Sonda SNP11-VIC

<400> 81
ctagtacaag gaagatagag 20

<210> 82
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Sonda SNP11-FAM

<400> 82
tagtacaagg aaaatagag 19

<210> 83
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Sonda SNP5-VIC

<400> 83
ccttatgccc tcgaactg 18

<210> 84
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Sonda SNP5-FAM

<400> 84
cttatgcccc cgaactg 17

<210> 85
<211> 15
<212> DNA
<213> Artificial Sequence

<220>
<223> Sonda SNP12-VIC

<400> 85
tcggcacgcg tgtag 15

<210> 86
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<223> Sonda SNP12-FAM

<400> 86
ctcggcacgt gtgtag 16

<210> 87
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Sonda SNP13-VIC

<400> 87
aattgctaac tgtataccc 19

<210> 88
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Sonda SNP13-FAM

<400> 88
aattgctaac tatataccc 19

<210> 89

<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<223> Sonda SNP6-VIC

<400> 89
cactcactgg ccttag 16

<210> 90
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Sonda SNP6-FAM

<400> 90
acactcacta gccttag 17

<210> 91
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<223> Sonda SNP10-VIC

<400> 91
atggcacgaa agcgaa 16

<210> 92
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<223> Sonda SNP10-FAM

<400> 92
atggcacgaa agtgaa 16

<210> 93
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Sonda SNP7-VIC

<400> 93
caaattgtga aagagaacc 19

<210> 94
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Sonda SNP7-FAM

<400> 94
aaattgtgaa aaagaacc 18

<210> 95
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Sonda SNP16-VIC

<400> 95
tagagcactg gataatga 18

<210> 96
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Sonda SNP16-FAM

<400> 96
tagagcactg gttaatga 18

<210> 97
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Sonda SNP15-VIC

<400> 97
cattgaagaa gtctcatatt a 21

<210> 98
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Sonda SNP15-FAM

<400> 98
attgaagaag tgtcatatta 20

<210> 99
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<223> Sonda SNP14-VIC

<400> 99
agccccaccc tgagca 16

<210> 100
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<223> Sonda SNP14-FAM

<400> 100
agccccaccc cgagca 16

<210> 101
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Sonda SNP18-VIC

<400> 101
ttttaaaagg tgcactgaat 20

<210> 102
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Sonda SNP18-FAM

<400> 102
ttaaaaggtg cactaaat 18

<210> 103
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Sonda SNP17-VIC

<400> 103
aggagaggct gttttaca 18

<210> 104
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Sonda SNP17-FAM

<400> 104
aggagaggct attttaca 18

<210> 105
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador CytBI-7F (Jerome et al., 2003b)

<400> 105
ctaacccgat tctttgcctt ccacttcct 29

<210> 106
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador Trucytb-R (Jerome et al., 2003b)

<400> 106
ccgacttccg gattacaaga ccg 23

<210> 107
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador L14841 (Kocher et al. 1989)

<400> 107
aaaaagcttc catccaacat ctccagcatg atgaaa 36

<210> 108
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Cebador H15149 (Kocher et al. 1989)

<400> 108
aaactgcagc ccctcagaat gatatttgtc ctca 34

OFICINA ESPAÑOLA
DE PATENTES Y MARCAS

ESPAÑA

(21) N.º solicitud: 201130534

(22) Fecha de presentación de la solicitud: 04.04.2011

(32) Fecha de prioridad:

INFORME SOBRE EL ESTADO DE LA TECNICA

(51) **Int. Cl. :** ***C12Q1/68*** (2006.01)

DOCUMENTOS RELEVANTES

Categoría	(56) Documentos citados	Reivindicaciones afectadas
A	PAKAKI, V. et al., 'New polymorphic microsatellite loci for population studies in the European anchovy, Engraulis encrasicolus (L.)', MOLECULAR ECOLOGY RESOURCES, 2009, Vol. 9, No. 5, páginas 1406-1409. ISSN: 1755-098X, todo el documento.	1-18
A	MAGOULAS, A. et al., 'Mitochondrial DNA reveals a mosaic pattern of phylogeographical structure in Atlantic and Mediterranean populations of anchovy (Engraulis encrasicolus)', MOLECULAR PHYLOGENETICS AND EVOLUTION, 2006, Vol. 39, No. 3, páginas 734-746, ISSN: 1055-7903, todo el documento.	1-18
A	BORSA, P. et al., 'Nuclear-DNA markers confirm the presence of two anchovy species in the Mediterranean', COMPTES RENDUS BIOLOGIES, 2004, Vol. 327, No. 12, páginas 1113-1123, ISSN: 1631-0691, todo el documento.	1-18
A	WO 2011009975 A2 (UNIV. CADIZ) 27.01.2011, todo el documento.	1-18

Categoría de los documentos citados
X: de particular relevancia
Y: de particular relevancia combinado con otro/s de la misma categoría
A: refleja el estado de la técnica
O: referido a divulgación no escrita
P: publicado entre la fecha de prioridad y la de presentación de la solicitud
E: documento anterior, pero publicado después de la fecha de presentación de la solicitud

El presente informe ha sido realizado

☒ para todas las reivindicaciones ☐ para las reivindicaciones nº:

Fecha de realización del informe	Examinador	Página
27.11.2012	J. L. Vizán Arroyo	1/4

Documentación mínima buscada (sistema de clasificación seguido de los símbolos de clasificación)

C12Q

Bases de datos electrónicas consultadas durante la búsqueda (nombre de la base de datos y, si es posible, términos de búsqueda utilizados)

INVENES, EPODOC, BIOSIS, MEDLINE, EMBASE,EBI

Fecha de Realización de la Opinión Escrita: 27.11.2012

Declaración

Novedad (Art. 6.1 LP 11/1986)	Reivindicaciones	1-18	**SI**
	Reivindicaciones		**NO**
Actividad inventiva (Art. 8.1 LP11/1986)	Reivindicaciones	1-18	**SI**
	Reivindicaciones		**NO**

Se considera que la solicitud cumple con el requisito de aplicación industrial. Este requisito fue evaluado durante la fase de examen formal y técnico de la solicitud (Artículo 31.2 Ley 11/1986).

Base de la Opinión.-

La presente opinión se ha realizado sobre la base de la solicitud de patente tal y como se publica.

1. Documentos considerados.-

A continuación se relacionan los documentos pertenecientes al estado de la técnica tomados en consideración para la realización de esta opinión.

Documento	Número Publicación o Identificación	Fecha Publicación
D01	PAKAKI, V. et al., *Mol. Ecol. Resour.*, (2009), 9(5): 1406-9.	2009
D02	MAGOULAS, A. et al., *Mol. Phylogenet. Evol.*, (2006), 39(3): 734-46.	2006
D03	BORSA, P. et al., *C. R. Biol.*, (2004), 327(12): 1113-23.	2004
D04	WO 2011009975 A2 (UNIV. CADIZ)	27.01.2011

En D1-D4 se describen diferentes métodos de identificación genética de la anchoa europea (*Engraulis encrasicolus*).

2. Declaración motivada según los artículos 29.6 y 29.7 del Reglamento de ejecución de la Ley 11/1986, de 20 de marzo, de Patentes sobre la novedad y la actividad inventiva; citas y explicaciones en apoyo de esta declaración

1. NOVEDAD (Art. 4.1. y Art. 6.1. de la Ley de Patentes) y ACTIVIDAD INVENTIVA (Art. 4.1. y Art. 8.1. de la Ley de Patentes).

1.1. Reivindicaciones independientes 1 y 3.

1.1.1. El objeto de las reivindicaciones 1 y 3 consiste respectivamente en el conjunto de 15 secuencias constituido por SEQ ID NO 1 a SEQ ID NO 15 caracterizadas porque comprenden las SNPs identificadas en la descripción (cf. Tabla 1) y en un método de identificación genética del origen geográfico de la anchoa europea (*Engraulis encrasicolus*) basado .en la detección de las SNPs de la reivindicación 1. En el estado de la técnica más próximo, constituido por los documentos D1-D4, no se ha divulgado ningún conjunto de secuencias que comprenda las SNPs concernidas, ni método de identificación del origen geográfico de *Engraulis encrasicolus* que comparta las mismas características técnicas del procedimiento reivindicado en la solicitud de patente. Además, dicho procedimiento no se deduce de una manera obvia del estado de la técnica mencionado.

1.2. La presente solicitud satisface el criterio establecido en el Art. 4. 1. de la Ley de Patentes, pues el objeto de las reivindicaciones 1-18 es nuevo y tiene actividad inventiva de acuerdo con los Arts. 6.1. y 8.1. de la Ley de Patentes.