



UNIVERSIDAD
POLITECNICA
DE VALENCIA

Nuevas Estrategias de Análisis de la Diversidad Genética Natural: *Identificación de Variantes Alélicas en Genes de Interés Mediante Ecotilling*

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



1. Resumen de las ideas clave

Los recientes avances en Genética y Genómica están proporcionando una gran cantidad de información sobre la secuencia y la función de los genes, lo que supone la base para el desarrollo de nuevas estrategias de estudio de la diversidad. En este artículo se describe una nueva metodología de uso creciente, utilizada para la detección de polimorfismos genéticos en poblaciones naturales: EcoTILLING. Esta herramienta se ha empleado con éxito en diversas especies, no sólo para estudiar la variabilidad existente, sino con fines de mejora vegetal o animal. En este artículo se incluye una breve revisión teórica y una explicación práctica detallada. Con la explicación planteada se pretende facilitar el aprendizaje de esta nueva estrategia de análisis de la diversidad al alumno de Ciencias de la vida (Agronomía, Biología, Veterinaria, Medio ambiente, Biotecnología..), tanto a nivel teórico como práctico.

2. Objetivos

Una vez que el alumno haya leído con detenimiento este documento, será capaz de:

1. Describir y explicar los fundamentos básicos de la técnica de EcoTILLING.
2. Dar ejemplos y aplicaciones de esta técnica en diversos campos.
3. Diseñar y desarrollar-aplicar un experimento de este tipo.
4. Analizar e interpretar resultados obtenidos.

3. Introducción

La técnica a estudiar, EcoTILLING, es una adaptación de la herramienta de genética reversa y genómica funcional denominada TILLING (Targeting Induced Local Lesions in Genomes).

El TILLING se inicia generando una colección de individuos mutantes con etilmetanosulfonato (EMS) u otros mutágenos, que posteriormente se analizan de forma sistemática ("screening"), para identificar mutaciones puntuales en genes-diana seleccionados. La identificación de mutaciones no requiere la secuenciación de cada genotipo, sino que se lleva a cabo a gran escala obteniendo moléculas homodúplex o heterodúplex a partir de un proceso de desnaturalización-renaturalización mezclando parejas de amplicones. Para un amplicón concreto (correspondiente a todo el gen diana o a un fragmento del mismo), cada pareja incluye un amplicón procedente de un individuo cuyo genotipo se desea conocer y el mismo amplicón procedente del genotipo de referencia (genotipo empleado para obtener la colección de mutantes; suele elegirse una variedad elite que pretende mejorarse para aspectos concretos). Los fragmentos resultantes de la renaturalización que son de tipo heterodúplex (emparejamiento entre dos moléculas monocordales que no son totalmente complementarias) se distinguen de los homodúplex (emparejamiento entre moléculas complementarias en toda su longitud) mediante digestión con endonucleasas que reconocen los emparejamientos anómalos entre bases ("mismatch"). Estos "mismatches" puntuales ocurren cuando entre los dos amplicones existen polimorfismos de tipo SNP o indel (Single nucleotide



Polymorphism o pequeñas inserciones-delecciones). Una vez cribada toda la población para cada amplicón, aquellos individuos que presenten mutaciones con respecto al control se seleccionan, se fija la mutación (que está en heterocigosis en el individuo original), y se estudia su fenotipo. La función del gen se esclarece relacionando la mutación identificada con el fenotipo mostrado por el individuo en cuestión. Esta técnica se empleó por primera vez en plantas en la especie modelo *Arabidopsis thaliana*¹ y presenta ciertas ventajas respecto a otras aproximaciones de genética reversa.

El EcoTILLING utiliza un procedimiento muy similar para la identificación de mutaciones en genes concretos, pero empleando poblaciones naturales. En este caso, por tanto, no se inducen mutaciones, sino que se parte de una colección de individuos (colección de germoplasma) lo más diversa posible y representativa de la variación natural existente en la especie. El genotipo de referencia suele ser también una variedad de élite, una línea mejorada de interés o uno de los parentales del mapa genético de la especie, frente al cual se quieren estudiar las variaciones existentes en la naturaleza. En especies en las que la mutagénesis química no es eficiente este método es una alternativa para estudiar la función génica, aunque en este caso es necesario tener en cuenta la disparidad entre los fondos genéticos de los genotipos analizados para estudiar la correspondencia mutación/fenotipo. El EcoTILLING permite evaluar muchas muestras rápidamente en busca de SNPs e indels, reduciendo el tiempo y coste del descubrimiento de SNPs frente a otras técnicas. El primer estudio publicado en el que se aplicó la técnica EcoTILLING también fue llevado a cabo en *A. thaliana*², permitiendo la identificación de 55 haplotipos (combinaciones de mutaciones) en los 5 genes analizados. Posteriormente, se ha puesto a punto en cultivos de interés como *Populus trichocarpa*³, cebada (*Hordeum vulgare*)⁴ y melón (*Cucumis melo*)⁵, para examinar genes de resistencia a hongos y a virus, y en otras especies como *Vigna radiata*, para realizar estudios de diversidad en colecciones de germoplasma. A continuación se explican los pasos del procedimiento con detalle. En especies de valor agronómico tiene especial importancia para la mejora, por ejemplo, en el caso del melón se confirmó la presencia de una mutación no silenciosa leucina-histidina en el factor de inicio de la traducción eIF4E, relacionada con la resistencia al virus del cribado necrótico del melón, MNSV.

4. Desarrollo

Los avances en genética y genómica de los últimos años están siendo muy rápidos. Para entender y aplicar la estrategia de EcoTilling 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 el desarrollo de una nueva estrategia de análisis de la variación. Aplicando dicha estrategia podrán descubrirse nuevas mutaciones en plantas, animales y humanos, relacionadas con características de interés en los sectores de la agronomía, medio ambiente, la producción animal o la medicina. Son impactantes los resultados obtenidos hasta la fecha: cereales, frutas o verduras resistentes a patógenos o de más calidad, o mutaciones poco frecuentes asociadas a enfermedades en animales y humanos⁶.

El contenido se ha estructurado utilizando un caso práctico de EcoTilling desarrollado por el grupo de Mejora Genética de Cucurbitáceas de la UPV. El objetivo de este estudio es analizar la variabilidad natural en el gen CmACO1 (ACC oxidasa1 de *Cucumis melo* L.), gen que forma parte de la ruta enzimática de producción de etileno, estando implicado, por tanto, en el proceso de maduración y la calidad del fruto. Se describen los fundamentos básicos de cada paso, planteándose una cuestión relevante al final de cada uno para validar lo aprendido. Al final se presenta un esquema resumen que permite confirmar el aprendizaje de la globalidad del proceso. El alumno se puede apoyar durante el proceso de aprendizaje en los trabajos, software y páginas web citadas durante el desarrollo del mismo.

PASO 1. Selección de una colección de germoplasma representativa de la variación en la especie

Uno de los aspectos clave para la aplicación con éxito del EcoTILLING es la selección del germoplasma a analizar. Es recomendable conocer la distribución de la variación en la especie de trabajo (domesticación, centros de origen y diversificación, etc.). Una estrategia adecuada es emplear colecciones nucleares (que representen en un número mínimo de individuos la mayor parte de la variación de la especie). En cualquier caso, es conveniente realizar un análisis de la población con marcadores neutros para conocer su estructura.

Ejemplo: El melón es una especie muy variable, especialmente en cuanto a caracteres de fruto. La colección empleada consiste en 212 entradas, incluyendo representantes de las 2 subespecies y de los 16 morfotipos ⁷, procedentes de 54 países, preferentemente de los centros de origen y diversificación (África, Asia, región mediterránea). Se encuentran tipos silvestres, naturalizados, razas locales, líneas de mejora y cultivares comerciales (Figura 1). El genotipo de referencia empleado, es decir, aquel frente al cual se quieren identificar las mutaciones, es un tipo Piel de sapo (PS), la clase comercial más importante a nivel internacional de los melones de tipo *inodorus*.



Figura 1. Diversidad geográfica y morfológica de la colección de germoplasma de melón empleada en EcoTILLING. La fotografía del extremo izquierda corresponde con el tipo Piel de sapo, genotipo de referencia.

¿Cuáles son los aspectos clave para la selección de la población natural?

PASO 2. Preparación del DNA

En primer lugar, se debe realizar una extracción de DNA genómico de calidad para el posterior "screening". Pueden utilizarse mezclas de DNA para acelerar

el proceso de identificación. Cuando se utiliza esta estrategia es necesario, en el caso de encontrarse una mutación, repetir el análisis en las muestras individuales para identificar el individuo mutante.

Ejemplo: El melón es una especie parcialmente alógama, por lo que una "accesión" (muestra de semillas mantenida en un banco de germoplasma procedente de una determinada colecta) no tiene por qué ser una colección de individuos homocigotos. Por ello, cada "accesión" debe estar representada por varios individuos. Para optimizar el análisis es conveniente realizar "pools" (mezclas equimolares de DNA de distintos individuos), tras la normalización. De esta forma, en un pocillo de una placa se tiene la mezcla del DNA de, por ejemplo, 8 individuos, teniendo solo 1 placa en vez de 8 (Figura 2). Esto ahorra tiempo y dinero cuando se quieren analizar cientos de muestras.

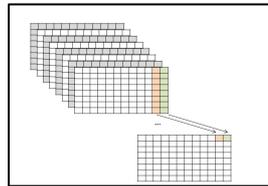


Figura 2. Generación de "pools" de DNA procedente de 8 individuos ("eightfold pools").

¿Cómo se puede maximizar el número de genotipos analizados reduciendo el tiempo invertido?

PASO 3. Amplificación del gen o fragmento del gen de interés

Para llevar a cabo un experimento de EcoTILLING es esencial disponer de secuencias genómicas de genes candidatos involucrados en procesos metabólicos de interés. Esta situación es hoy en día una realidad, tanto para especies modelo como para especies menores, gracias a las mejoras de la metodología de Sanger, pero sobre todo a las nuevas técnicas de secuenciación masiva que están ampliando de forma impresionante el rango de especies para las cuales se dispone de colecciones de secuencias.

A partir de la secuencia del gen diana elegido (disponible en base de datos, <http://www.ncbi.nlm.nih.gov/genbank/>⁸) se diseñan cebadores específicos para amplificar la región deseada (Figura 3). Es conveniente que el gen diana tenga sólo una copia en el genoma, pues facilita la amplificación e interpretación de los resultados. También es necesario tener en cuenta que en los últimos 100-200 pb al final del amplicón no se detectan bien los SNPs, por lo que muchas veces se han de estudiar varios amplicones solapantes. Los genes largos y con muchos intrones se suelen dividir en amplicones de entre 1-1.5 kb (tamaño adecuado para la amplificación, para una adecuada renaturalización, y para la optimización de la detección de mutaciones).

Ejemplo: En el caso del gen CmACO1, con un único amplicón se permite la identificación de mutaciones en prácticamente toda la secuencia del gen, a excepción de los extremos de los exones 1 y 4. La amplificación se realiza con cebadores marcados con dos fluoróforos (IRD-700 Y IRD-800), lo que permite la posterior visualización de los fragmentos digeridos y el análisis de su tamaño.

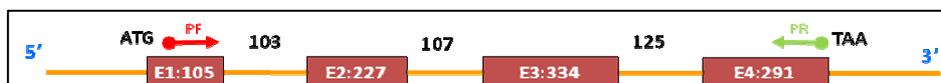


Figura 3. Estructura del gen CmACO1. Los exones son las cajas y los intrones las líneas, indicándose el tamaño en pb. Los cebadores empleados se localizan en el primer y último exón, estando marcado el cebador directo (PF) con IRD-700 y el reverso (PR) con IRD-800 en el extremo 5'.



En realidad, suelen emplearse los cebadores marcados para amplificar el genotipo de referencia y no marcados en el resto de genotipos. De esta forma, sólo se visualizan los fragmentos con extremos del genotipo de referencia. El genotipo de referencia es una línea pura "*homocigota para todos sus genes*", así todos los fragmentos resultantes de la digestión procederán de moléculas de tipo heterodúplex entre el genotipo de referencia y cada uno de los genotipos evaluados. Por otro lado, como no puede asegurarse la homocigosis de todos los genotipos evaluados, al no tratarse en su mayoría de líneas fijadas, sino de tipos tradicionales o silvestres, conviene comprobar si son heterocigotos. Por ello, habría que amplificar con cebadores marcados aquellos genotipos que presentaran una mutación respecto al de referencia y realizar la renaturalización "*consigo mismo*".

¿Por qué deben emplearse varios amplicones por gen para cribar toda su secuencia? Elige de la base de datos del Genbank un gen que te interese y trata de diseñar unos amplicones para su análisis por EcoTILLING.

PASO 4. Obtención de heterodúplex, digestión con Cel I y precipitación

Una vez cuantificados los productos de la amplificación, se mezcla la misma cantidad del amplicón de cada genotipo con el amplicón de referencia. Esta mezcla se desnaturalizará, renaturalizándose después lentamente, para que no sólo se formen homodúplex (XX o PSPS), sino también heterodúplex o moléculas híbridas (XPS) (Figura 4).

Cuando existe un emparejamiento anómalo puntual ("*mismatch*") la enzima endonucleasa corta en ese punto la cadena, resultando en fragmentos de distinto tamaño según la posición del polimorfismo. Existen diversas nucleasas que cortan sólo una de las hebras cuando detectan un "*mismatch*" y, debido a que no son eficientes al 100%, permiten detectar más de una mutación en un amplicón. Entre las más empleadas están S1 (incluye Endo I) y P1 de origen fúngico, y MBE y Cel I de origen vegetal. La nucleasa Cel I se extrae del apio, habiéndose demostrado su eficacia tanto purificada como en extracto crudo⁹. Cada nucleasa tiene un pH y una temperatura de trabajo.

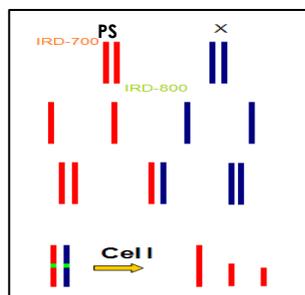


Figura 4. Esquema de la detección de mutaciones por EcoTILLING. Se muestra la mezcla de los pPCR de una entrada cualquiera X con el genotipo de referencia PS. Se desnaturaliza y renaturaliza de forma que algunas hebras X se emparejan con las de PS (heterodúplex). Posteriormente, al tratar con una endonucleasa que reconoce emparejamientos anómalos (en verde) se obtienen fragmentos de distinto tamaño según la posición del punto de corte.

Tras la digestión se procede a la precipitación de los productos digeridos para eliminar los restos de la PCR y digestión. Es muy importante en cada experimento incluir una serie de controles: el genotipo de referencia desnaturalizado y renaturalizado consigo mismo (para comprobar que es homocigoto), el genotipo de referencia no desnaturalizado-renaturalizado pero sí digerido con Cel I (para comprobar que la endonucleasa no corta



cuando no hay moléculas híbridas) y el producto de PCR de este genotipo de referencia (cualquier banda presente se tomará como ruido de fondo).

¿Qué procedimiento se utiliza para detectar polimorfismos puntuales entre amplicones tras el proceso de desnaturalización y renaturalización? ¿Qué controles deben emplearse en un procedimiento de EcoTILLING?

PASO 5. Electroforesis y visualización en el analizador de fragmentos LI-COR.

Los fragmentos generados en la digestión se separan en un gel de acrilamida en condiciones desnaturalizantes. La electroforesis se realiza en un analizador de fragmentos LI-COR para obtener una elevada resolución. El tamaño de los fragmentos nos indicará la posición en el amplicón de las mutaciones respecto al genotipo de referencia. Así, las bandas visualizadas corresponderán a las hebras de PS (son las únicas marcadas con fluorescencia IRD) cortadas en diferentes puntos del amplicón. Cuando existen muchos SNPs en la secuencia existe el riesgo de no visualizar aquéllos más alejados del extremo marcado (5'). Se obtiene una imagen para cada uno de los 2 canales (marcaje IRD-700 e IRD-800 respectivamente), de modo que en realidad se visualizan los 2 extremos del amplicón (5' y 3') y se pueden comprobar las bandas procedentes de un mismo corte con tamaños complementarios que suman el tamaño del amplicón.

Ejemplo: En el amplicón ACO1 (1148 pb) encontramos una banda de 365 pb en el canal-700, entonces debe existir una banda en el canal-800 cuyo tamaño sea 783 pb (Figura 5). Además de cargar las muestras con el control PS, se ha de incluir el control consigo mismo (PSPS), desnaturalizado y no desnaturalizado (pero sí digerido con Cel I), y el pPCR de PS. Se carga también en los extremos un marcador de pesos para saber los tamaños de las bandas.

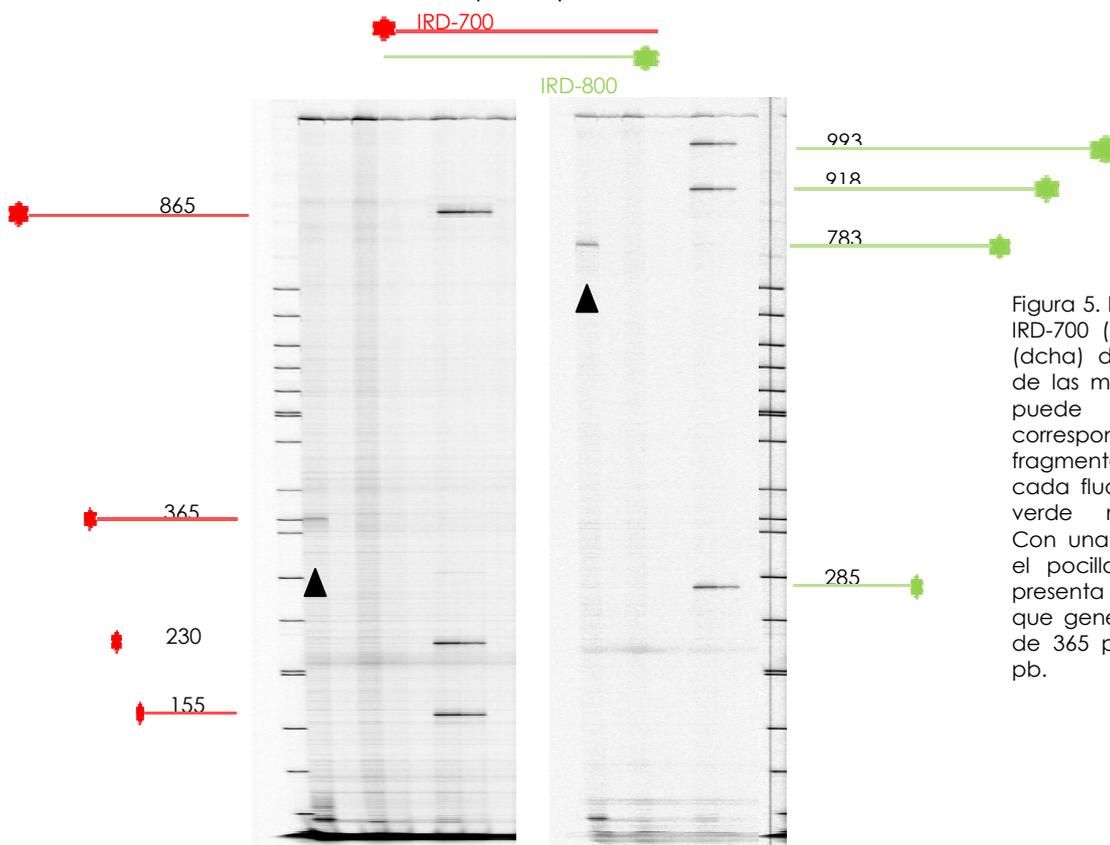


Figura 5. Imagen del canal IRD-700 (izqda.) e IRD-800 (dcha) del LI-COR con 8 de las muestras donde se puede observar la correspondencia de los fragmentos marcados con cada fluoróforo (en rojo y verde respectivamente). Con una flecha se indica el pocillo cuyo genotipo presenta una mutación tal que genera un fragmento de 365 pb y otro de 783 pb.

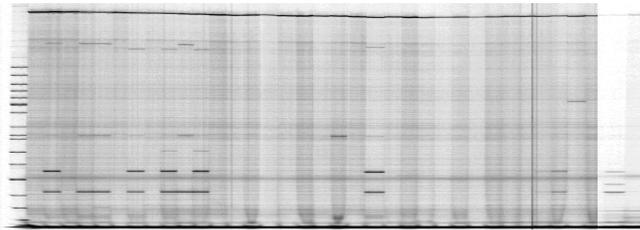


Figura 6. Imagen del canal-IRD700 para el amplicón ACO1. Se muestran las carreras correspondientes a 44 genotipos de melón.

Una vez se tienen los diferentes gels con todos los genotipos (o pools), se analizan en busca de bandas consistentes (mutaciones), agrupando los diferentes genotipos en haplotipos según el patrón de bandas visualizado. En el ejemplo del gel anterior (Figura 6).

¿En qué posición se encuentran las mutaciones con respecto al genotipo de referencia? ¿Cuántos haplotipos identificas?

PASO 6. Análisis e Interpretación de resultados

Tras entender el fundamento de cada uno de los pasos anteriores, estamos en situación de interpretar los resultados obtenidos. Podemos comprobar nuestro aprendizaje revisando la respuesta a la pregunta planteada en el punto anterior. **¿Cuántos haplotipos identificas? Respuesta:** Podemos agrupar los 44 genotipos en 8 haplotipos (H0-H7). El haplotipo H0 engloba a todos los genotipos que no presentan ninguna mutación respecto al control PS. El resto presentan las mutaciones que se indican en el esquema de la Figura 7. Se indica, además, la posición de los SNPs localizados en exones (desde ATG).

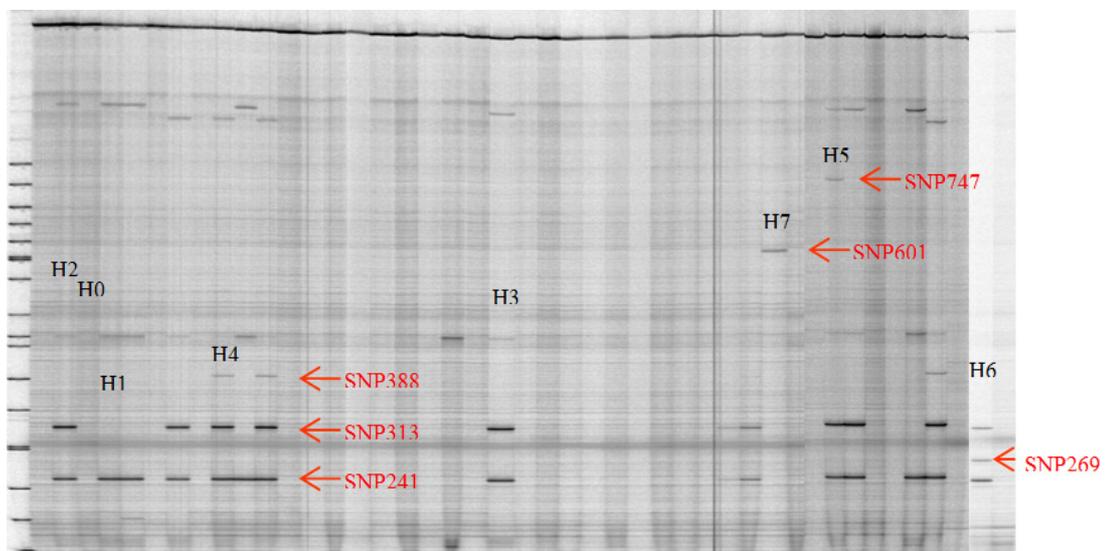


Figura 7. Haplotipos encontrados y localización de SNPs para el amplicón ACO1.

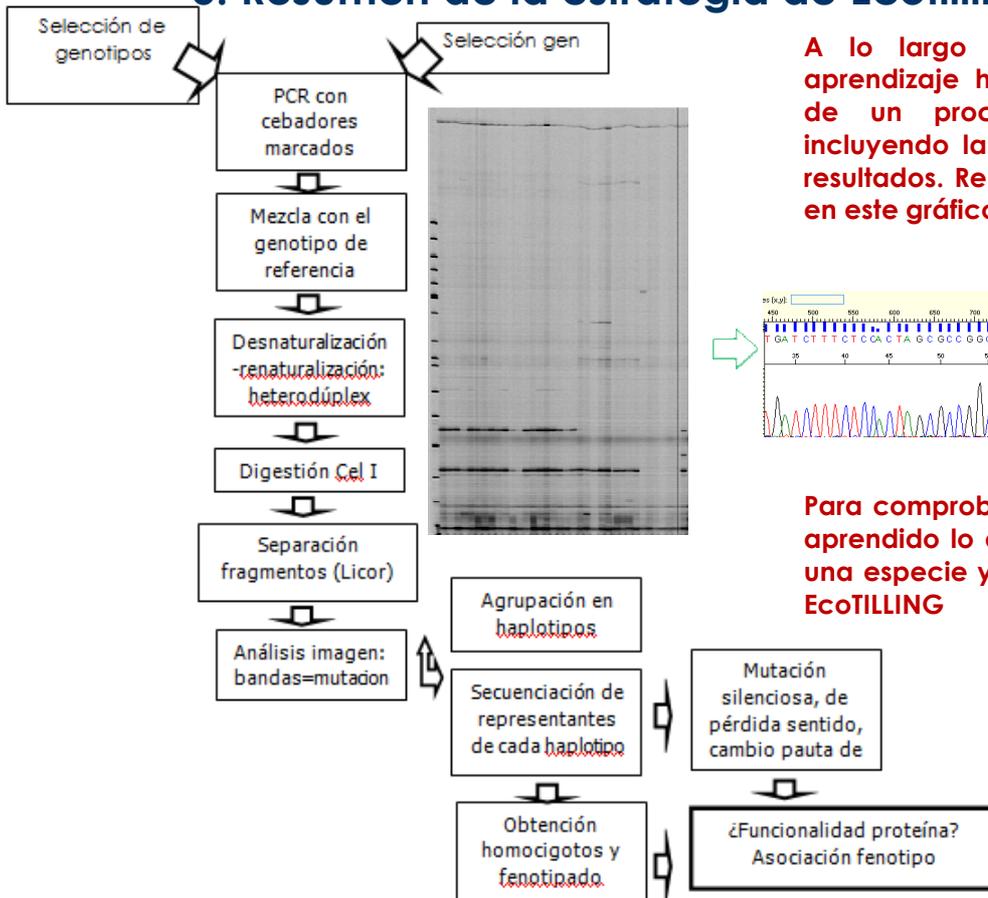
Ventaja: Hemos identificado todas las variantes alélicas naturales de este gen en la especie. Este análisis puede realizarse de forma rápida en un número muy amplio de individuos y clasificarlos en los distintos haplotipos. Sólo es necesario confirmar las mutaciones por secuenciación, para asociar la banda visualizada a un SNP o indel concreto, en un individuo de cada haplotipo.



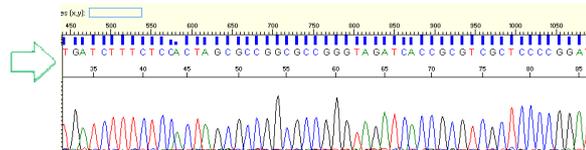
Las mutaciones comprobadas por secuenciación se analizan utilizando varios programas disponibles *on line* (clustalw2:<http://www.ebi.ac.uk/Tools/clustalw2/index.html>; blast2seq:<http://blast.ncbi.nlm.nih.gov/>). Tras conocer el tipo de cambio y la posición exacta de cada mutación, se estudia cómo se vería afectada la proteína. Las mutaciones en regiones intrónicas o del promotor pueden suponer diferencias fenotípicas al afectar el nivel de expresión génica. Las mutaciones en las regiones codificantes (exones) se analizan en función de si conducen o no a un cambio aminoacídico. Incluso en el caso de que ocurra un cambio el efecto puede ser muy variable dependiendo de la posición del aa y del tipo de cambio. Hay mutaciones que truncan totalmente la proteína, por ejemplo si la mutación supone un codón de parada. A veces el cambio de aa no afecta a la funcionalidad de la proteína, debido a que es un aa del mismo tipo o a que esa región proteica no contiene el centro activo. Existen programas online (SIFT, http://blocks.fhcrc.org/sift/SIFT_seq_submit2.html) que predicen cómo se puede ver afectada la proteína. La determinación se lleva a cabo mediante alineamientos con genes ortólogos en las especies en las que se tiene información. Mutaciones en las zonas más conservadas, menos tolerantes a los cambios de secuencia, es probable que conduzcan a drásticos efectos por ser importantes para la actividad proteica. Los indels conllevan cambio de pauta de lectura, por lo que también son especialmente graves las consecuencias.

Ejemplo: En el caso en estudio, sólo 3 mutaciones suponen un cambio de aa, siendo sólo 2 las que según el programa SIFT pueden afectar a la funcionalidad de la ACCoxidasa (cambio aminoacídico no tolerado). El último paso sería obtener individuos homocigotos para la mutación en cuestión realizando autofecundaciones (en el caso de ser heterocigotos) y realizar un fenotipado de éstos para asociar mutación–fenotipo.

5. Resumen de la estrategia de EcoTilling



A lo largo de este objeto de aprendizaje hemos visto los pasos de un proceso de EcoTILLING incluyendo la interpretación de los resultados. Repasamos la estructura en este gráfico.



Para comprobar qué realmente has aprendido lo explicado, selecciona una especie y aplica el proceso de EcoTILLING



6. Bibliografía

6.1. Citas en texto:

- 1 Colbert T, Till BJ, Tompa R, Reynolds S, Steine MN, Yeung AT, McCallum CM, Comai L, Henikoff S. 2001. High-throughput screening for induced point mutations. *Plant Physiology*, 126: 480-484.
- 2 Comai L, Young K, Till BJ, Reynolds SH, Greene EA, Codomo CA, Enns LC, Johnson JE, Burtner C, Odden AR, Henikoff S. 2004. Efficient discovery of DNA polymorphisms in natural populations by EcoTilling. *Plant J.* 37: 778-786.
- 3 Gilchrist EJ, Haughn GW, Ying CC, Otto SP, Zhuang J, Cheung D, Hamberger B, Aboutorabi F, Kalynyak T, Johnson L, Bohlmann J, Ellis BE, Douglas CJ, Cronk QCB. 2006. Use of EcoTilling as an efficient SNP discovery tool to survey genetic variation in wild populations of *Populus trichocarpa*. *Mol. Ecol.* 15: 1367-1378.
- 4 Mejlhede N, Kyjovska Z, Backes G, Burhenne K, Rasmussen SK, Jahoor A. 2006. EcoTILLING for the identification of allelic variation in the powdery mildew resistance genes *mlo* and *Mla* of barley. *Plant Breed.* 125: 461-467.
- 5 Nieto C, Piron F, Dalmais M, Marco CF, Moriones E, Gómez-Guillamon ML, Truniger V, Gomez P, Garcia-Mas J, Aranda MA, Bendahmane A. 2007. EcoTILLING for the identification of allelic variants of melon *eIF4E*, a factor that controls virus susceptibility. *BMC Plant Biol.* 7:34.
- 6 Till BJ, Zerr T, Bowers E, Greene EA, Comai L, Henikoff S. 2006. High-throughput discovery of rare human nucleotide polymorphisms by EcoTilling. *Nucleic Acids Res.* 34: e99.
- 7 Pitrat M, Hanelt P, Hammer K. 2000. Some comments on infraspecific classification of cultivar of melon. *Acta Hortic.* 510, 29-36.
- 8 Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL. 2008. *Nucleic Acids Research*, 2008 Jan;36(Database issue):D25-30.
- 9 Till BJ, Burtner C, Comai L, Henikoff S. 2004. Mismatch cleavage by single-strand specific nucleases. *Nucleic Acids Res.* 32, no 8: 2632-2641.

6.2. No citados en texto:

Libros: Meksem K, Liu S, Liu XH, Jamai A, Mitchum MG, Bendahmane A, El-Mellouki T. 2008. TILLING: A reverse genetics and a functional genomics tool in soybean. En: Kahl G, Meksem K. 2008. *The handbook of plant functional genomics: concepts and protocols*. WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

Artículos: Barkley NA, Wang ML. 2008. Application of TILLING and EcoTILLING as Reverse Genetic Approaches to Elucidate the Function of Genes in Plants and Animals. *Curr Genomics.* Jun;9(4):212-26.

Comunicaciones en congresos: Esteras C, Pascual L, Saladie M, Dogimont C, Garcia-Mas J, Nuez F, Picó B. 2009. Use of EcoTilling to identify natural allelic variants of melon candidate genes involved in fruit ripening. 8 *Plant Genomics European Meetings (Plant Gem)*. 7-10th October, Lisboa (Portugal).