



UNIVERSIDAD
POLITECNICA
DE VALENCIA

Marcadores moleculares basados en PCR: Marcadores SSR o STR

(Simple Sequence Repeats or Short Tandem Repeats).

Microsatélites

Apellidos, nombre	Picó Sirvent, María Belén (mpicosi@btc.upv.es) Esteras Gómez, Cristina (criesgo@upvnet.upv.es)
Departamento	Departamento de Biotecnología
Centro	Universitat Politècnica de València



1. Palabras clave

Microsatélite – SSR – polimorfismo – PCR – electroforesis – interpretación

2. Objetivos

Una vez que el alumno haya estudiado con detenimiento este documento y los recursos de apoyo asociados, será capaz de:

1. Definir y explicar cómo surge y en qué consiste el polimorfismo en las secuencias microsatélite.
2. Distinguir y aplicar las diferentes metodologías empleadas para la el uso de estas secuencias como marcadores genéticos.
3. Analizar e interpretar los resultados obtenidos.

3. Introducción

Las secuencias de tipo microsatélite (SSR o STR, Simple Sequence Repeats o Short Tandem Repeats), muy abundantes en los genomas de eucariotas y algunos procariotas, están constituidas por unidades cortas (motivos básicos) de 1 a 6 pares de bases, que se repiten en tándem un elevado número de veces. Cada secuencia SSR se define por el tipo de unidad repetida (lo más frecuente mono, di, tri o tetra, aunque también penta o hexa nucleótidos) y por el sitio que ocupan en el genoma (*locus*). Su frecuencia y tipo de repetición varía en los genomas de distintas especies. Por ejemplo, se sabe que son muy abundantes en peces, insectos himenópteros y mamíferos, y menos en los genomas de aves, en plantas y en lepidópteros (González, 2003).

Se trata de secuencias altamente variables, entre y dentro de individuos. La variación se manifiesta normalmente como diferencias en longitud entre los distintos alelos del mismo *locus*. Estas diferencias en longitud surgen de la existencia de un número diferente de repeticiones del motivo básico en cada caso. Se ha estimado que la tasa de mutación en los microsatélites varía entre 10^{-2} y 10^{-5} por generación y el mecanismo que explica mejor su alto grado de polimorfismo en tamaño es la acumulación de errores producidos por el deslizamiento de la polimerasa durante la replicación del ADN (Ellegren, 2004).

El uso de las secuencias microsatelites como marcadores genéticos se inicia con el desarrollo y generalización de la PCR (Polymerase Chain Reaction) (Mullis et al, 1986). A pesar de que los microsatélites poseen altas tasas de mutación, las regiones flanqueantes están más conservadas y se emplean para la amplificación específica de los alelos de cada *locus*. Al final de los 80 es cuando se publican los primeros trabajos sobre el aislamiento y caracterización de microsatélites (Tautz, 1989). A partir de entonces su uso se ha difundido rápidamente, revolucionando los campos de la biología molecular, la genética cuantitativa y la genética de poblaciones. Los SSRs han sido durante muchos años los marcadores preferidos para múltiples objetivos, genética forense, test de paternidad, análisis poblacionales, estudios de diversidad e identificación varietal, construcción de mapas genéticos y estudios de asociación.



Su detección se basa en la amplificación por PCR de la región que los contiene, empleando cebadores complementarios a las regiones flanqueantes, y la posterior visualización de la diferencia de tamaño de estos amplicones. Sin embargo, su desarrollo requiere información de secuencia y por tanto, hasta hace poco tiempo, no eran marcadores muy frecuentes en especies no modelo. Inicialmente, para conocer la secuencia de las regiones flanqueantes de los SSR se empleaban sondas de los motivos repetidos para cribar genotecas completas, secuenciándose los clones positivos. Esta estrategia de aislamiento era efectiva en taxones en los que la cantidad de secuencias microsatélites era alta o cuando se requería un número reducido de marcadores. Pronto se desarrollaron métodos alternativos para aumentar la eficiencia en la identificación de microsatélites. El método más ampliamente utilizado ha sido la utilización de genotecas genómicas enriquecidas en cada especie, en las que los fragmentos utilizados para construir la genoteca ya son ricos en microsatélites.

Con el avance de las nuevas tecnologías de secuenciación (NGS, Next Generation Sequencing, fundamentalmente las tecnologías, 454 y Illumina (Perez de Castro et al., 2012), que han llevado al abaratamiento del proceso de obtención de secuencias, actualmente se dispone de grandes colecciones de secuencias, tanto genómicas como de ESTs (Expressed Sequence Tags) para muchas especies, lo que ha supuesto un incremento considerable del número de este tipo de marcadores, que son identificados *in silico*, mediante el empleo de distintos algoritmos informáticos, y después validados experimentalmente (Blanca et al., 2011; Zalapa et al., 2012).

Las ventajas que ofrecen los microsatélites se deben, en parte, al empleo de amplificación PCR con cebadores largos, específicos de cada *locus*, ya que el tejido que se utiliza no necesita ser de mucha calidad, e incluso ADN en estado avanzado de degradación es suficiente para ser analizado. Su naturaleza codominante, que permite la distinción de homocigotos y heterocigotos, su amplia distribución en el genoma, su reproducibilidad y su elevada variabilidad (multialélicos) los han convertido en uno de los sistemas de marcadores genéticos más informativos y empleados. A pesar de que los SNPs (Single Nucleotide Polymorphism) están reemplazando a los marcadores tipo SSR en estudios masivos, dado que pueden ser automatizados *via* plataformas de genotipado de alto rendimiento, los microsatélites siguen siendo marcadores muy empleados, especialmente en estudios con un menor número de individuos.

4. Desarrollo

Para entender y aplicar este tipo de marcadores es necesario tener conocimientos previos de Genética, Genómica y técnicas de Genética Molecular (PCR, electroforesis, secuenciación). El contenido que se va a presentar facilitará que el alumno utilice estos conocimientos previos para la interpretación de resultados de este tipo de marcadores.

El contenido se ha estructurado teniendo en cuenta diferentes metodologías para la detección de estos polimorfismos y diferentes casuísticas que el alumno puede encontrar en un experimento real.



Otras páginas web como "IMEx-web: Imperfect Microsatellite Extraction webserver" (http://imex.cdfd.org.in/IMEX/imex_genome.php) etc., permiten la localización de SSR imperfectos.

Escoge un gen de la base de datos NCBI y averigua si presenta motivos SSR.

Entra en OMIM y escoge un gen responsable de una enfermedad humana, causada por una expansión de trinucleótidos, busca la región microsatélite causante de la mutación.

PASO 2. Diseño de cebadores

A pesar de que los microsatélites poseen altas tasas de mutación, las regiones flanqueantes están más conservadas y por ello se usan para la amplificación del locus correspondiente en los distintos individuos de la especie. La transferencia de cebadores entre especies sólo es posible cuando se trabaja con taxones cercanos, pertenecientes al mismo género o a géneros separados recientemente. La transferencia de cebadores entre especies no cercanas puede resultar en la aparición de alelos nulos, debido a sustituciones, deleciones e inserciones en las secuencias que corresponden a los sitios de unión de los cebadores específicos. Esto puede implicar la asignación errónea de individuos heterocigóticos como homocigóticos, y puede suponer a la larga la acumulación de grandes errores en los análisis de datos si los alelos nulos son abundantes.

Una vez detectado el SSR en la secuencia, el siguiente paso es el diseño de cebadores para amplificar la región que lo contiene. Programas libres como "primer 3" (<http://frodo.wi.mit.edu/>) permiten el diseño de cebadores permitiendo la manipulación de ciertos parámetros como: el rango de tamaños del amplicón, el tamaño de los oligos, la temperatura de fusión o melting (T_m), el porcentaje de GC en la secuencia de los oligos y penalización en caso de posibilidad de formación de estructuras secundarias entre los oligos o consigo mismos.

Intenta diseñar cebadores para el SSR anterior.

¿Pueden utilizarse cebadores de loci SSR de una especie modelo a una especie cercana del mismo género? Justifique su respuesta

PASO 3. Visualización de SSRs e interpretación

Tras la amplificación del fragmento que contiene el SSR, hay que proceder a la separación y visualización de estos fragmentos. A continuación se comentarán diversas metodologías de visualización y la interpretación de resultados.

CASO 1. Electroforesis en gel de agarosa

Ciertos alelos SSR con suficiente diferencia de tamaño pueden ser visualizados mediante un simple gel de agarosa con la concentración óptima (2-3%) y posterior tinción con bromuro de etidio (Figura 2).

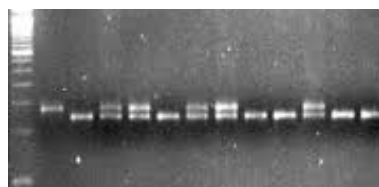


Figura 2. Gel de agarosa que muestra los 3 genotipos existentes para un locus SSR con dos alelos en una población F2. Imagen obtenida de: www.cdfd.org.in/SILKSAT/ssrags.html

CASO 2. Electroforesis en geles de poliacrilamida

En otros casos se opta por geles con mayor capacidad de resolución, como los geles de poliacrilamida. La tinción posterior puede realizarse con sales de plata (Figura 3), pudiendo conservar el gel mediante procedimientos especiales, o bien se pueden emplear cebadores marcados con P^{32} (radiactividad) o con los fluoróforos IRD-700 e IRD-800, para su visualización en analizador de fragmentos LI-COR (Figura 4). Esta última opción resulta la más sencilla si se dispone del equipo, pues se obtiene la imagen directamente sin necesidad de tinciones ni lavados y se pueden cargar al menos 2 muestras por pocillo mediante el empleo de los 2 tipos de fluoróforos (una imagen para cada uno de los dos canales).

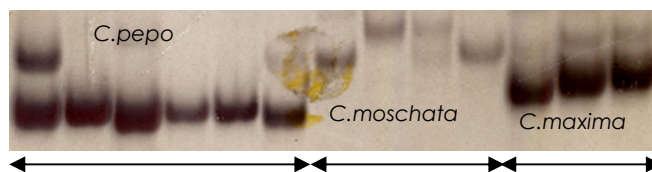


Figura 3. Gel de poliacrilamida teñido con nitrato de plata que muestra los diferentes alelos presentes en tres especies de calabaza (*C. pepo*, *C. moschata* y *C. maxima*).

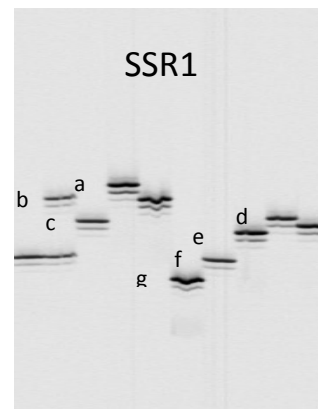


Figura 4. Imagen del canal IRD-700 que muestra los diferentes alelos (a-g) del SSR1 de melón en 10 individuos pertenecientes a los morfotipos más importantes de la especie.

CASO 3. Electroforesis capilar (analizador de fragmentos)

También se pueden emplear analizadores de fragmentos más sofisticados como secuenciadores automáticos, basados en electroforesis capilar. El empleo de diferentes fluorocromos para el marcaje de los cebadores permite la detección por emisión de fluorescencia, al mismo tiempo, de varios productos de PCR, es decir de varios marcadores.

Esto ha hecho que se generalicen las reacciones de amplificación múltiple (Hayden et al, 2008). Este método consiste en amplificar varios *loci* SSR simultáneamente, para lo que se requiere la optimización de las condiciones de PCR, seleccionando cebadores que no interaccionen entre sí, con similar temperatura de unión de cebadores y con alelos en diferentes rangos de tamaño, etc.

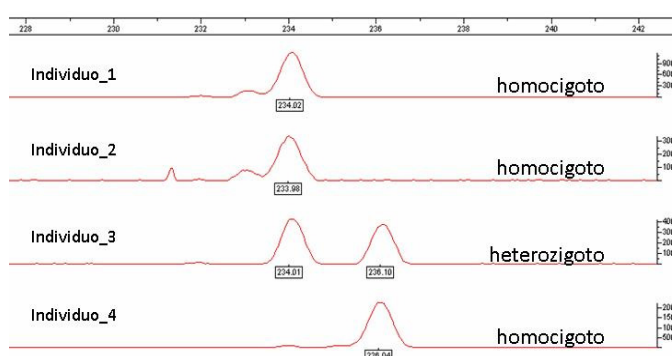


Figura 5. Electroferograma obtenido mediante secuenciador automático donde se muestran 4 individuos portadores de uno o dos alelos de un locus SSR (tamaños 234 y 236 pb).

CASO 4. Detección por HRM (High Resolution Melting)

Recientemente la técnica HRM (High Resolution Melting) está siendo muy utilizada para genotipado de marcadores tipo SNP y SSR (Mackay et al., 2008). El método implica la desnaturalización gradual de los amplicones y la detección de pequeños cambios en la fluorescencia emitida gracias al uso de fluoróforos de unión a DNA de doble cadena. El uso de instrumentos con control de la temperatura muy preciso permite que ligeras variaciones de secuencia entre las muestras puedan ser detectadas mediante las diferencias en las curvas de fusión o melting. La curva de fusión se logra mediante la desnaturalización lenta de los amplicones problema. La fluorescencia es alta al principio, cuando todo el DNA está en forma de doble cadena y el fluoróforo está intercalado, y disminuye durante la desnaturalización, al aumentar la temperatura, debido a que la molécula intercalante (fluoróforo) es liberada pasando a la solución. Esta curva de fusión y T_m (temperatura de fusión: temperatura en la que el 50% de las moléculas de DNA se encuentran como cadena doble) (Figura 6) es característica de cada producto de PCR, variando con la secuencia de bases, el contenido de GC y la longitud del amplicón.

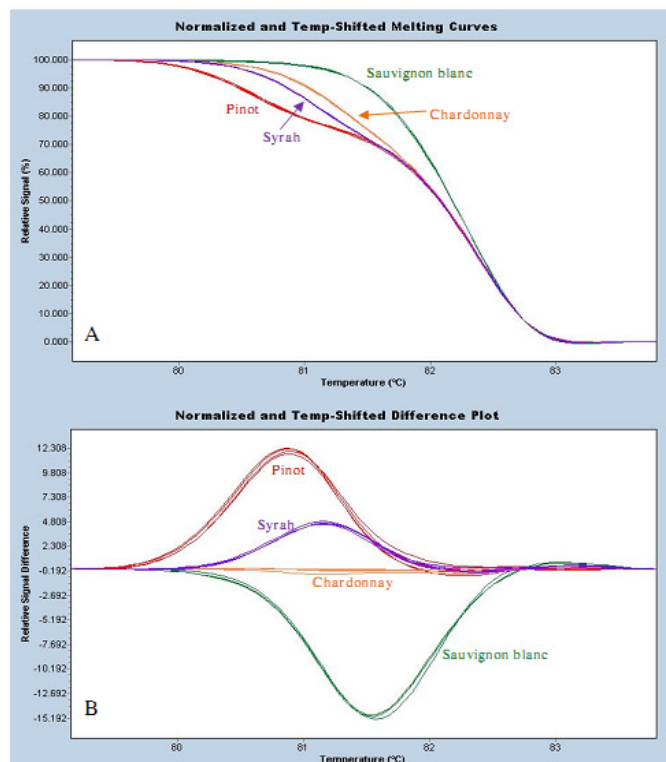


Figura 6. Identificación de microsatélites mediante HRM. Curva de fusión y gráfico diferencial. Se observan las diferencias entre varios genotipos para el SSR ZAG79 (Mackay et al., 2008).

Interpretación

1. Dominancia-codominancia

Aunque en la mayoría de los casos los marcadores SSR se analizan como codominantes y los diferentes alelos se distinguen por su tamaño, en algunos se interpretan como dominantes. Esto ocurre cuando los



cebadores no amplifican en algún genotipo, dando lugar a un alelo nulo, debido a que la secuencia de unión al cebador no se ha conservado en estos genotipos. Normalmente se trata de estudios en los que se han empleado marcadores diseñados para otras especies o subespecies. En general la transferibilidad de los SSR no es buena. En estos casos también suelen darse ejemplos de amplificación preferencial, en la que en unos genotipos la amplificación se produce, pero no de forma eficiente debido a alguna mutación en la secuencia de unión al cebador, de forma que cuando están los dos alelos en un individuo prácticamente sólo amplifica uno, haciendo pensar que es homocigoto para el alelo preferencial y no heterocigoto.

2. Alelos de tamaños similares y homoplasia

En algunos casos, y dependiendo del sistema de separación de fragmentos empleado o detección, alelos de tamaños similares no se resuelven bien interpretándose como un único alelo. Esto ocurre normalmente cuando se emplean geles de agarosa, en estos casos técnicas de mayor resolución como la electroforesis capilar o HRM serían más adecuadas.

Otro caso es el de la homoplasia, productos de PCR de igual tamaño, y que por tanto consideramos como un mismo alelo, pero que en realidad no son idénticos y presentan cambios de estructura, presencia de inserciones o deleciones, cambios de bases o variaciones en la región flanqueante. En algunos casos podrían llegar a distinguirse mediante HRM.

3. SSR multiloci

Debido a procesos de duplicación en el genoma, en algunas ocasiones los cebadores presentan secuencias complementarias a varias regiones, de forma que se visualizan varios *loci* amplificados, que a su vez pueden presentar varios alelos debido a su diferente evolución (Figura 7). Esto dificulta la interpretación de resultados.

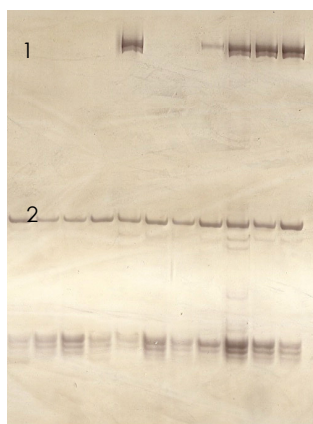


Figura 7. Gel de acrilamida en el que se observan 2 *loci* (1 y 2) fruto de la amplificación con una misma pareja de cebadores. El *locus* 1 presenta polimorfismo con un alelo nulo (ver punto 1 de la Interpretación), mientras que el *locus* 2 es monomórfico para los genotipos probados.

4. Artefactos

En muchos casos la interpretación de resultados también se ve dificultada por la aparición de artefactos y bandas extra generados en el proceso de amplificación o durante la electroforesis.



El efecto "sonrisa" en los geles normalmente se debe a diferencias en la concentración del tampón que constituye el gel y el de los electrodos, o a diferencias de temperatura entre los extremos y la región central del gel mientras corren las muestras (Figura 8).

La visualización de los llamados "espejos", bandas extra, (Figura 9) puede tener varias causas. La primera es el deslizamiento de la cadena sintetizada en la amplificación que conduce a la obtención de productos de PCR de diferente tamaño, pero también puede deberse a la estructura secundaria de DNA que retarde la migración en el gel o el alineamiento incorrecto de la secuencias con motivos repetidos durante la renaturalización de la doble cadena durante la amplificación.

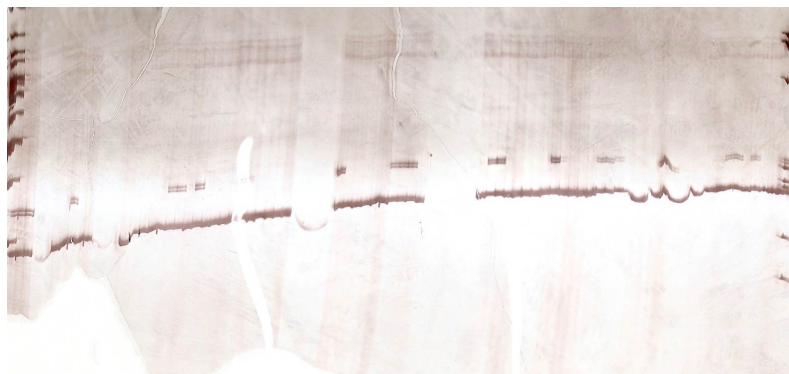


Figura 8. Gel de acrilamida en el que se observan las muestras con efecto de sonrisa.

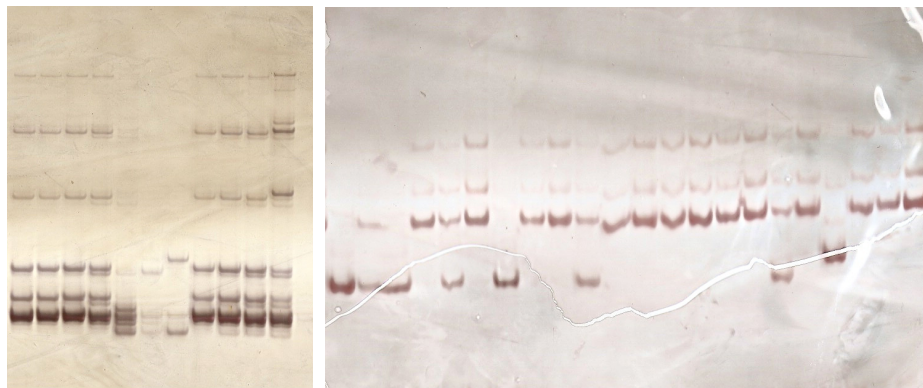
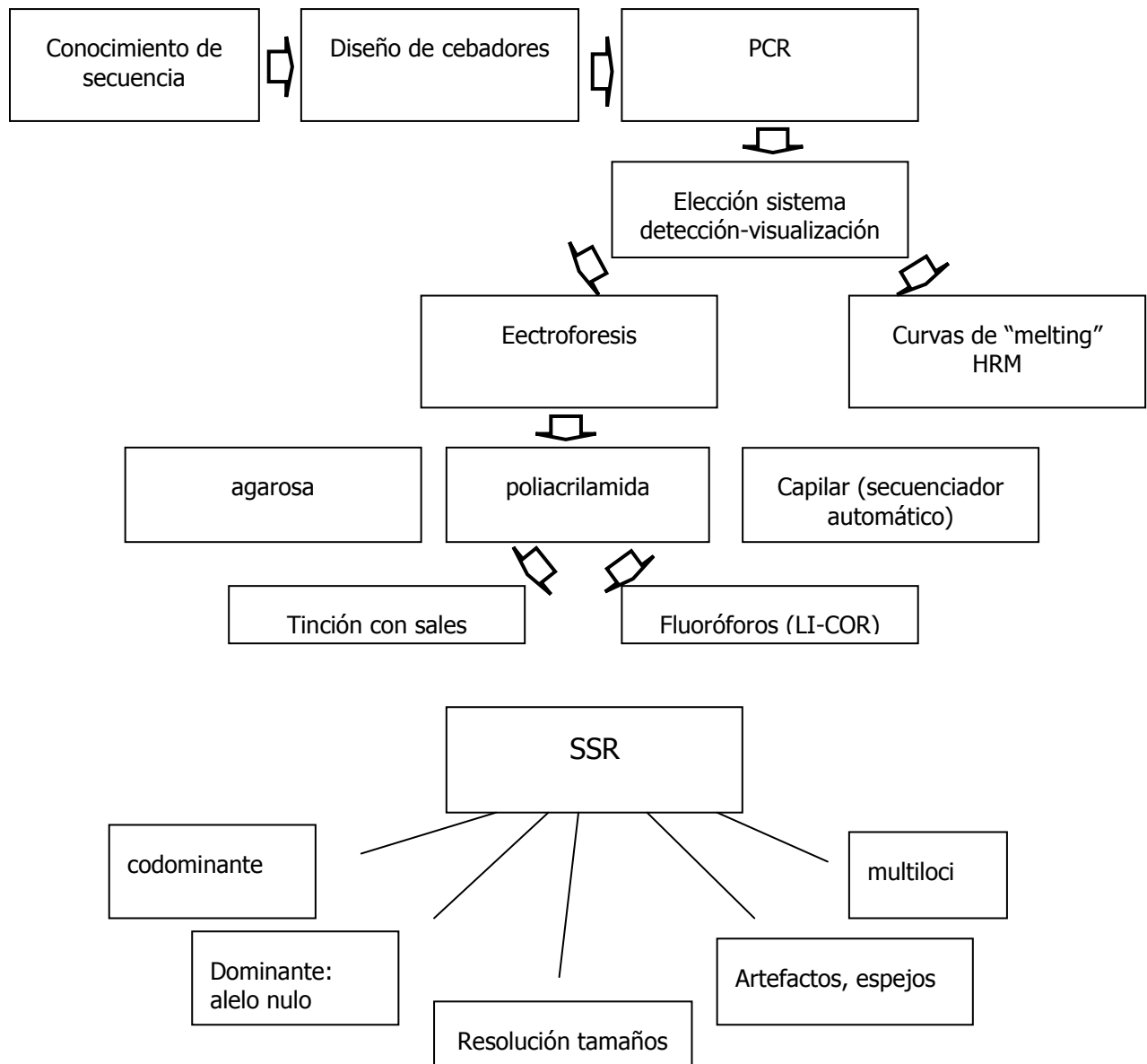


Figura 9. Geles de acrilamida en los que se observan espejos para 2 SSRs evaluados en dos poblaciones diferentes.



5. Resumen-Esquema

REPASA LO APRENDIDO CON EL ESQUEMA SIGUENTE





6. Bibliografía

6.1. Citas y webs citadas en el texto:

-Blanca J, Cañizares J, Ziarolo P, Esteras C, Mir G, Nuez F, Garcia-Mas J, Picó B. 2011. Melon transcriptome characterization. SSRs and SNPs discovery for high throughput genotyping across the species. *The Plant Genome* 4(2):118-131.

-Ellegren H. 2004. Microsatellites: simple sequences with complex evolution. *Nat Rev Genet* 5(6):435-445.

-González EG. 2003. Microsatélites: sus aplicaciones en la conservación de la biodiversidad. *Graellsia* 59(2-3):377-388.

-Hayden MJ, Nguyen TM, Waterman A, Chalmers KJ. 2008. Multiplex ready PCR: a new method for multiplexed SSR and SNP genotyping. *BMC Genomics* 9:80.

-Mackay JF, Wright CD, Bonfiglioli RG. 2008. A new approach to varietal identification in plants by microsatellite high resolution melting analysis: application to the verification of grapevine and olive cultivars. *Plant Methods* 4:8.

-Mullis K, Faloona F, Scharf S, Saiki R, Horn G, Erlich H. 1986. Specific enzymatic amplification of DNA *in vitro*: The polymerase chain reaction. *Cold Spring Harbor Symp Quant Biol* 51:263-273.

-Pérez-de-Castro A, Vilanova S, Cañizares J, Pascual L, Blanca JM, Díez MJ, Prohens J, Pico B. 2012. Application of Genomic Tools in Plant Breeding. *Current Genomics* 13(3):179-195 (17).

-Tautz D. 1989. Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Research* 17: 6463-6471.

-Zalapa JE, Cuevas H, Zhu HY, Steffan S, Senalik D, Zeldin E, McCown B, Harbut R, Simon P. 2012. Using Next-Generation Sequencing approaches to isolate Simple Sequence Repeat (SSR) loci in the plant sciences. *American Journal of Botany* 99(2):193-208. DOI: 10.3732/ajb.1100394.

-http://imex.cdfd.org.in/IMEX/imex_genome.php

-<http://frodo.wi.mit.edu/>

-<http://www.gramene.org/db/markers/ssrtool>

6.2. No citados en texto:

-Romero C. 2009. Identificación de marcadores microsatélites. Análisis e interpretación de resultados. En "Prácticas de Biodiversidad y Marcadores Moleculares". Picó B, Díez MJ (eds.) Editorial Universidad Politécnica de Valencia.