

Primeros pasos para el análisis de las lesiones en secuencias específicas del ADN de ciervo: secuenciación del ADN

Jéssica Alonso Molero¹

Alumna de 5º de Licenciatura en Biología y Becaria en el Instituto de Desarrollo Ganadero y Sanidad Animal de la Universidad de León. Curso 2012/2013

1.- jalonmoo@estudiantes.unileon.es

Trabajo de investigación dirigido y supervisado por el Dr. Felipe Martínez Pastor (felipe.martinez@unileon.es).

El ciervo rojo tiene una gran proyección como especie ganadera: pero se conoce poco de su genética en comparación con otras especies domésticas. Tanto es así que, tres años después de anunciar la completa secuenciación del genoma de vaca (*Bos taurus*) apenas hay secuencias nucléares de cérvidos disponibles. En el Instituto de Desarrollo Ganadero y Sanidad Animal (INDEGSAL) de la Universidad de León existe una línea de investigación sobre la calidad espermática de esta especie. En estos estudios estamos intentando aplicar técnicas de biología molecular para determinar lesiones en sus secuencias específicas de ADN. Esto requiere la disponibilidad de la secuencia de los genes de interés, motivo por el cual estamos en proceso de secuenciar varios fragmentos de su ADN. En este artículo repasamos la técnica de secuenciación que estamos utilizando, los obstáculos con los que nos hemos encontrado y algunos resultados que vamos obteniendo.

Palabras Clave: Ciervo rojo, secuenciación, ADN, PCR, electroforesis capilar.

Introducción

Mi nombre es Jéssica Alonso, y como muchos de los lectores de esta revista, soy estudiante en la Facultad de Ciencias Ambientales y Biológicas. En este momento estoy cursando 5º de Licenciatura de Biología, pero ya llevo algún tiempo en el mundo de la investigación. Llegué a este mundo al ponerme en contacto con el Dr. Felipe Martínez, investigador del INDEGSAL y del Departamento de Biología Molecular. Él me recibió, hace ya dos años, en su grupo de investigación, incluyéndome en algunas de las líneas de investigación que coordina. Nuestra intención es estudiar cómo afectan los daños del ADN a diferentes funciones y procesos celulares, centrándonos en los espermatozoides.

En este momento me encuentro colaborando en un proyecto que tiene como objetivo utilizar técnicas de biología molecular para evaluar estos daños en

el ADN espermático. Estas técnicas se basan en la utilización de PCR en tiempo real (Rothfuss et al., 2010), comparando la velocidad de acumulación de producto de la reacción en la muestra problema y en una muestra control, ya que la progresión de la polimerasa se verá entorpecida, e incluso impedida, si la cadena molde presenta daños.

Con este tipo de estrategia, un requisito esencial es conocer las secuencias de ADN de las especies con las que trabajamos, para poder diseñar cebadores (primers) específicos. Este es un problema si no están disponibles estas secuencias (por ejemplo en NCBI (<http://www.ncbi.nlm.nih.gov/nuccore>)), en cuyo caso debemos diseñar cebadores basándonos en las secuencias publicadas de especies filogenéticamente cercanas. Con el tiempo, me he convertido en la responsable de esta tarea. Tengo que decir que no trabajo sola en ello. He tenido a mucha gente enseñándome y colaborando conmigo en esta tarea, y actualmente me ayudan Adolfo Velázquez y Cristina Fuentes, estudiantes de Biología. María Mata y Leticia Ordás, doctorandas del grupo, se encargan de realizar los análisis utilizando los cebadores diseñados gracias a las secuencias que obtenemos.

Comenzando por lo más básico, debemos recordar que el ADN (ácido desoxi-ribonucleótido) es una molécula definida por una secuencia de nucleótidos mediante los cuales se cifra una información genética. El orden de las bases nitrogenadas es crítico para cifrar esta información. La secuenciación del ADN tiene como objetivo estudiar el orden de estas bases mediante una serie de métodos y técnicas bioquímicas, de modo que con estas secuencias se puedan realizar diversos estudios.

Determinar la secuencia de ADN es muy útil en muchos campos de investigación y ha ayudado a lograr diversos descubrimientos en la biología. A nivel práctico, sabemos que la información proporcionada por la secuenciación del genoma humano podría ser de gran interés para conocer la predisposición de una persona concreta a padecer ciertas enfermedades. En investigación, hay muchas técnicas que se basan en la utilización de secuencias específicas para lograr distintos fines: utilización de esas secuencias como sondas, realización de PCR para detectar expresión génica, etc.

Hasta el momento, se ha conseguido secuenciar el genoma de más de 300 organismos, generalmente los más relevantes para distintos campos de conocimiento. Desafortunadamente para nosotros, apenas se dispone de secuencias nucleares del ciervo rojo (*Cervus elaphus*). En el INDEGSAL estamos trabajando en la reproducción de esta especie por el valor que representa como ganadería alternativa. Si bien es posible utilizar la secuencia de una especie

similar o secuencias muy conservadas evolutivamente, hay muchas técnicas que requieren de secuencias muy específicas. Este es nuestro caso, ya que utilizamos PCR cuantitativa en tiempo real (Real Time-PCR) para detectar modificaciones (daños) que pueden tener una frecuencia muy baja. Hemos comprobado que las técnicas que utilizamos se ven muy afectadas por la calidad de los cebadores, debido a que amplificamos secuencias de alrededor de 1 kb ("semi-largas"), y la eficiencia de la reacción disminuye considerablemente si los cebadores no hibridan correctamente.

El hecho de que apenas se disponga de secuencia de ciervo, hace que tengamos que recurrir a especies cercanas a él, como son la oveja doméstica (*Ovis aries*) y la vaca doméstica (*Bos taurus*) para poder obtener posibles oligonucleótidos que nos permitan conseguir el material necesario para llevar a cabo la secuenciación. Todas estas especies pertenecen al orden de los artiodáctilos (Artiodactyla). Recordemos que este orden agrupa a mamíferos ungulados cuyas extremidades terminan en un número par de dedos, de los cuales los más desarrollados son los que ocupan la tercera y la cuarta posición (realmente, Artiodactyla se considera parafilético, siendo filogenéticamente correcto Cetartiodactyla, que incluye a los cetáceos). Habitán en todos los continentes, excepto en la Antártida, aunque los que se encuentran en Australia han sido introducidos por el hombre. En este orden se incluyen unas 235 especies no extintas, repartidas en 10 familias. La **Fig. 1** muestra la relación filogenética entre estas familias, y donde se puede apreciar la proximidad de las especies que estamos usando en nuestro trabajo.

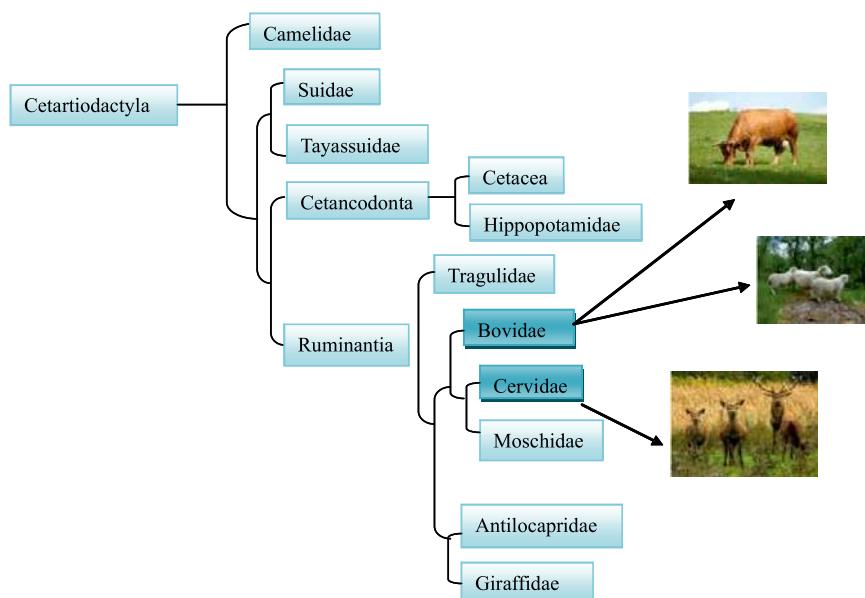


Figura 1. Árbol filogenético del orden Cetartiodactyla.

En este artículo voy a describir las técnicas que estamos realizando para secuenciar algunos genes de ciervo ibérico, y los primeros resultados que hemos obtenido.

Metodología

El protocolo de secuenciación se puede resumir en varias etapas que detallaré a continuación y que no tienen por qué ir en este orden, ya que dependiendo del experimento podrían cambiar (la **Fig. 2** muestra un esquema general de distintas estrategias para la preparación de la muestra antes de la secuenciación). Estas etapas son:

1. Extracción de ADN (en mi caso, de espermatozoides o de otros tejidos de ciervo rojo).
2. Búsqueda de cebadores utilizando secuencias de oveja o vaca.
3. Realización de las diferentes PCR (primero para comprobar si los cebadores amplifican secuencias del ADN de ciervo y posteriormente para obtener una cantidad de producto aceptable para secuenciar).
4. Obtención de bandas en un gel, mediante electroforesis de los productos de las PCR.
5. Purificación de las bandas obtenidas.
6. Secuenciación.

Según lo comentado anteriormente, lo que hacemos es obtener el ADN a partir, en nuestro caso, de espermatozoides de ciervo rojo, puesto que es nuestro principal material de trabajo. Estos espermatozoides los tenemos congelados y reservados en tanques de nitrógeno líquido, a -196 °C, en dosis de 250 µl, con una concentración de 100 millones de espermatozoides por ml. La descongelación de estas dosis la realizamos en un baño a 37 °C durante 30 segundos.

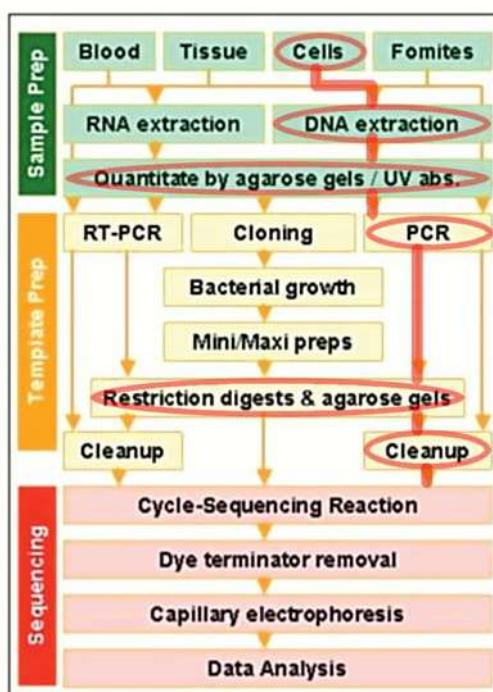


Figura 2. Esquema de distintas estrategias de secuenciación, dependiendo del tipo de ácido nucleico y del origen de la muestra. La estrategia seguida en nuestro trabajo está resaltada en rojo.

El protocolo de extracción del ADN es sencillo. El primer paso que realizamos es una reducción de puentes disulfuro de las protaminas y digestión de las proteínas con la proteinasa K. Posteriormente, separamos el ADN del resto de los componentes celulares, mediante el método del fenol:cloroformo y la precipitación con etanol. Para ello, partimos de la muestra acuosa procedente de la digestión con proteinasa K, y realizamos una serie de centrifugaciones en presencia de fenol: cloroformo, que disuelve las proteínas y los componentes lipídicos, quedando el ADN en la fase acuosa. Por último, añadimos acetato sódico y etanol absoluto sobre la fase acuosa, para hacer que el ADN precipite y forme un ovillo visible, el cual recogemos e introducimos en un eppendorf vacío estéril hasta que se evapora el etanol. Finalmente, lo diluimos en agua de Biología Molecular (agua BioMol), identificamos el tubo y lo conservamos a -20 °C.

Posteriormente, se lleva a cabo la obtención de cebadores o primers. Para ello, lo que hacemos es acudir a diferentes bases de datos, como el NCBI o el Ensembl Genome Browser (<http://www.ncbi.nlm.nih.gov/tools/primer-blast/>). Localizamos los genes que nos interesan, generalmente en vaca, y determinamos la estructura de intrones y exones. Debemos tener en cuenta que nuestro objetivo final es amplificar una secuencia de ~1 kb. Debido a esto, generalmente debemos diseñar los cebadores en dos exones (la secuencia estará más conservada —más similar entre especies— en los exones) separados por un intrón de menos de 1 kb.

Estas secuencias las tratamos con programas bioinformáticos (Primer-Blast, <http://www.ncbi.nlm.nih.gov/tools/primer-blast/>, nos ha dado muy buenos resultados), en los que se introduce la secuencia de ADN y devuelven una lista de posibles cebadores, de acuerdo con unas condiciones predefinidas (longitud del amplicón, Tm, etc.). Para descartar problemas (tales como que los cebadores hibriden entre ellos, que formen horquillas, etc.), utilizamos la herramienta Oligoanalyzer, que nos ofrece una serie de análisis que nos permiten determinar la calidad de los cebadores. Una vez seleccionados, los cebadores se piden a una casa comercial que los sintetiza y nos los envía liofilizados.

Con todo este material, realizamos la PCR para determinar si los cebadores hibridan y se obtienen productos del tamaño deseado. Esta PCR de prueba nos permite hacer una selección de los cebadores que mejor funcionan, es decir, si tras la electroforesis en gel obtenemos una banda intensa, definida y del tamaño deseado. Nuestra rutina de PCR es la habitual (Barlett y Stirling, 2003), con una mezcla de reacción incluyendo cebadores (forward y reverse), el ADN extraído de los tejidos de ciervo y la polimerasa (nosotros utilizamos una Taq

polimerasa —de *Thermophilus aquaticus*— optimizada). A la reacción le añadimos otra serie de compuestos necesarios para la reacción: MgCl₂, que funciona como cofactor, los dNTPs (desoxirribonucleótidos-trifosfato), los cuales son utilizados para sintetizar la secuencia complementaria a la cadena de ADN y el tampón óptimo para el funcionamiento de la enzima.

Todo ello lo introducimos en el termociclador, un aparato que permite realizar los ciclos de temperaturas necesarios para que la polimerasa pueda realizar la amplificación. Normalmente utilizamos entre 25 y 35 ciclos, y en cada uno de ellos deben programarse tres fases (**Fig. 3**):

- 1- Desnaturalización
- 2- Hibridación
- 3- Extensión

Antes de comenzar con dichas etapas, solemos incluir una fase preliminar, denominada fase de inicio. En ella el termociclador alcanza la temperatura de 94°C y la mantiene durante 5 minutos para asegurar la activación de la Taq polimerasa. Acabada esta última fase, en el último ciclo se llevan a cabo otras dos fases, una en la que se asegura la total amplificación de la región deseada del ADN, denominada elongación final y que se lleva a cabo durante 7 minutos a 72°C, y otra que es la de conservación, en la que la temperatura desciende a 4°C, de modo que el ADN no se degrade hasta que retiremos los tubos del termociclador.

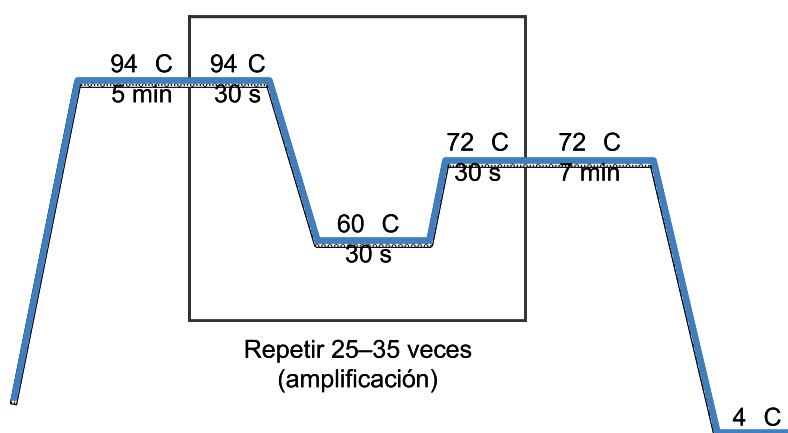


Figura 3. Esquema de un programa de PCR típico. La temperatura se representa en vertical y el tiempo en horizontal (escalas no proporcionales).

Tras realizar este paso, realizamos una electroforesis del producto obtenido en un gel de agarosa. El porcentaje de ésta depende del tamaño de producto esperado. Como para la secuenciación se usan productos de un tamaño

semi-largo (~1kb), la concentración utilizada suele ser del 1%. Este gel se examina primero para comprobar si las bandas son definidas y del tamaño esperado, descartando el experimento si se presenten bandas inespecíficas o smear (un borrón debido a la presencia de muchos fragmentos con tamaños muy variados. No obstante, a menudo encontramos bandas bien definidas pero que se sitúan en una posición que no corresponde al tamaño esperado. Esta diferencia de tamaño se debe muchas veces no a una amplificación inespecífica, sino a que en el genoma del ciervo ese gen tiene un intrón mayor o menor que el ADN de la especie en la que hemos obtenido los cebadores.

Una vez que comprobamos que los cebadores funcionan correctamente, repetimos la PCR con las condiciones óptimas que hemos seleccionado en las pruebas anteriores, pero con un mayor número de ciclos y triplicando las reacciones, para conseguir una mayor concentración y un volumen final de 60 µL y no de 20 µL, que son los que se utilizan en cada reacción. El producto se somete de nuevo a electroforesis en un gel de agarosa al 1% y, en el transiluminador, se corta el fragmento de gel que contiene la banda deseada. La extracción del ADN (para secuenciar) se realiza mediante un kit comercial (NucleoSpin Gel and PCR Clean-Up, de Macherey-Nagel). Colocamos el fragmento del gel en un tampón y lo sometemos a una incubación en caliente para licuar el gel. El producto se pasa por una columna que atrapa el ADN, el cual se recupera tras una serie de lavados. Finalmente, se eluye el ADN en agua BioMol.

Con ese ADN repetimos la PCR, exactamente igual que antes, pero en este caso, el ADN utilizado es el obtenido de la amplificación anterior. Lo que buscamos es obtener un gran número de copias de esa zona, para poder mandarlo a secuenciar. El procedimiento es idéntico al descrito anteriormente.

Tras esto, se pone en marcha la primera de las tres fases de cada ciclo, la de desnaturación. Esta etapa se lleva a cabo a 94 °C durante 30 s, de modo que las dos hebras de ADN se separen y permitan la acción posterior de la enzima. Posteriormente, se lleva a cabo la segunda fase, la de hibridación. La temperatura en ésta suele variar y depende principalmente de los primers diseñados, puesto que el objetivo es que estos se unan a la hebra molde de ADN. Nosotros solemos diseñar los cebadores con una temperatura de fusión (melting temperature, Tm) de 60 °C, por lo que las PCR de prueba las realizamos a tres temperaturas: 58, 60 y 62 °C. El tiempo que dura esta fase es de 30 s. En la tercera fase, extensión, la polimerasa se une y sintetiza la hebra complementaria a la hebra de ADN original mediante la unión de dNTPs al extremo 3' del primer. Para ello se requiere de una temperatura de 72 °C y un tiempo que depende del

tamaño de producto para el que esté diseñado el primer (en nuestro caso, normalmente basta con 30 s).

El método de secuenciación que utilizamos se denomina “secuenciación por electroforesis capilar” (Wilkinson et al., 2000, Katherine, 2002). Para ello utilizamos el Servicio de Secuenciación del Laboratorio de Técnicas Instrumentales de la ULE. Este servicio está a cargo del técnico Benjamín Rabanal, a quien se le entrega el ADN a secuenciar y los cebadores específicos de esa secuencia. El secuenciador de ADN que utilizan es un MEGABACE 500 (Amersham Bioscience) (**Fig. 4**), basado en electroforesis capilar, que usa el kit de secuenciación DYEnamic ET Dye Terminator Kit (GE Healthcare, 2012). La Universidad de León usa este tipo de secuenciador por varias razones, como que la secuenciación se lleva a cabo de forma rápida y eficaz, en una sola reacción y con bajo coste, en comparación con otros métodos de secuenciación.



Figura 4. Secuenciador MEGABACE 500 del Servicio de Secuenciación de la Universidad de León.

El método de secuenciación está basado en el llamado "método de terminación de cadena", seguido por una electroforesis capilar. El método de terminación de cadena (desarrollado por Sanger y Coulson, 1975) consiste en realizar una PCR en la cual se ha añadido, además de los cuatro desoxinucleótidos trifosfato (dNTPs), una pequeña concentración de los cuatro nucleótidos modificados (ddNTPs). Esta modificación consiste en la eliminación del hidroxilo 3', con lo cual, si la polimerasa añade un ddNTP, la extensión termina. Cada uno de estos ddNTPs está marcado con un colorante fluorescente

distinto. Por lo tanto, tras realizar una PCR en estas condiciones tendremos fragmentos de la secuencia molde de distintas longitudes y marcados con un fluorocromo (molécula fluorescente) distinto.

En un segundo paso, el producto de la PCR es sometido a electroforesis a través de un capilar que está lleno con un polímero que funciona del mismo modo que el gel de agarosa, de modo que es capaz de separar los diferentes fragmentos de ADN por tamaños. En el extremo del capilar que coincide con el polo negativo se localiza un láser, que ilumina cada uno de los fragmentos de ADN que recorren el capilar, detectándose una longitud de onda diferente dependiendo del ddNTP añadido en la terminación de ese fragmento. Posteriormente, esas señales luminosas son tratadas con un software que las convierte en datos digitales, permitiéndonos conocer de qué nucleótido se trata y, por tanto, la secuencia del fragmento de ADN en estudio (**Fig. 5**).

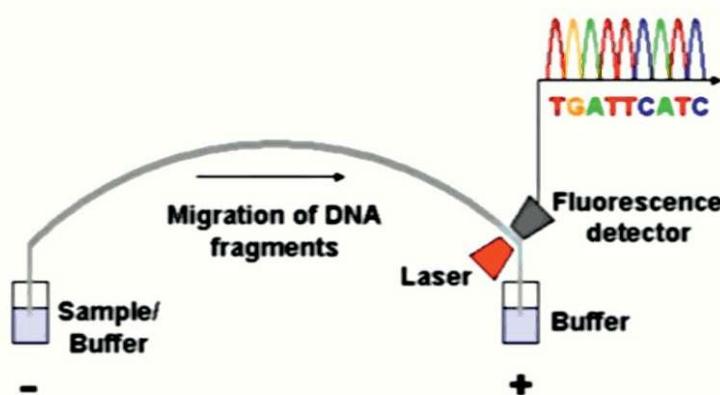


Figura 5. Esquema de la secuenciación capilar. La muestra, tras la PCR con los nucleótidos terminadores marcados, migra a través de un capilar y la fluorescencia de los fragmentos sucesivos es leída mediante la excitación con un láser.

El DYEnamic ET Dye Terminator Kit es utilizado por el Servicio de Secuenciación porque permite una secuenciación más sensible, a pesar de que la muestra no esté completamente purificada, de modo que los resultados no se ven alterados por la presencia de ciertas sustancias, como las sales. Además muestra una eficiente incorporación del fluorocromo a los diferentes terminadores, lo cual permite largos tiempos de lectura, con los que se obtienen picos de fluorescencia uniformes, con lo cual se obtienen menos errores al procesar los resultados con el software.

Primeros resultados

Mediante el protocolo descrito anteriormente hemos conseguido, de momento, secuenciar un fragmento del gen que codifica para la β -globina (956 pb entre los exones 2 y 3; **Fig. 6**) y dos fragmentos del gen h19 (312 pb entre el exón 1 y el intrón 2, y 575 pb entre el intrón 3 y el exón 5). Además, hemos conseguido diseñar cebadores, utilizando principalmente secuencias de toro, con las que podemos amplificar fragmentos de distintos genes en ADN de ciervo (**Fig. 7**), y que estamos utilizando para conseguir secuenciar esos fragmentos.

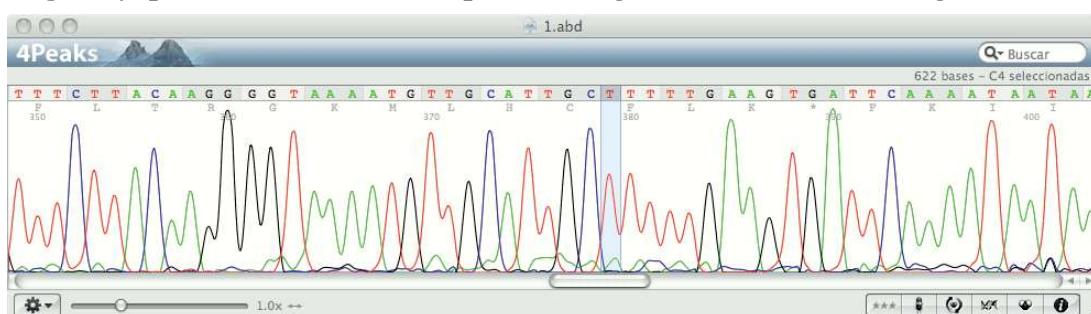


Figura 6. Análisis del resultado de la secuenciación capilar utilizando el software 4Peaks (<http://nucleobytes.com/index.php/4peaks>). Este fragmento corresponde al gen de la β -globina, uno de los primeros en ser secuenciados en ciervo por nuestro grupo.



bandas inferiores corresponden a tamaños de banda de 500 pb y 1 kb, mientras que el resto de los carriles corresponden alternativamente a reacciones en las que se utilizó ADN de toro y de ciervo, a tres temperaturas de hibridación diferentes (58 °C, 60 °C y 62 °C). En este caso, obtuvimos una banda definida y del tamaño esperado en todas las reacciones con ADN de toro, y en el caso del ADN de ciervo, obtuvimos una banda aceptable y de tamaño ligeramente superior cuando realizamos la reacción con una temperatura de hibridación de 60 °C.

Figura 7. Imagen de un gel tras realizar una electroforesis a los productos de seis reacciones de PCR. En estas reacciones se utilizaron cebadores para una secuencia de 843 pb del gen dlk1, utilizando secuencias publicadas para vaca (NCBI Gene ID 281117). El primer carril a la izquierda es una escalera cuyas

Hemos empleado otra técnica para incrementar la calidad y cantidad del ADN, pero fue abandonada al no obtener unas ventajas claras. En este caso, buscamos amplificar el fragmento aislado del gel tras PCR, clonándolo en un vector y transformando bacterias. De esta manera, se obtienen finalmente muchas copias del vector con el fragmento a secuenciar, pudiendo además aprovechar la secuencia del vector para diseñar otros cebadores, utilizar endonucleasas específicas, etc. Debemos agradecer la ayuda de Marta Fernández, que está realizando la tesis doctoral en el INDEGSAL con la Dra. Vanesa Robles, por enseñarme a realizar esta técnica.

El vector que utilizamos fue el pUC19 (Vieira y Messing, 1982; amablemente cedido por Antonio Rodríguez, del INBIOTEC), el cual posee una región que codifica una proteína que proporcionará a las bacterias que asimilen el vector resistencia frente a ampicilina (**Fig. 8**). Una vez extraído del gel, introdujimos el fragmento a secuenciar en el vector, y transformamos con él la cepa α de *Escherichia coli*. Para ello, mediante el uso de cloruro de calcio y con un golpe de calor, abrimos poros en las membranas de las bacterias, por los cuales puede penetrar ADN exógeno (en este caso, nuestro vector). Cultivando las bacterias en un medio con ampicilina, solo aquellas bacterias que hayan incorporado el vector formarán colonias. Estas colonias se analizaron mediante PCR para comprobar si el vector contenía el fragmento de interés (mediante la aparición en el gel de electroforesis de la banda correspondiente), y con las colonias positivas se realizó una miniprep. Finalmente se extraía el ADN correspondiente al vector, se linearizaba y se mandaba a secuenciar.

Si bien conseguimos transformar las bacterias y obtener colonias positivas al vector con nuestros fragmentos de ADN de ciervo, los resultados de la secuenciación han sido muy irregulares. Paralelamente, estuvimos mejorando nuestro otro protocolo, consiguiendo mejores resultados con la amplificación mediante PCR, por lo que decidimos abandonar, al menos de momento, la técnica basada en la transformación bacteriana (mucho más larga y más costosa).

Conclusión

Hemos conseguido fragmentos de dos genes que cumplen las condiciones que buscábamos (diseño de cebadores para amplificar un fragmento ~ 1 kb), la β -

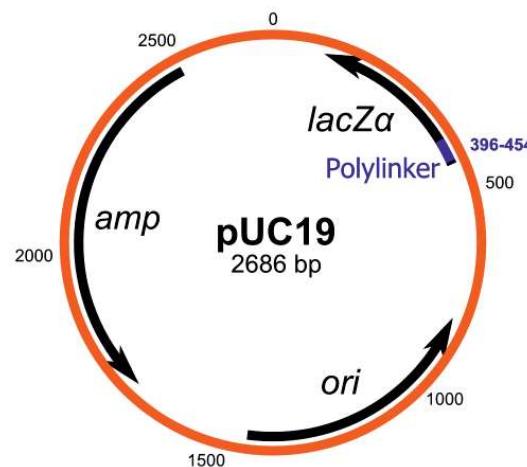
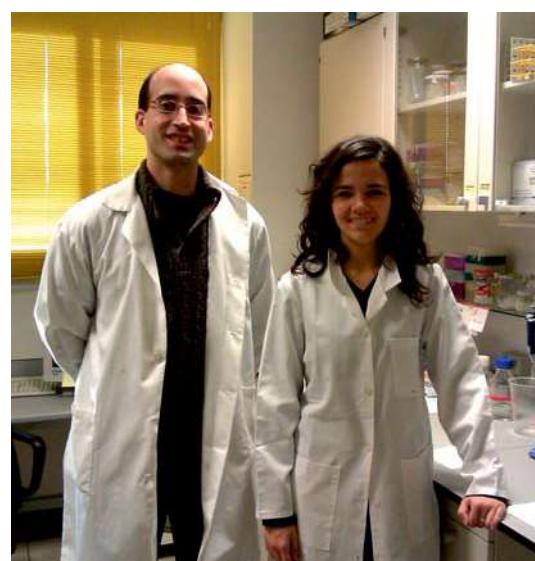


Figura 8. Esquema del vector pUC19, con sus características más relevantes.

globina y el *h19*. Actualmente estamos amplificando varios fragmentos de otros genes, de manera que finalmente tengamos una colección de cebadores para analizar el daño en varias regiones del ADN espermático del ciervo rojo.

Bibliografía

- Bartlett, J.M.S. y Stirling, D. 2003. A short history of the polymerase chain reaction. *PCR Protocols* 226:3–6.
- GE Healthcare. DYEnamic ET Dye Terminator Kit (MegaBACE). 2012. <http://www.gelifesciences.com/webapp/wcs/stores/servlet/productById/en/GELifeSciences-US/25020371> (7 dic. 2012).
- Katherine, L. 2002. Analytical chemistry comes to the rescue: solving the human genome and other biological mysteries with capillary electrophoresis. *The Journal of Young Investigators*, 5 (9), <http://www.jyi.org/features/ft.php?id=460> (8 dic. 2012).
- Huffman, B. 2011. www.ultimateungulate.com. http://www.ultimateungulate.com/whatis_anungulate.html (7 dic. 2012).
- Rothfuss, O., Gasser, T. y Patenye, N. 2010. Analysis of differential DNA damage in the mitochondrial genome employing a semi-long run real-time PCR approach. *Nucleic Acids Research* 38:24.
- Sanger, F. y Coulson, A. R. 1975. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *Journal of Molecular Biology* 94:441–8.
- Vieira, J. y Messing, J. 1982. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene* 19:259–268.
- Wilkinson, D. 2000. Capillary Action. *The Scientist* 14, 10:21.



La autora del trabajo, Dña. Jéssica Alonso, junto con el supervisor del mismo, el Dr. Felipe Martínez Pastor.