

CONVENIO

**TIPIFICACIÓN GENÉTICA DE MATERIAL
FORESTAL DE ALTO INTERÉS COMERCIAL**

INFORME FINAL

NOVIEMBRE DE 2011

NOVIEMBRE DE 2011

BOSQUES NATURALES S.A.-UNIVERSIDAD POLITECNICA DE MADRID

Convenio

TIPIFICACIÓN GENÉTICA DE MATERIAL FORESTAL DE ALTO INTERÉS COMERCIAL

Informe final

Bosques Naturales S.A. – Universidad Politécnica de Madrid

Índice

Planteamiento general	3
Resultados e Hitos	5
Tarea 1	5
Tarea 2	9
Tarea 3	21
Tarea 4	31
Validación metodológica	39
Resultados más relevantes	40
Conclusiones	41

Planteamiento general

El objetivo central de este proyecto de I+D ha sido desarrollar un sistema eficiente para la tipificación genética del material vegetal de la empresa. Dicho sistema debería conjugar idealmente sencillez, reproducibilidad, especificidad y bajo coste. El *scale-up* debería ser asimismo simple, dado el gran número de muestras potenciales en las fincas productivas. Se pretendía en última instancia desarrollar un sistema para tipificar cada genotipo (árbol, planta) de interés, discriminándolo de todos los demás. Dentro de este planteamiento, la metodología concreta se seleccionó de acuerdo con tres consideraciones generales:

Potencial. Se optó desde la concepción del proyecto por un sistema de genotipado basado en el uso de marcadores moleculares, concretamente de tipo microsatélite o SSR. Los marcadores moleculares son regiones específicas del genoma que pueden vincularse a características de interés aplicado –de ahí el término *marcadores*. Éstos pueden utilizarse para fines tan diversos como genotipado (este proyecto), estudios de estructura de poblaciones, determinación de relaciones de parentesco o identificación de flujos génicos, entre otros. En el ámbito forestal, la selección asistida por marcadores es una de las áreas que más interés suscita en este momento en el sector. Su potencial para la mejora es ingente a medio-largo plazo. En el momento de iniciarse el proyecto se disponía de un número bastante limitado de marcadores para *Juglans* (nogal). Lo contrario ocurría en *Prunus* (cerezo), dado el interés comercial de sus frutos.

Innovación. Los marcadores moleculares de tipo microsatélite han demostrado ser particularmente útiles para el genotipado a escala individual –de hecho, son los marcadores que utiliza la policía científica. Con algunas excepciones notables, su empleo en el mundo forestal es todavía escaso y ello a pesar de su gran interés aplicado. Los microsatélites son secuencias muy cortas de DNA, que se repiten en tándem un número variable de veces. Entre sus características más interesantes como marcadores genéticos cabe destacar que:

- 1) En general presentan un polimorfismo elevado en cuanto a su longitud, que es fácil de detectar con métodos e instrumentación apropiada; es necesaria una variabilidad grande si se pretende diferenciar individuos.

2) Presentan una abundancia relativa muy alta y están distribuidos regularmente por el genoma. En otras palabras, tienen un potencial muy alto para la mejora genética.

3) El grado de especificidad a nivel *locus* es grande en general y ello se debe a que los microsatélites suelen estar flanqueados por secuencias únicas, lo que facilita el desarrollo de *primers* para su amplificación específica mediante PCR.

4) No es raro que los *primers* desarrollados para una especie pueden ser transferibles a otras, sobre todo si no están muy alejadas filogenéticamente.

5) Los microsatélites son además marcadores co-dominantes, esto es, se diferencian fácilmente las variantes alélicas de ambos parentales, si las hubiera; todo ello facilita su identificación y el establecimiento de relaciones de parentesco entre individuos.

Investigación. Como se acaba de apuntar, había una diferencia grande de partida entre nogal y cerezo, las dos especies en que se centra este proyecto. En el momento de su inicio la disponibilidad de marcadores moleculares de nogal era muy escasa, concentrándose ésta en buena medida en especies de poco interés para la empresa, como *J. nigra* o *J. cinerea*. Era preciso desarrollar marcadores *de novo* para los nogales de Bosques Naturales, lo que supondría un esfuerzo adicional considerable respecto a *Prunus*. En el caso de nogal el proyecto era por tanto mucho más ambicioso científicamente y costoso. En el caso de cerezo se mantenía el mismo objetivo final, pero no se iban a construir genotecas ni se iban a desarrollar marcadores nuevos. Se utilizarían marcadores ya descritos, disponibles a través de publicaciones científicas y bases de datos. La importancia del cerezo como frutal y su elevado nivel de domesticación subyacen a esta diferencia de partida.

Este documento evita repetir la información contenida en el Anexo I (memoria técnica) del convenio en cuanto a líneas generales de actividad y su justificación, o detalles metodológicos para las distintas tareas y subtareas. Por el contrario, se centra en presentar un resumen suficientemente minucioso de los resultados y su puesta en valor. El grado de cumplimiento de los hitos del convenio se expone de forma pormenorizada.

Resultados e Hitos

Para facilitar su seguimiento, se utiliza la misma estructura de tareas y subtareas que en la memoria técnica del convenio.

Tarea 1. Genotecas enriquecidas en microsatélites: construcción y análisis

Subtarea 1.1. Purificación de DNA genómico de nogal

Esta subtarea se completó con éxito, si bien los protocolos habituales de purificación de DNA resultaron poco eficientes: se obtenían rendimientos bajos (μg de DNA/g de tejido fresco), el producto final no era de suficiente calidad (contaminación con polisacáridos o compuestos fenólicos) o el DNA purificado no era fácilmente digerible con endonucleasas de restricción. Dicha digestión es imprescindible para poder completar las subtareas siguientes. Los kits de QIAGEN específicos para plantas produjeron los mejores resultados (Figura 1 y Tabla I).

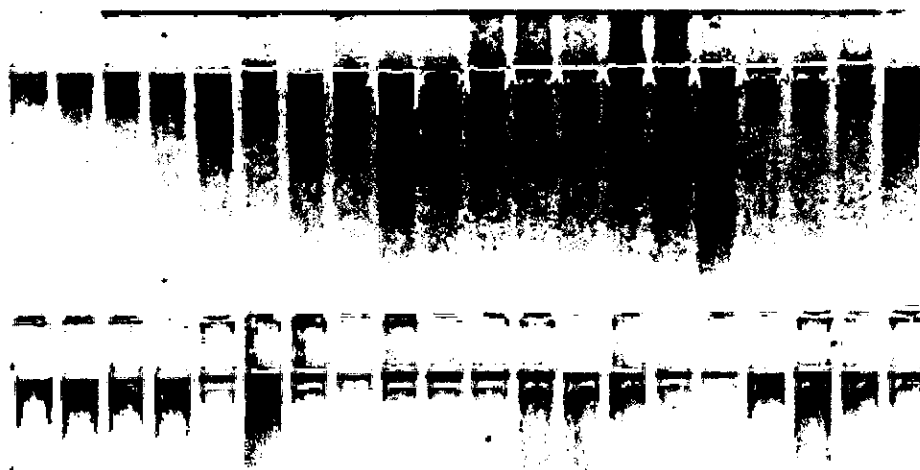


Figura 1. Comprobación de la integridad del DNA genómico mediante electroforesis en agarosa-TBE. Cada carril contiene DNA de un nogal distinto (en esta imagen se analizan 40 nogales). La purificación se llevó a cabo con el sistema DNeasy Plant Mini Kit de QIAGEN.

Tabla I. La calidad del DNA genómico se valoró también mediante espectroscopia UV. Para las purificaciones partimos de 100 mg de tejido fresco (hoja). La elución final se realizó en un volumen final de 100 μ l. Para todos los nogales analizados pudimos obtener DNA de buena calidad y en cantidad suficiente.

muestra	ng/ μ l	260/280	260/230	muestra	ng/ μ l	260/280	260/230
D-A	10.32	1.72	0.81	D-19	12.77	1.79	1.24
D-B	44.52	1.84	1.63	D-20	14	1.88	0.88
D-C	16.10	2.00	1.25	D-21	36.03	1.81	1.57
D-D	15.46	1.83	1.05	D-22	40.31	1.73	0.50
D-E	10.81	1.52	1.29	D-23	9.35	1.48	0.84
D-F	11.61	2.12	1.29	D-28	20.09	1.91	1.48
D-G	40.57	1.78	1.89	D-31	21.24	1.90	1.19
D-H	9.43	1.81	0.50	D-32	15.31	1.48	1.15
D-L	20.28	1.59	1.08	D-33	25.65	1.76	1.36
D-M	22.30	1.69	0.79	D-34	20.38	1.82	0.81
D-N	22.99	1.71	1.34	D-48	14.94	1.71	0.83
D-O	15.57	1.91	1.78	D-49	8.85	1.56	0.73
D-01	19.81	1.84	1.14	D-50	46.97	1.75	1.54
D-03	9.36	1.7	1.48	D-51	14.69	1.81	0.88
D-04	15.53	1.81	1.20	D-53	5.84	1.29	0.79
D-05	12.17	1.65	1.20	D-54	11.89	1.75	1.06
D-06	11.19	1.92	1.16	D-55	13.44	1.82	0.85
D-07	11.12	1.88	2.31	D-56	13.17	1.52	0.90
D-08	9.23	1.67	1.22	D-57	17.68	1.80	1.18
D-12	9.99	1.74	1.41	D-58	8.91	1.68	1.09
D-13	31.3	1.95	1.52	D-59	12.51	1.91	0.63
D-14	6.79	2.02	1.34	D-60	6.24	1.76	1.04
D-15	11.69	1.70	0.91	D-61	8.49	1.60	0.48
D-16	10.05	1.43	0.90	D-62	10.64	2.19	0.41
D-18	18.45	1.65	0.75	D-63	31.33	1.78	1.55

Subtarea 1.2. Fragmentación específica y selección por tamaño

Partiendo de más de 15 mg de DNA de excelente calidad de nogal híbrido, ensayamos diversas endonucleasas de restricción, a veces de forma aislada y a veces en dobles e incluso triples digestiones, hasta conseguir el grado deseado de fragmentación: una mezcla con fuerte predominio de tamaños en el intervalo de 100 – 300 pares de bases (pb). Esta subpoblación se purificó mediante filtración molecular en columna.

Subtarea 1.3. Enriquecimiento en microsatélites

Utilizamos en total cuatro tipos de secuencias microsatélite, dos con repeticiones tri-nucleotídicas y dos con repeticiones di-nucleotídicas: AAT, ATG, CA y GA. Fueron estas últimas las que produjeron mejores resultados en cuanto a nivel de enriquecimiento de las genotecas (véase la subtarea 1.5.). En todos los casos se utilizaron hexadecanucleótidos para capturar los fragmentos apropiados entre la población de moléculas de DNA genómico obtenida en la subtarea 1.2. Los heterodúplex resultantes sirvieron como material de partida para el clonaje descrito en la siguiente subtarea, tal y como estaba programado.

Subtarea 1.4. Clonaje en plásmido

Los heterodúplex seleccionados en el apartado anterior se clonaron en el vector plasmídico pUC19 (UC, University of California). Este plásmido confiere resistencia al antibiótico *ampicilina*. La secuencia del vector, de uso muy extendido, es de dominio público. Utilizamos métodos estándar para introducir los plásmidos recombinantes en la cepa DH5 α de la bacteria *Escherichia coli* (ElectroMaxJ, Invitrogen) mediante electroporación. Así cada genoteca se podría amplificar de manera sencilla y eficaz.

Subtarea 1.5. Análisis primario

Tras plaquear por separado cada una de las cuatro genotecas enriquecidas, seleccionamos cincuenta colonias (clones) al azar y secuenciamos los doscientos (50 x 4) insertos resultantes mediante la lectura de ambas hebras. El análisis de dichas secuencias reveló que los insertos tenían, en su inmensa mayoría, el tamaño esperado. Un análisis detallado de las secuencias nos permitió además determinar el nivel de enriquecimiento de cada genoteca. Las dos mejores, con diferencia, correspondían a los microsatélites CA y GA, cuyos niveles de enriquecimiento rondaban el 80% y cuya redundancia era inferior al 5%. En conjunto, podemos afirmar que ambas genotecas son de excelente calidad para detectar microsatélites.

HITOS - TAREA 1

HITOS PROPUESTOS EN EL CONVENIO (Anexo I)

Construcción de 1-2 genotecas en vector plásmido, enriquecidas en microsatélites de nogal (al menos un 60%). En cada caso se determinará

- el tamaño promedio de inserto
- el nivel de enriquecimiento en microsatélites
- el grado de redundancia de secuencia

HITOS CONSEGUIDOS:

Hemos construido 4 genotecas en vector plásmido, enriquecidas en microsatélites de nogal. Dos de ellas, que corresponden en concreto a las secuencias CA y GA, muestran niveles de enriquecimiento en torno al 80%.

Además de este parámetro, se ha cuantificado también el tamaño promedio de inserto (en torno a 240 pb) y el grado de redundancia de secuencia (inferior al 5%). Las dos genotecas seleccionadas son por tanto de magnífica calidad.

Tarea 2. Identificación de loci

Subtarea 2.1. Secuenciación nucleotídica a gran escala

El siguiente paso fue individualizar clones de las genotecas CA y GA para obtener la secuencia nucleotídica de sus insertos. A partir de bacterias capaces de crecer en medio selectivo para pUC19 –suplementado con *ampicilina*– seleccionamos clones al azar. Tras analizar el tamaño de los insertos mediante PCR y electroforesis (ver párrafo siguiente), obtuvimos finalmente 680 secuencias de buena calidad (**Figura 2**). Los clones correspondientes se depositaron por duplicado en microplacas de 96 pocillos, identificados con un código numérico. Cada réplica se almacenó en un congelador distinto por precaución. La secuenciación de insertos se llevó a cabo con los *primers* M13 *forward* y *reverse*, apropiados para clonajes en el vector pUC19. Dado el bajo nivel de redundancia que encontramos y la alta proporción de clones positivos, no fue necesario obtener más secuencias para completar el proyecto.

Se muestrearon en total cerca de 800 colonias (**Figura 2**). Para analizar el tamaño de los insertos, cada colonia se transfirió a un tubo eppendorf con la siguiente mezcla:

Tampón PCR 10X	1.0 μ l
MgCl ₂ (25 mM stock)	1.2 μ l
dNTPs mix (2.5 mM cada dNTP)	1.2 μ l
RNAsa A (20 mg/ml stock)	0.21 μ l
Primer forward (20 mM stock)	0.75 μ l
Primer reverse (20 mM stock)	0.75 μ l
Solución de rojo cresol-sacarosa	1.5 μ l
H ₂ O	3.39 μ l

La solución de rojo cresol-sacarosa (10 mM rojo cresol; 20% sacarosa p/v) no interfiere con la reacción PCR y permite cargar los geles de electroforesis directamente después de la amplificación. Además de un marcador coloreado, la sacarosa proporciona densidad a las muestras.

Secuencias de los *primers*:

Forward: 5'-AGG AAA CAG CTA TGA CCA TG-3'

Reverse: 5'-ACG ACG TTG TAA AAC GAC GG-3'

La mezcla anterior se incubó 10 min a 100°C para conseguir la lisis de las bacterias de la genoteca enriquecida. Entonces se añadieron los siguientes componentes a la mezcla:

4.425 µl H ₂ O
0.5 µl PCR Buffer 10X
0.075 µl Taq polimerasa (5 U/µl)

Tras centrifugar brevemente, las mezclas de reacción se sometieron al siguiente programa en un termociclador PCR:

(paso 1) 94°C 30"

(paso 2) 57°C 30"

(paso 3) 72°C 30"

Se llevaron a cabo 25 ciclos y después se realizó una extensión adicional de 5 min a 72°C. Las reacciones se mantuvieron entonces a 10°C hasta su análisis electroforético. Éste se llevó a cabo en geles de agarosa de alta densidad (3%, p/v) para conseguir una buena separación de los productos de PCR y una estima suficientemente precisa de su tamaño. Tras descartar los clones con insertos demasiado pequeños, los clones restantes se inocularon en medio LB con *Ampicilina* y se cultivaron 18 h a 37°C con una agitación de 250 rpm. Estos cultivos se utilizaron para purificar los vectores recombinantes de interés –es decir, los vectores con un inserto de DNA de nogal con un tamaño apropiado. Los insertos presentes en dichos plásmidos se secuenciaron con los *primers* indicados más arriba, para obtener una lectura de ambas hebras. Solo se consideraron para este proyecto secuencias en que ambas lecturas fueron consistentes y de buena calidad.

Subtarea 2.2. Diseño de primers

El análisis inicial de los datos globales de secuenciación se centró en tres puntos: (1) eliminar las secuencias de baja calidad o inconsistentes; (2) eliminar las secuencias redundantes; y (3) eliminar las secuencias procedentes del vector pUC19, situadas a ambos lados de cada inserto. Construimos entonces una base de datos de secuencias potencialmente únicas, identificadas con ayuda del programa AUTO-ASSEMBLER. La edición de los datos de secuencia se hizo con el programa BIOEDIT (Figura 3). Finalmente, las posibles parejas de primers para amplificaciones PCR se diseñaron con el programa PRIMER 3.0 (Tabla II). En esta ronda inicial de diseño permitimos un cierto intervalo de temperaturas de fusión (T_M) y de tamaños para el producto final de PCR, lo que nos llevó a 360 parejas de primers potencialmente útiles. Más adelante acotamos la población inicial en función de los resultados obtenidos en la subtarea 2.3. Aunque eran posibles otras alternativas, los primers que diseñamos resultaron ser más que suficientes para alcanzar los objetivos propuestos.

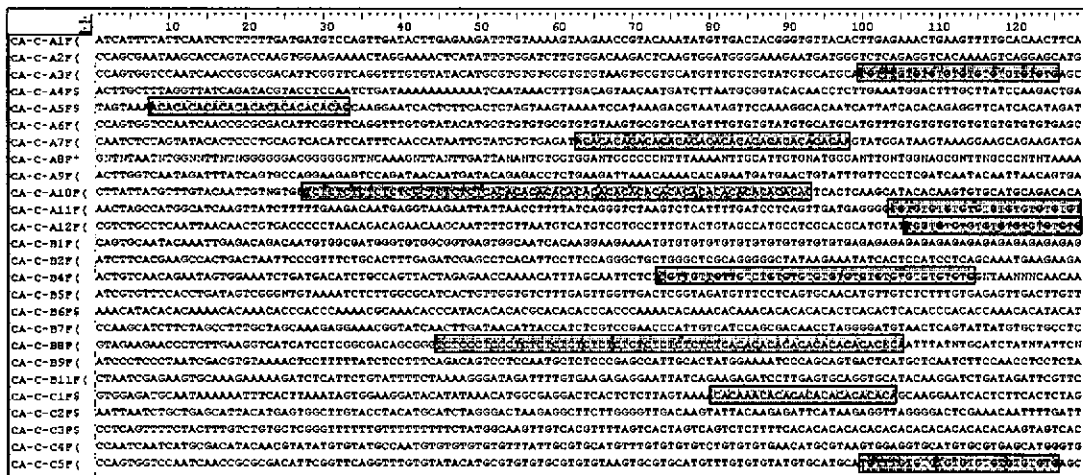


Figura 3. Captura de pantalla del programa BIOEDIT, en la que aparecen 26 secuencias correspondientes a otros tantos fragmentos de DNA genómico de nogal. Cada secuencia proviene de un clon individual de una de las genotecas enriquecidas. Aunque sólo se muestra una parte de cada secuencia en esta imagen, es suficiente para apreciar microsatélites en varios de los fragmentos (reticulado).

Tabla II. Condiciones utilizadas para el diseño de *primers* con el programa PRIMER 3.0. Todos los valores numéricos son modificables por el investigador.

Parámetros de selección de *primers*

Primer Length Minimum	18
Primer Length Maximum	22
Primer Tm Minimum	53 - 55
Primer Tm Maximum	60
Primer Tm Variance	± 1 - 2
Primer % GC Minimum	35
Primer % GC Maximum	65
Primer Concentration (μM)	50
Salt Concentration (mM)	50
Product Length Minimum	200
Product Length Maximum	300 - 350

En la **Tabla III** se muestran, a modo de ejemplo, algunos de los *primers* que diseñamos tras analizar exhaustivamente los datos de secuenciación para las librerías CA y GA con los programas informáticos antedichos. Para cada posible locus microsatélite se indica su nombre, el tamaño esperado del producto de PCR, la secuencia de la pareja de *primers* y la temperatura teórica de fusión (Tm). Como se detalla en el subapartado siguiente, analizamos centenares de posibilidades hasta optimizar la lista de locus microsatélite en función de la consistencia de los resultados y del polimorfismo encontrado con las muestras proporcionadas por la empresa.

Tabla III. Algunos de los primers ensayados para detectar locus microsatélite de interés (cuatro páginas siguientes).

Primers genoteca CA

Nombre	Tamaño del producto	Temperatura optima	Secuencias	Tm
A101	138	55.8	Fw: 5'-GCT-TTA-TTC-CAA-ATT-CTG-TCG-'3 Rv: 5'-AAA-TGA-AGG-TCC-AAC-TTA-GAG-C-'3	56.6 56.7
A102	293	57.2	Fw: 5'-CCC-TAA-IGT-TCG-CTC-TAG-C-'3 Rv: 5'-CCA-AAG-GTG-AGG-GTA-CAA-G-'3	55.2 55.6
A103	247	56.0	Fw: 5'-AAG-AGG-GTG-ATT-CCT-CAC-'3 Rv: 5'-GTA-AAG-CAC-AGT-TTC-ATG-TAA-G-'3	52.0 52.1
A105	255	57.0	Fw: 5'-TGA-AGC-AAC-ATT-TCC-CTT-AC-'3 Rv: 5'-GAG-CCG-TGG-TCT-CTA-ACC-'3	55.4 55.6
A109	216	57.5	Fw: 5'-GAG-CTT-CGT-CGA-TCA-GTT-AGT-C-'3 Rv: 5'-TTT-GGT-TAG-TCA-TGC-TGT-TGA-G-'3	58.2 58.0
A111	280	57.9	Fw: 5'-CCT-CCG-TCA-TCA-CTC-CTG-'3 Rv: 5'-TGT-GCC-TTT-GGA-ACT-ATT-ACG-'3	57.5 57.9
A112	192	56.5	Fw: 5'-CGT-GTC-GAA-TAT-CTA-TGT-GAT-G-'3 Rv: 5'-TTT-CGC-TTA-AAT-AGT-GGA-AGG-3	55.7 56.3
A114	228	57.0	Fw: 5'-ACG-GAG-AGA-AAG-TGT-ATG-GAG-'3 Rv: 5'-TCT-GGT-GCT-GGA-ATG-CTA-'3	56.0 56.2
A115	132	56.0	Fw: 5'-ACC-GTT-TAG-GTT-CTC-TCA-AGC-'3 Rv: 5'-AAG-TCA-TGC-ACT-TGG-CAT-AAC-'3	57.6 57.8
A115a	254	58.4	Fw: 5'-GGC-TTT-GGC-ATT-GGC-TTA-'3 Rv: 5'-TTG-GGG-GTT-CTG-CTT-CTG-'3	59.8 59.8
A116	253	57.7	Fw: 5'-ATT-TCC-GAG-GTT-GAA-GGT-G-'3 Rv: 5'-TTG-GGA-GGG-AGT-GAG-AGG-'3	57.6 58.7
A122	270	58.0	Fw: 5'-TCC-CAA-CAG-TTA-CAT-GAG-GC-'3 Rv: 5'-GGG-GAG-ATG-GAA-ATC-AAA-AC-'3	58.6 58.3
A123	141	56.8	Fw: 5'-CAC-CTT-TGC-ATC-TTT-TGT-TCA-C-'3 Rv: 5'-AGG-GTT-TTC-AAG-TGT-CGT-TTT-C-'3	59.7 59.5
A125	222	56.8	Fw: 5'-TTT-CAA-TAT-GGA-ATC-CTG-TAC-C-'3 Rv: 5'-AGG-CAT-TCT-CAA-CAG-AAC-AA-'3	55.7 55.9
A128	201	56.7	Fw: 5'-TGG-AAG-TGA-AGC-AAT-ATG-TTT-C-'3 Rv: 5'-CCT-TTT-GGA-ATC-GTA-CTA-CAA-A-'3	56.9 56.1
A129	182	56.3	Fw: 5'-TCG-CAC-CTA-CAC-TGG-TAT-TAG-'3 Rv: 5'-GCT-ATT-GAG-CAA-CAA-AAT-GAG-'3	55.5 55.3
A129	188	56.3	Fw: 5'-TTT-GTT-GCT-CAA-TAG-CTT-GTA-C-'3 Rv: 5'-AGC-TCT-ACG-GAA-CCC-TTT-AC-'3	55.0 55.2
A131	147	55.8	Fw: 5'-CAA-ACT-CAA-ATC-TTC-CTC-AAA-C-'3 Rv: 5'-GCT-TGT-CCA-AAA-GTA-IGG-TG-'3	55.7 55.7
A132	212	56.8	Fw: 5'-GAA-TGT-TAG-GCT-TTG-GTT-GAC-'3 Rv: 5'-GGG-AAG-GTT-TTA-GAC-ACC-C-'3	56.4 56.0
A134	251	57.0	Fw: 5'-TGT-AGG-AAT-CCG-AAG-AGT-TG-'3 Rv: 5'-AGG-TAA-AGA-GTG-TTG-CCA-CTA-C-'3	55.4 55.3
A201	186	56.2	Fw: 5'-GCA-TCC-AAA-GGT-AAG-TTA-GAA-G-'3 Rv: 5'-TGG-ACC-CTA-GAA-CAC-AAC-A-'3	55.0 54.8
A202	201	56.7	Fw: 5'-TGG-GAT-TTT-GTT-GTA-AGG-G-'3 Rv: 5'-ACC-AGT-TGT-CAC-CAT-CTG-TC-'3	55.9 55.8
A203	180	56.4	Fw: 5'-TCC-CAG-TCA-CCT-CAG-AAA-C-'3 Rv: 5'-GGT-TCC-CAC-ACT-AAT-CAT-AGA-A-'3	56.0 55.8
A204	164	56.5	Fw: 5'-CAG-TTC-CTT-TTA-GGG-GTT-CTC-'3 Rv: 5'-TCA-GGT-GAG-GGA-TTT-GAT-G-'3	57.0 56.9
A206	191	56.6	Fw: 5'-ACA-TCG-TTT-AGG-GGT-GAA-TT-'3 Rv: 5'-TTC-CCT-TCC-ATT-CTT-TGC-'3	56.0 56.1
A210	285	57.8	Fw: 5'-CAT-GGT-TCT-CAA-GTA-CCT-CTC-C-'3 Rv: 5'-TGC-AAG-ATT-CAG-AAA-CTC-TAC-G-'3	57.8 57.3
A211	171	56.6	Fw: 5'-TCT-CCT-TCT-TCT-TCC-TCA-TTT-C-'3 Rv: 5'-ATC-GGT-CYC-GGC-ATA-ACT-AC-'3	56.8 56.8
A212	259	57.2	Fw: 5'-TCC-CAC-GAA-TAA-CTT-TCT-GTT-'3 Rv: 5'-TTT-AGC-TCT-TTG-CCA-CAA-TC-'3	56.4 55.7
A213	239	56.8	Fw: 5'-CTA-TGG-TTT-TCT-TTG-TGA-GAC-C-'3 Rv: 5'-GTG-CCT-TTG-GAA-CTA-TTA-CG-'3	55.6 55.0
A214	212	56.8	Fw: 5'-TAG-CAG-CGT-CCC-TAT-AAC-TG-'3 Rv: 5'-GCA-CTC-TCG-AAT-TGT-CAA-AC-'3	55.8 55.9
A216	136	56.5	Fw: 5'-GCA-CGA-CAA-AGA-ATC-ACC-C-'3 Rv: 5'-TCG-GGG-AGG-GAA-ATA-GAA-C-'3	59.1 58.9
A221	212	57.3	Fw: 5'-AGC-CAC-GAT-ACA-AAC-ACA-AAC-'3 Rv: 5'-AGG-ATA-AAA-IGC-TTG-CGA-ATA-G-'3	57.7 57.7

A222	157	56.2	Fw: 5'-AAC-ACC-ACC-ACC-ACT-AAG-TTC-'3 Rv: 5'-TTC-TGC-TCT-TCC-CTC-ACT-AAG-'3	56.5 56.4
A223	268	57.1	Fw: 5'-GAG-CAC-TGA-GAG-AAA-GAG-AAA-G-'3 Rv: 5'-CCC-AAA-ACA-GAG-GAA-AGT-AAC-'3	55.3 55.5
A224	119	55.5	Fw: 5'-CAT-AAG-AGG-GAA-ATC-CTC-AGA-G-'3 Rv: 5'-GAC-GCA-TGA-GTC-CAT-AAA-TAT-G-'3	57.1 57.2
A226	125	55.3	Fw: 5'-CTA-GCT-GTA-ACG-TCC-ATA-GCT-C-'3 Rv: 5'-ATT-GCT-TCG-GTA-TCA-GTA-AGT-G-'3	56.1 56.2
A227	166	56.2	Fw: 5'-GGT-AGG-GTC-ATT-GGA-ACA-TAG-'3 Rv: 5'-GGC-GTA-CTC-TGG-AGA-CAT-ATA-G-'3	56.0 55.7
A230	160	55.9	Fw: 5'-AGT-GGC-AAT-CTC-AAG-GAA-G-'3 Rv: 5'-TCT-CCG-TTC-ATC-TCA-TTC-TC-'3	55.4 55.2
A231	274	58.4	Fw: 5'-TTC-AGC-CAA-CTT-TGA-AGA-CAA-C-'3 Rv: 5'-GGA-GCA-TAT-TGC-AGA-TCA-TCA-G-'3	59.4 59.7
A233	189	56.4	Fw: 5'-GCA-CTG-TTC-AGG-CGT-AAT-T-'3 Rv: 5'-ATC-CAG-GAA-AGG-GTA-ATA-AGT-C-'3	56.4 55.5
A235	268	57.5	Fw: 5'-CAC-CAT-CTG-TCT-AGG-TTC-AAT-G-'3 Rv: 5'-AGG-ACC-TTA-GAG-TTC-CAG-GAG-'3	57.2 56.6
A236	230	57.1	Fw: 5'-CAG-TTG-AAT-AAG-TTG-CCA-CTC-T-'3 Rv: 5'-TGG-AAG-TGA-AGC-AAT-ATG-TTT-C-'3	56.2 56.9
A237	293	57.5	Fw: 5'-ATC-TCT-CCT-TCC-TTT-CCT-CTC-T-'3 Rv: 5'-GAC-AGG-TTC-TTC-AAC-GAG-TCT-'3	56.8 56.0
A238	207	57.2	Fw: 5'-CAG-CAG-TTA-CCA-TGC-AAT-TAA-G-'3 Rv: 5'-GCT-GCA-ATA-CAC-AAG-ACA-ATT-C-'3	57.6 57.4
A241	185	57.2	Fw: 5'-GGA-ACC-AGC-AGA-AAC-GAA-C-'3 Rv: 5'-CAT-TTG-GCA-CAC-CTG-AGA-A-'3	58.3 58.2
A243	169	57.1	Fw: 5'-TTG-CAT-CCA-AAT-AGC-TGT-CC-'3 Rv: 5'-TTG-CTG-GGA-ACT-ACT-GCA-TC-'3	58.7 58.9
A248	172	56.2	Fw: 5'-TGC-CTC-AAT-AGG-AGA-GTT-AGA-G-'3 Rv: 5'-GCT-GTC-AGA-TTA-CCC-AGA-GTA-G-'3	55.6 55.7

Primers genoteca GA

Nombre	Tamaño del producto	Temperatura óptima	Secuencias	Tm
B102	210	56.8	Fw: 5'-TGA-CAT-GGT-GTT-CAT-AAC-TCA-G-'3 Rv: 5'-GTC-AAC-AGG-CAA-AGA-ACT-AGA-C-'3	56.1 55.8
B103	213	57.7	Fw: 5'-ACC-CAT-TGT-TGG-GGT-TAG-TG-'3 Rv: 5'-CTA-CGG-GTG-CAA-TTT-CGT-TA-'3	59.6 58.7
B105	207	57.2	Fw: 5'-TAG-CCT-CGT-TCC-AGG-TTC-TAT-A-'3 Rv: 5'-CCT-AGA-GGT-AGA-TGC-GGT-AGA-G-'3	58.1 57.3
B107	188	57.2	Fw: 5'-TAG-CCA-ACC-TTT-GTT-TGA-CTG-'3 Rv: 5'-TCG-GAC-ACA-ACT-TAC-AAC-CAC-'3	57.9 58.5
B109	133	56.0	Fw: 5'-CCT-TGT-TGC-TTT-GGG-AGT-AA-'3 Rv: 5'-CGC-ATT-TGA-ACA-GAA-TAA-AGT-G-'3	57.9 57.5
B109a	152	55.9	Fw: 5'-AAT-GCG-ATA-CTT-TGC-TGT-TG-'3 Rv: 5'-GGA-GTG-AGA-TGA-GAT-TCT-GAA-G-'3	56.5 55.6
B111	193	56.5	Fw: 5'-GCC-TGT-GGT-CAG-TGT-TCA-'3 Rv: 5'-GAC-GGG-GAA-GAA-ATC-AAG-'3	56.4 55.5
B113	141	56.3	Fw: 5'-ACC-GAT-AAG-GCT-GCG-ATA-C-'3 Rv: 5'-CCT-CGA-CAT-CAC-ATG-AAG-C-'3	57.8 57.7
B114	200	57.7	Fw: 5'-CTT-CCC-TCA-GCC-AAT-TAG-AGA-C-'3 Rv: 5'-TCC-TTG-TTC-GAT-GAA-GAC-TTT-G-'3	59.4 59.4
B115	246	57.7	Fw: 5'-AAA-GAG-AGA-GTC-AGC-GAG-TGA-C-'3 Rv: 5'-GTT-AGA-TAA-TGG-GCG-TTG-TGA-C-'3	57.9 58.5
B118	250	57.8	Fw: 5'-GCT-TCT-CTT-CAT-CTC-GTT-GAT-C-'3 Rv: 5'-ACG-GAC-GAA-CCA-AGT-CAT-ATA-G-'3	58.1 58.1
B120	183	56.3	Fw: 5'-AGG-TGG-GAA-TAT-GAC-ACT-GTA-C-'3 Rv: 5'-AAT-CGT-GGT-GCT-TAT-TAG-ATT-C-'3	55.5 55.2
B121	144	56.9	Fw: 5'-AGC-CAA-TAA-AAC-AGG-AAA-TTG-C-'3 Rv: 5'-CCC-TCT-CCC-AGT-CAT-AAG-AAT-G-'3	59.5 60.0
B122	219	57.9	Fw: 5'-CAT-CCT-CCC-CAC-CAT-AAA-AG-'3 Rv: 5'-TTG-ACC-CAC-TCA-CCA-TCA-AC-'3	59.2 59.4

B125	279	57.7	Fw: 5'-TTC-GTC-TTC-TTC-CTC-TTT-GTG-'3 Rv: 5'-CAG-TTG-GTA-GTT-GTG-CGT-TAT-T-'3	57.2 57.0
B127	237	57.7	Fw: 5'-GAC-GAT-AGC-CCT-GAG-TTC-TTC-'3 Rv: 5'-ATG-AAT-CTG-CGA-AAC-CAA-TAA-G-'3	58.0 58.3
B128	248	56.9	Fw: 5'-GGC-TCT-CAG-CCA-CTA-GAA-C-'3 Rv: 5'-CTG-GGC-AGT-AAT-TTA-GGA-TTG-'3	55.1 56.5
B129	134	56.6	Fw: 5'-GGG-TCT-TTG-TGG-GTC-AAA-AC-'3 Rv: 5'-TCT-CCC-GCA-TCT-CTT-CTT-TC-'3	59.3 59.5
B131	193	56.8	Fw: 5'-ATG-AAC-CCT-TGT-ACG-ATG-ATC-'3 Rv: 5'-ATT-TCC-TCA-AAT-TCT-TCC-AGA-G-'3	56.4 56.6
B133	162	55.9	Fw: 5'-GGA-GGA-TGA-GAT-TGT-GAG-C-'3 Rv: 5'-TAA-CCT-TAT-GCG-GCA-CTT-AC-'3	55.5 55.2
B135	184	56.7	Fw: 5'-CAT-TTC-AGC-CTC-TTG-AAG-TTC-'3 Rv: 5'-ATA-GCG-TAA-CCA-GCT-CCT-TG-'3	56.7 57.2
B135a	170	57.2	Fw: 5'-ACA-GCA-CTC-TCA-GGC-TAG-AGC-'3 Rv: 5'-CTG-CCG-AAT-CTG-GTA-AAA-CC-'3	59.0 59.6
B136	194	56.5	Fw: 5'-AAA-GCT-CGA-TCC-TCT-CCT-T-'3 Rv: 5'-GGA-TGA-TGA-AAG-AAG-AAT-ATG-C-'3	55.7 55.5
B137	247	56.9	Fw: 5'-GGC-ACC-ATA-GAA-GCA-ATG-'3 Rv: 5'-TAA-GCG-GGG-TGT-TAC-AAA-'3	55.0 55.1
B201	297	57.5	Fw: 5'-GGG-GGC-ICT-CTC-TTT-GTC-'3 Rv: 5'-TGG-CTC-TTG-CTC-AGA-CTG-'3	57.2 56.0
B203	291	57.9	Fw: 5'-GAG-AGG-CAC-ACT-GAA-AGA-TTG-'3 Rv: 5'-CCA-CCA-TTA-GTA-AAC-CAC-AGC-'3	57.5 57.2
B204	275	58.5	Fw: 5'-TTC-TTG-CAG-GAT-GGA-CCT-AGA-'3 Rv: 5'-TTG-TTG-CTA-CTT-TTG-GGA-CTC-A-'3	59.8 59.8
B205	299	57.6	Fw: 5'-CGT-TGT-CGC-TTC-ACT-TAG-TC-'3 Rv: 5'-GTT-TGA-TTT-GGT-CTG-GAT-TGT-'3	56.2 56.5
B206	194	57.0	Fw: 5'-GGC-TCC-GTT-TCC-CTC-TAA-'3 Rv: 5'-AGG-TGC-TCC-GTG-AGA-CAC-'3	57.8 57.1
B206a	171	56.1	Fw: 5'-ATC-GTC-TCT-CTC-TCA-GTG-TCT-C-'3 Rv: 5'-AAC-CGT-CTA-TAT-CAG-TCT-CAG-C-'3	55.2 55.3
B208	172	56.7	Fw: 5'-TAC-GCC-TTT-AGG-GAA-GCC-'3 Rv: 5'-CAA-TGA-TGA-GTG-TTG-GTG-ATG-'3	57.9 57.4
B209	265	57.5	Fw: 5'-AGA-GGA-ATG-AGA-GGA-ACT-TCA-G-'3 Rv: 5'-TAT-TAG-CTT-CCC-CAA-AGT-CAT-C-'3	56.7 57.4
B210	213	57.4	Fw: 5'-GCA-AGC-ACG-ATA-GAA-AGG-AC-'3 Rv: 5'-GGG-ATA-CCT-CTG-ATG-CCT-AAG-'3	57.6 57.8
B211	175	56.5	Fw: 5'-CTC-AGG-CAA-GCA-CAA-ATG-'3 Rv: 5'-GCA-GGG-ATG-GTA-GAA-GAT-TC-'3	56.3 56.2
B212	156	56.1	Fw: 5'-CGT-TCA-TCT-CAT-TCT-CTC-TGT-C-'3 Rv: 5'-AGT-GGC-AAT-CAC-AAG-GAA-G-'3	56.6 56.2
B213	180	56.4	Fw: 5'-GCG-TCT-ATT-ATT-GTT-TTC-CAA-C-'3 Rv: 5'-GCT-TAT-TGT-GAA-CAT-CAT-CAT-G-'3	55.7 55.7
B214	247	57.0	Fw: 5'-GCA-GCA-CTT-GAC-GGT-TAG-'3 Rv: 5'-GAC-CAA-TCA-GGA-TAA-GAA-TGT-G-'3	55.4 56.2
B215	107	54.5	Fw: 5'-TTC-CTC-CCT-CAC-ACT-TGT-AG-'3 Rv: 5'-GCC-ATT-IGG-ICT-ATG-TTC-TAT-C-'3	55.3 55.4
B216	146	55.9	Fw: 5'-TCC-CTT-TAG-CAT-CTA-TGA-ACT-G-'3 Rv: 5'-GAT-GGT-GAA-TGT-ICC-TGT-ACT-G-'3	56.2 57.0
B217	212	57.4	Fw: 5'-CAG-CAA-TAC-TTT-CCC-TCG-TG-'3 Rv: 5'-TTG-CGA-AGG-ACT-TCA-AAT-TC-'3	58.4 57.9
B219	100	54.7	Fw: 5'-TTT-GAG-GTT-GAA-TGA-GAG-TGA-G-'3 Rv: 5'-ACC-ATT-TCA-CCC-ATA-TCT-ICT-C-'3	57.1 57.0
B220	257	57.2	Fw: 5'-CCT-CCA-ACA-ACT-GCT-TTA-CTA-C-'3 Rv: 5'-CCT-GGG-GAA-TCA-ACT-GAC-'3	55.8 56.2
B221	260	57.0	Fw: 5'-CCT-GCT-TAG-CCT-TGT-CAT-AG-'3 Rv: 5'-AAC-CTA-GAA-AGC-CCT-AAA-AGT-C-'3	55.3 55.3
B222	140	55.9	Fw: 5'-TAA-CAT-ACA-CAC-GCA-AAC-ACA-G-'3 Rv: 5'-GAG-GGC-ACA-CCC-ACT-AAG-'3	57.3 56.5
B223	231	57.0	Fw: 5'-CGT-GGG-TCC-ATA-AGT-CCT-A-'3 Rv: 5'-AAA-CAC-ATT-TGA-TGC-GTA-AGT-C-'3	56.0 56.9
B224	122	55.4	Fw: 5'-CCA-GCA-GGA-GAG-TCT-TCG-'3 Rv: 5'-AGC-ACA-CCA-TAG-AGA-GAA-AAC-A-'3	57.1 56.6
B226	271	58.1	Fw: 5'-CTC-GTT-GAG-CCA-AGA-AAA-TG-'3 Rv: 5'-CCA-AGT-GGA-TGT-CTG-TCT-GC-'3	58.5 59.3
B227	109	55.1	Fw: 5'-AAT-ITA-ACG-CCC-CTT-TCA-TC-'3 Rv: 5'-CAC-GCA-TGG-GAA-GAA-ACT-'3	58.1 57.1
B227a	240	57.5	Fw: 5'-ATG-CGT-GCT-TAT-CTG-TTT-AGC-'3 Rv: 5'-TCA-AGT-GTG-GAT-GGG-ACT-ATC-'3	57.3 57.4
B232	145	56.2	Fw: 5'-CTC-GAT-TAG-GGG-GTG-AGT-C-'3 Rv: 5'-GAG-GAG-GTG-GCA-GAA-TAC-C-'3	57.1 57.1
B233	153	56.0	Fw: 5'-GGA-CTA-CGC-GGA-TTT-AGT-G-'3 Rv: 5'-CTG-GTC-CCC-GTT-TAG-GTA-'3	55.9 56.0

B234	174	57.1	Fw: 5'-TCT-CCC-TCT-CAG-ATT-TAT-CGG-'3 Rv: 5'-CCA-ATG-CCA-GAA-ACA-ATA-CG-'3	58.3 59.0
B235	237	56.9	Fw: 5'-GAA-GCT-CAG-CAT-CTT-CTT-TG-'3 Rv: 5'-AGG-CGA-TTA-AGT-TGG-GTA-AC-'3	55.4 55.9
B236	174	56.2	Fw: 5'-CAG-GTC-CTC-CTT-TCT-CTT-TC-'3 Rv: 5'-GCC-TCT-TCG-TAT-CTG-TTT-CTC-'3	55.6 55.8
B240	189	57.3	Fw: 5'-TAA-CCG-GAT-ACT-TCC-CTT-TCT-T-'3 Rv: 5'-ACT-GAG-AGA-GGA-ATG-GTG-TGA-C-'3	58.2 58.2
B240a	279	57.3	Fw: 5'-CCA-CCA-CAC-AAA-GTC-ACA-C-'3 Rv: 5'-TAC-CCA-CTA-CAA-GAG-GAA-AAG-A-'3	55.6 55.7
B244	198	56.6	Fw: 5'-GCA-AAG-AGA-GTT-ACA-ATG-GTT-C-'3 Rv: 5'-TAG-GAG-TAT-GAC-ACG-GAA-TGT-T-'3	55.7 56.3
B245	129	55.3	Fw: 5'-CTC-CCA-GTC-TCT-CTG-ATT-CTC-'3 Rv: 5'-AAC-CCT-AAT-GAC-CAT-AAC-ACA-G-'3	55.6 55.7
B246	178	56.7	Fw: 5'-CGC-ATC-TAC-TTC-CAC-CTT-G-'3 Rv: 5'-GGA-CCA-CCC-ATA-CCC-TCT-A-'3	56.8 57.2
B248	256	58.1	Fw: 5'-CCA-AAT-GGC-AAA-TAT-CAT-CAA-G-'3 Rv: 5'-TTC-CAC-TGC-CTT-GTT-TCA-AG-'3	59.3 58.9

Subtarea 2.3. Optimización de parámetros y validación de la metodología.

Como se ha indicado anteriormente, uno de los objetivos clave del proyecto era desarrollar un sistema eficaz y de coste asequible para el análisis de muestras a gran escala. Dicho objetivo se pudo cumplir mediante la integración de dos estrategias experimentales:

1.- La primera estrategia se refiere a la extracción de DNA genómico. No hay que confundir con la subtarea 1.1., cuyo propósito era extraer gran cantidad de DNA para construir genotecas. Ahora se trataba de obtener cantidades suficientes de DNA de buena calidad, a partir de ejemplares individuales, para llevar a cabo una tipificación genética mediante PCR. Los *primers* utilizados fueron los diseñados en el apartado anterior. Como el nogal es caducifolio, no se extrajo únicamente DNA de hojas sino también de yemas y ramas, para constatar que el genotipado podía realizarse en cualquier momento del año.

2.- La segunda estrategia, más compleja y extensa, se centró en la evaluación de los *primers* diseñados *de novo*. Para ello utilizamos DNA genómico de ejemplares individuales: híbridos MJ209 x RA de la empresa, uno de ellos de la finca de Toledo y el resto de Cáceres; también incluimos *J. regia* de la E.T.S. de Ingenieros de Montes. Estos experimentos nos permitieron descartar *primers* de baja especificidad, que daban lugar a bandas artefactuales, que producían inconsistencias en las réplicas, etc.

Para cada pareja de *primers* se determinaron las condiciones óptimas de PCR, con particular énfasis en la temperatura de *annealing* (el paso 2 de un ciclo típico, justo después de la desnaturalización térmica). Los resultados de las amplificaciones se evaluaron inicialmente mediante electroforesis en geles de agarosa MetaPhor de alta resolución, ideal para el análisis de microsátélites (**Figura 4**). De esta forma pudimos clasificar los *primers* en función de la calidad y reproducibilidad de los resultados. En algunos casos mejoramos notablemente los patrones electroforéticos reajustando las condiciones teóricas de PCR para aumentar la especificidad.

Nuestra capacidad discriminante fue aumentando lógicamente a medida que refinamos la selección de *primers*, buscando el máximo nivel de polimorfismo. Muchos de los diseñados a nivel teórico en la subtarea 2.2 fueron descartados en esta fase, como es habitual en proyectos de esta naturaleza. Al mismo tiempo, simplificamos y miniaturizamos el proceso de análisis, pudiendo reducir la cantidad necesaria de DNA a tan sólo 5 ng por muestra –por debajo de lo publicado para otras especies de *Juglans*– sin efectos negativos en la consistencia o reproducibilidad de los resultados. De esta forma se necesitarán cantidades mínimas de material recogido en campo para proceder a genotipar. La consistencia de los resultados no se vió afectada tampoco porque el material procediese de hoja, yema o rama, si bien los rendimientos de DNA tienden a ser más altos en el primer caso.

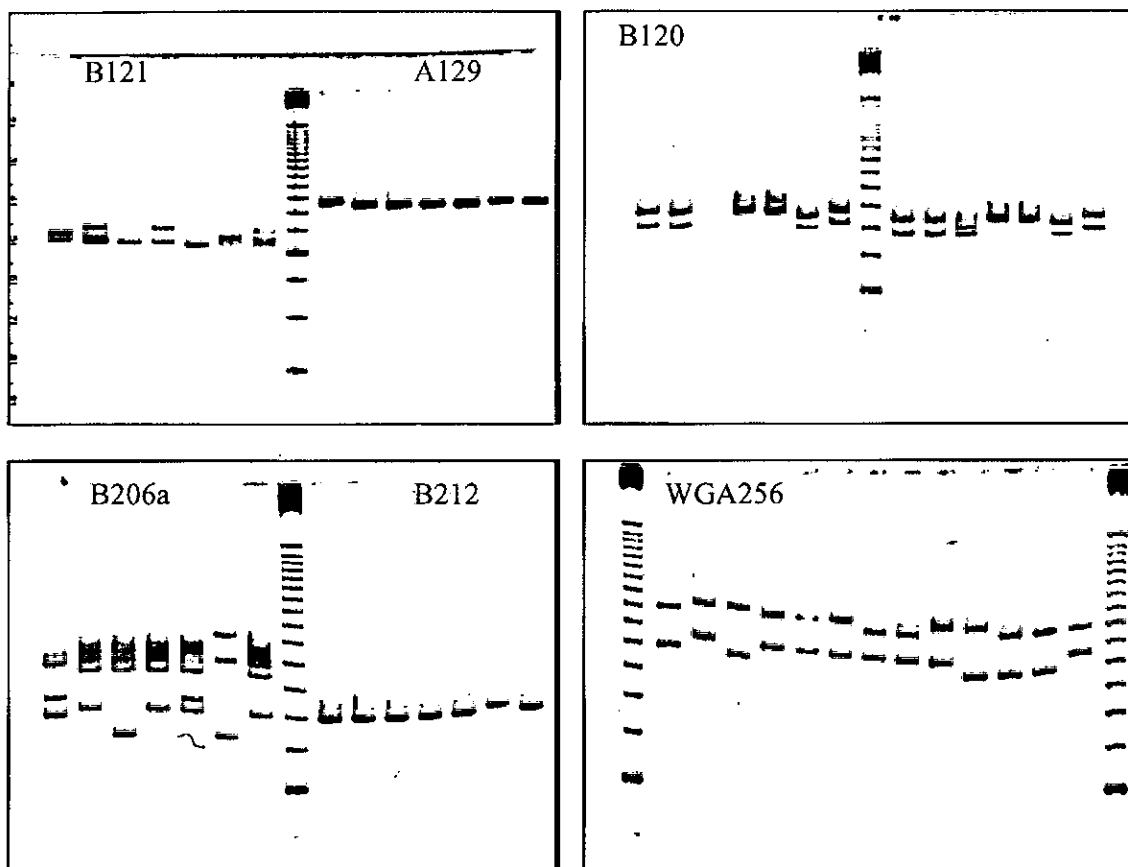


Figura 4. Separación de productos de PCR mediante electroforesis de alta resolución en geles de agarosa MetaPhor, específica para el análisis de microsatélites. El *locus* concreto analizado se indica en la parte superior del gel. WGA 256 es uno de los diez *loci* de *Juglans cinerea* descritos por Ross-Davis y Woeste (2008). Los carriles de marcadores (escaleras de fragmentos de tamaño conocido) se sitúan en posición central (paneles 1-3) o en los extremos (panel 4) y sirven para calibrar el gel. Los restantes carriles corresponden a reacciones individuales de amplificación con DNA genómico (5 ng) de nogales individuales. Nótese las diferencias entre ejemplares (se han incluido en la figura marcadores con distintos niveles de polimorfismo). Los datos obtenidos con WGA-256 son comparables a los obtenidos con los marcadores desarrollados en este proyecto.

HITOS - TAREA 2

HITOS ACORDADOS EN EL CONVENIO (Anexo I)

- Diseño y testado de *primers* para nogal.
- Puesta a punto de un sistema para el análisis de microsatélites a gran escala.
- Detección de *loci* polimórficos. Evaluación del nivel de variabilidad.

HITOS CONSEGUIDOS:

Las cifras concretas sobre diseño de *primers* son: de las 180 parejas diseñadas inicialmente a partir de cada genoteca, 360 parejas en total, han pasado la primera ronda de selección 47 para la genoteca CA y 38 para la genoteca GA. Es decir, un total de 85 parejas rindieron resultados de buena calidad utilizando geles MetaPhor a partir de una muestra de nogales de Cáceres, Toledo y Madrid.

Hemos puesto a punto un sistema de análisis simplificado y de alta reproducibilidad, como revelan las Figuras de este informe, que sólo requiere 5 ng de DNA por ejemplar de nogal (1 ng = 10^{-9} g). Las muestras pueden provenir indistintamente de hoja, yema o rama, lo que permite genotipar a lo largo de todo el año.

Respecto a los análisis de variabilidad, la mostrada por los *loci* ensayados en esta fase nos ha permitido culminar el proyecto con éxito. Esto es, no ha habido que evaluar *primers* adicionales a los que se describen en la Tabla III. Las Tablas IV y V al final de la sección siguiente muestran la selección definitiva de *loci* utilizada para genotipar los nogales de la empresa.

Tarea 3: Genotipado de material clonal seleccionado (nogal)

Aunque el número de donadores de polen era desconocido cuando se inició el proyecto, no era esperable un nivel de polimorfismo demasiado elevado, al provenir toda la población híbrida analizada de un único clon materno de *J. major*. Para soslayar este problema, inherente al hecho de estar analizando una F1 híbrida, reevaluamos las parejas de *primers* seleccionadas en la tarea anterior con una muestra más amplia de nogales y reajustamos las condiciones experimentales de las amplificaciones PCR y las electroforesis. Utilizamos en concreto 18 nogales procedentes de las fincas de Cáceres para este segundo análisis, lo que nos permitió depurar la selección inicial. Eliminamos así *loci* para los que un único alelo mostraba frecuencias muy elevadas, en algunos casos cercanas a 1, o *loci* que no producían resultados consistentes tras aumentar el tamaño de la muestra (Figuras 5 y 6).

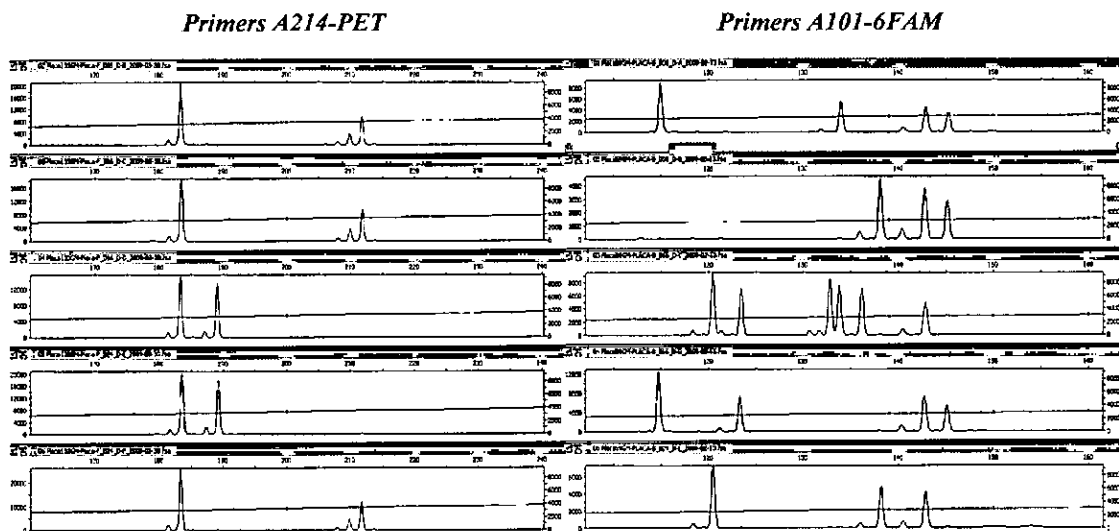


Figura 5. De la librería CA se descartó por ejemplo la pareja A214, porque uno de sus alelos era igual para todas las muestras. También la pareja A101, porque rendía patrones demasiado heterogéneos, un problema que no pudo solventarse modificando la temperatura. En ambos paneles los productos de PCR se fraccionaron mediante electroforesis capilar. Dichos productos están marcados con fluoróforos (PET y 6FAM en este caso), lo que permite visualizarlos mediante excitación con láser y mediciones de fluorescencia a la longitud de onda apropiada. Cada panel horizontal corresponde a un nogal individual. Si se comparan estos paneles con los carriles de la electroforesis de la Figura 6, las mejoras son evidentes: los productos de PCR están mucho mejor definidos (picos en vez de bandas) y su tamaño puede ser determinado con precisión a partir de su posición en el eje horizontal.

Primers B102-6FAM

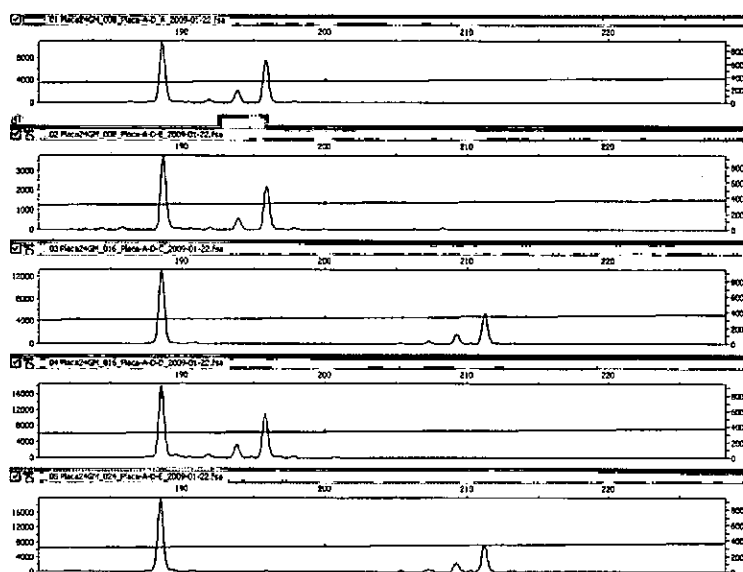


Figura 6. De la librería GA se descartó por ejemplo la pareja B102, porque uno de sus alelos era común a todas las muestras.

En paralelo con las mejoras antedichas incorporamos también dos innovaciones, no previstas en la memoria inicial del proyecto, que incrementaron de forma significativa nuestra capacidad discriminante global:

1) Robotización del proceso y *multiplexing*

Las principales modificaciones que introdujimos en este punto fueron:

1.- Los productos de PCR se fraccionaron de forma automatizada, utilizando el mismo sistema multicapilar que los secuenciadores automáticos de DNA. Esto proporciona información muy precisa y reproducible sobre el tamaño de los fragmentos amplificados, mejorando la detección de polimorfismos (compárense las Figuras 5 y 6, en las que se utiliza el nuevo sistema, con la Figura 4, correspondiente a una electroforesis en agarosa MetaPhor).

2.- Las reacciones de amplificación se llevaron a cabo con *primers* marcados con fluoróforos, uno por *locus*, para facilitar la detección de los productos mediante fluorescencia. Esto mejora la sensibilidad y facilita la informatización de los resultados.

3.- Las reacciones se realizaron en microplacas de 96 muestras, lo que facilita la robotización del análisis.

4.- La documentación y análisis de los resultados se informatizaron por completo.

El resultado final es muy superior a la electroforesis en gel, particularmente en cuanto a sensibilidad y especificidad (Figuras 5 y 6). También mejora de forma considerable la reproducibilidad (Figura 7). La combinación apropiada de fluoróforos permite además simultanear el análisis de varios *loci* (*multiplexing*), con el consiguiente ahorro de tiempo y materiales. La Figura 8 muestra el análisis de cinco nogales diferentes con cuatro marcadores especialmente polimórficos.

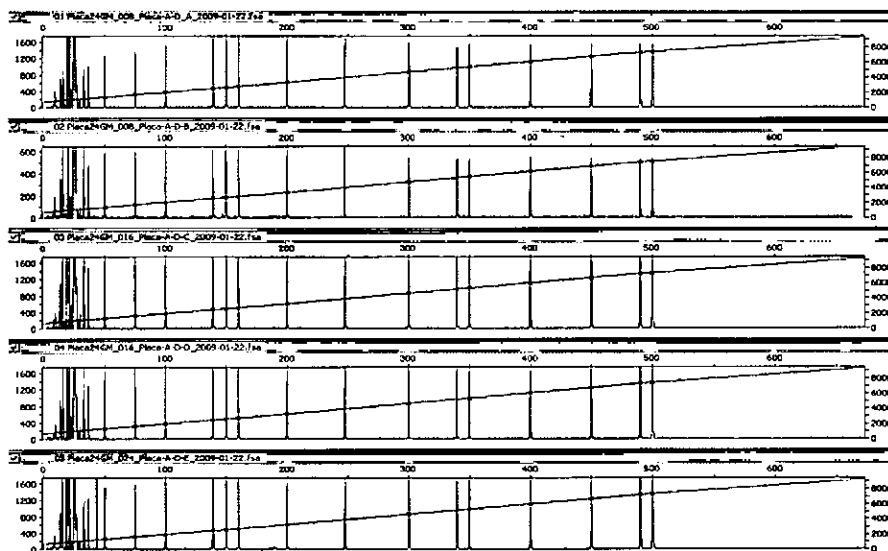


Figura 7. Cada vez que se analiza un nuevo lote de reacciones de amplificación PCR, se calibra el sistema utilizando mezclas de fragmentos de DNA de longitud conocida. Su posición respecto al eje de abscisas sirve para determinar con exactitud el tamaño de los fragmentos amplificados. Los paneles que se muestran corresponden a cinco lotes distintos. La reproducibilidad del sistema es evidente.

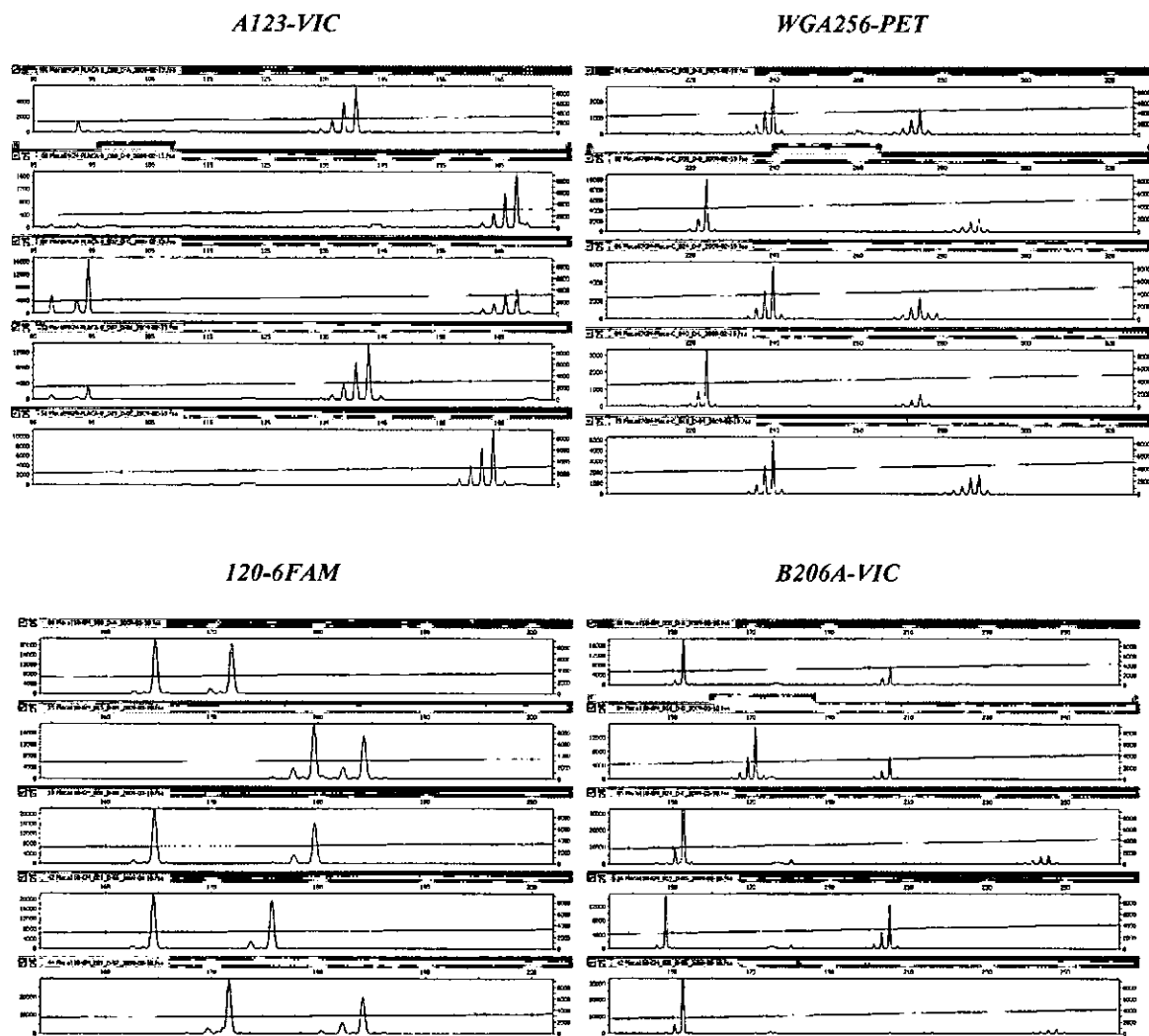


Figura 8. Las cuatro parejas de *primers* mostradas aquí resultaron ser particularmente polimórficas: A123, B120, B206a y WGA256.

La **Tabla IV** recoge la selección definitiva de *primers* nuevos utilizada en este proyecto. Nótese que todos se han desarrollado *de novo* a partir de genotecas enriquecidas, construidas específicamente para este fin. Nótese también que estos marcadores, 24 en total, son el resultado de un largo proceso de selección empírica que se inició con la caracterización detallada de cerca de 700 secuencias nucleotídicas.

Tabla IV. Marcadores moleculares (parejas de *primers*) utilizados para el genotipado definitivo de los nogales en estudio.

Nombre	Secuencia 5→3'(°)	Patrón repetido	Longitud (bp)	Temperatura óptima (°C)	Alelos obs.
A103	F AAGAGGGTGATTCCTCAC R GTAAAGCACAGTTTCATGTAAG	(TA) ₅ (CA) ₃₄	114-244	60.0	3
A123	F CACCTTTGCATCTTTTGTTCAC R AGGGTTTTCAAGTGTCGTTTTTC	(CA) ₂₆ (TA) ₂	138-168	60.0	5
A128	F TGGAAAGTGAAGCAATATGTTTC R CCTTTGGAATCGTACTACAAA	(AT) ₄ (GT) ₁₂	148-202	60.0	3
A201	F GCATCCAAAGGTAAGTTAGAAG R TGGACCCCTAGAACACAACA	(GT) ₁₉ TT(GT) ₂ T(GT) ₂	178-186	55.0	3
A214	F TAGCAGCGTCCCATAAATCG R GCACTCTCGAATTCGCAAAC	(CT) ₉ (CA) ₃ CC(CA) ₁₇	183-212	55.0	4
A221	F AGCCACGATACAAACACAAC R AGGATAAAATGCTTCGGAATAG	(CA) ₉	196-208	58.0	3
A222	F AACACCACCACCCTAAGTTC R TTCGTCCTTCCCTCACTAAG	(CA) ₁₂	149-155	58.0	2
A227	F GGTAGGGTCATTGGAACATAG R GGGGTACTCTGGAGACATATAG	(CT) ₁₅ (CA) ₁₄	155-167	58.0	3
A230	F AGTGGCAATCTCAAGGAAG R TCTCCGTTCATCTCATTCTC	(GT) ₁₂ (GA) ₁₇ GCAA(GA) ₃	160-163	55.0	2
A243	F TTGCATCCAAATAGCTGTCC R TTGCTGGAACTACTGCATC	(CA) ₈ CG(CA) ₁₁	157-172	55.0	3
B102	F TGACATGGTGTTCATAACTCAG R GTCAACAGGCAAGAACTAGAC	(CT) ₁₇ (CCCT) ₃	189-211	55.0	3
B105	F TAGCCTCGTTCAGGTTCTATA R CCTAGAGGTAGATGCGGTAGAG	(CT) ₉ CCACGTTAT(CT) ₂₃	184-211	62.0	3
B107	F TAGCCAACTTTGTTTGACTG R TCGGACACAACCTACAACCAC	(CT) ₂₇	171-205	60.0	6
B111	F GCCTGTGGTCAAGTGTCA R GACGGGAAGAAATCAAG	(CT) ₂₂	181-195	55.0	4
B120	F AGGTGGGAATATGACACTGTAC R AATCGTGGTGTCTATTAGATTTC	(GA) ₁₅	165-197	55.0	6
B135	F CATTTCAGCCCTTTGAAGTTC R ATAGCGTAACCAGCTCCTTG	(GT) ₇	184-191	55.0	3
B206a	F ATCGTCTCTCTCAGTGTCTC R AACCGTCTATATCAGTCTCAGC	(CT) ₁₈	148-246	60.0	7
B210	F GCAAGCACGATAGAAAGGAC R GGGATACTCTGATGCCAAG	(CT) ₂₅	195-221	60.0	4
B216	F TCCCTTAGCATCTATGAAGTTC R GATGGTGAATGTTCCGTACTG	(CT) ₁₄ (CA) ₁₄	127-151	57.0	4
B222	F TAACATACACAGCAACACAG R GAGGGCACACCCTAAG	(CA) ₂ (GA) ₂₁	130-155	55.0	5
B224	F CCAGCAGGAGAGTCTTCG R AGCACACCATAGAGAGAAAACA	(CT) ₁₄	114-125	55.0	3
B227a	F ATGCGTGCTTATCTGTTTTC R TCAAGTGTGGATGGGACTATC	(CT) ₂₆ (CA) ₉	197-239	55.0	5
B233	F GGACTACGGGATTTAGTG R CTGGTCCCGTTTAGGTA	(GT) ₇ (AG) ₁₅	150-171	55.0	4
B236	F CAGTCTCTCTTTCTTTTC R GCCTCTTCGTATCTGTTTCTC	(CT) ₂₃	160-174	55.0	4

* F = forward primer; R = reverse primer

2) Incorporación de marcadores de *Juglans cinerea*

También hemos incluido en nuestros análisis marcadores de *J. cinerea* publicados previamente. Destaca en este sentido un artículo de Ross-Davis y Woeste aparecido poco después de iniciarse el proyecto (Microsatellite markers for *Juglans cinerea* L. and their utility in other *Juglandaceae* species. *Conservation Genetics* 9: 465–469, 2008). Decidimos incorporar estos marcadores al proyecto por dos razones fundamentales: (1) nuestra capacidad de genotipar podría verse incrementada notablemente y eran de libre uso; y (2) tendríamos además un buen sistema para comparar la capacidad discriminante de nuestros propios marcadores. Esta comparación es importante en el contexto del proyecto.

Así, utilizando los mismos protocolos y muestras de DNA genómico que en los apartados anteriores, procedimos a la evaluación de los *primers* de *J. cinerea*. Siete de ellos arrojaron resultados de buena calidad (**Tabla V**), mientras que los restantes no funcionaron bien con los híbridos de la empresa.

Tabla V. Marcadores moleculares de *J. cinerea* seleccionados para el genotipado definitivo de nogal, por su buen comportamiento con el material de la empresa.

Nombre	Secuencia 5→3'(*)	Referencia	Longitud (bp)	Temperatura óptima (°C)	Alelos obs.
WGA004	F TGTTCATTGACCCACTTGT R TAAGCCAACATGGTATGCCA	Woeste et al. 2002	229–261	60.0	4
WGA082	F TGCCGACACTCCTCACTTC R CGTGATGTACGACGGCTG	Woeste et al. 2002	171–189	60.0	4
WGA090	F CTTGTAATCGCCCTCTGCTC R TACCTGCAACCCGTTACACA	Victory et al. 2006	134–180	60.0	2
WGA147	F TGGAACTIGTTCTGTGCGAG R CCGAGTCCCTTCACATCTA	Ross-Davis et al. 2008	175–199	60.0	4
WGA148	F GGTGAACICCCATAGGGGTA R CCAATGCTACTTGAGAACC	Ross-Davis et al. 2008	247–261	60.0	2
WGA204	F GGGTCTCGCCTTCTTTCTT R CACAGAGAGAAGCACGGGTA	Ross-Davis et al. 2008	168–187	60.0	3
WGA256	F TGAAGACAACAAACTGCCG R CCGCATTGTTTCTGAAAAT	Ross-Davis et al. 2008	224–289	60.0	5

La **Tabla VI** resume los resultados obtenidos al analizar 50 ejemplares de nogal híbrido, *J. major 209 x regia*, con los marcadores descritos en las Tablas IV y V. Los nogales se seleccionaron al azar en las fincas de la empresa en Cáceres. Se indican en la tabla los tamaños y las frecuencias de todos los alelos detectados. La probabilidad combinada de identidad al utilizar los 31 marcadores es del orden de 10^{-14} (**Tabla VII**). Esto significa que, en promedio, dos de cada cien billones de nogales compartirán por azar la misma huella genética. En términos prácticos ello implica que el sistema que hemos desarrollado permite discriminar ejemplares individuales de nogal, a pesar de que todos descienden de la misma madre y muchos del mismo padre. Los objetivos de esta tarea se han cumplido con creces.

Tabla VI. Resumen de resultados del análisis de nogal, desglosados por *locus* y alelo individual.

LOCUS	TAMAÑO ALELO (pb)	FRECUENCIA ALÉLICA			
				157	0.260
			A243-NED	165	0.260
				172	0.480
A103-NED	114	0.070		189	0.500
	198	0.630		196	0.310
	244	0.300	B102-6FAM	211	0.190
	138	0.020		184	0.380
A123-VIC	140	0.550		205	0.020
	143	0.020	B105-NED	211	0.600
	164	0.020			
	168	0.390		171	0.020
				191	0.550
A128-PET	148	0.310		193	0.040
	198	0.290	B107-VIC	201	0.360
	202	0.400		203	0.010
				205	0.020
A201-NED	178	0.060			
	180	0.460		181	0.560
	186	0.480		189	0.020
			B111-PET	191	0.060
A214-PET	183	0.490		195	0.360
	185	0.010			
	189	0.330		165	0.210
	212	0.170		172	0.280
				176	0.010
A221-6FAM	196	0.500	B120-6FAM	180	0.200
	202	0.230		184	0.290
	208	0.270		197	0.010
A222-NED	149	0.395		184	0.370
	155	0.604	B135-6FAM	189	0.460
				191	0.170
A227-VIC	155	0.280			
	163	0.440		148	0.020
	167	0.280		153	0.230
			B206a-VIC	171	0.250
A230-PET	160	0.448		197	0.010
	163	0.551		205	0.470

	246	0.010
	248	0.010
	195	0.020
	209	0.480
B210-PET	219	0.040
	221	0.460
	127	0.310
	141	0.370
B216-NED	149	0.130
	151	0.190
	130	0.193
	132	0.010
B222-VIC	139	0.551
	140	0.224
	155	0.020
	114	0.180
	120	0.060
B224-PET	125	0.760
	197	0.240
	203	0.260
B227a-6FAM	217	0.220
	235	0.040
	239	0.240
	150	0.330
	153	0.180
B233-VIC	156	0.250
	171	0.240
	160	0.240
	166	0.260
B236-PET	172	0.220
	174	0.280
	229	0.020
	231	0.480
WGA004-NED	235	0.370
	261	0.130
	171	0.210
	177	0.210
WGA082-6FAM	185	0.560
	189	0.020
	134	0.480
WGA090-6FAM	180	0.520
	175	0.010
	183	0.880
WGA147-VIC	191	0.100
	199	0.010
	247	0.480
WGA148-NED	261	0.520
	168	0.250
	176	0.500
WGA204-VIC	187	0.250
	224	0.220
	240	0.280
WGA256-PET	275	0.230
	285	0.020
	289	0.250

Tabla VII. Heterozigosidad observada y esperada, frecuencia estimada de alelos nulos, probabilidad de exclusión de paternidad y probabilidad de identidad tras analizar una amplia muestra de nogales híbridos seleccionados al azar (n=50). La probabilidad global de identidad (PI) y de exclusión de paternidad (PEP) al combinar todos los marcadores se indica al final.

Locus	Heterozigosidad Esp - Obs	Frecuencia de alelos nulos	Probabilidad de exclusión de paternidad	Probabilidad de identidad
A103-NED	0.508-0.160	0.230	0.245	0.470
A123-VIC	0.544-0.200	0.339	0.258	0.487
A128-PET	0.659-0.580	0.048	0.365	0.337
A201-NED	0.554-0.000	0.356	0.255	0.500
A214-PET	0.622-1.000	-0.233	0.340	0.360
A221-6FAM	0.624-1.000	-0.231	0.340	0.353
A222-NED	0.478-0.000	0.323	0.181	0.615
A227-VIC	0.649-0.560	0.054	0.358	0.341
A230-PET	0.494-0.000	0.331	0.186	0.622
A243-NED	0.634-0.520	0.069	0.347	0.347
B102-6FAM	0.617-1.000	-0.236	0.333	0.365
B105-NED	0.495-0.000	0.331	0.205	0.567
B107-VIC	0.565-0.020	0.348	0.291	0.430
B111-PET	0.552-0.000	0.356	0.273	0.454
B120-6FAM	0.753-1.000	-0.140	0.515	0.193
B135-6FAM	0.622-0.340	0.174	0.332	0.376
B206a-VIC	0.663-1.000	-0.202	0.400	0.287
B210-PET	0.556-0.000	0.357	0.260	0.494
B216-NED	0.714-1.000	-0.166	0.459	0.238
B222-VIC	0.607-0.326	0.174	0.350	0.326
B224-PET	0.386-0.040	0.249	0.192	0.501
B227a-6FAM	0.767-1.000	-0.131	0.538	0.176
B233-VIC	0.738-1.000	-0.150	0.490	0.212
B236-PET	0.748-1.000	-0.144	0.501	0.204
WGA004-NED	0.615-1.000	-0.238	0.333	0.375
WGA082-6FAM	0.597-0.420	0.111	0.337	0.340
WGA090-6FAM	0.499-0.000	0.332	0.187	0.624
WGA147-VIC	0.215-0.020	0.160	0.104	0.663
WGA148-NED	0.499-0.000	0.332	0.187	0.624
WGA204-VIC	0.625-1.000	-0.230	0.341	0.351
WGA256-PET	0.757-0.960	-0.115	0.520	0.189

Probabilidad de Exclusión de Paternidad (PEP): 0.999996

Probabilidad de Identidad (PI): 6.258×10^{-14}

HITOS - TAREA 3

HITOS ACORDADOS EN EL CONVENIO (Anexo I)

- Desarrollar un sistema eficaz de tipificación molecular para nogal basado en marcadores microsatélite (genotipado)
- Se consideran prioritarios en esta tarea (1) simplificar al máximo el análisis y (2) minimizar el coste total por muestra

HITOS CONSEGUIDOS

Se ha conseguido plenamente el objetivo. Hemos desarrollado y optimizado un sistema sencillo, reproducible y eficiente en todos sus términos, costes incluidos, para genotipar con fiabilidad el material de la empresa. Todo el análisis se ha informatizado y miniaturizado, permitiendo el procesamiento simultáneo de un gran número de muestras.

De los 31 microsatélites seleccionados en este apartado, 24 se han desarrollado *de novo* y 7 proceden de la literatura científica sobre *J. cinerea*. Su uso combinado permite genotipar ejemplares individuales de nogal con una fiabilidad comparable a la obtenida en el ámbito forense. La probabilidad de identidad global al combinar todos los marcadores es del orden de 10^{-14} , extraordinariamente pequeña.

Por lo que respecta a nogal, todas las tareas previstas en la memoria técnica del proyecto han culminado con éxito.

Tarea 4: Genotipado de material clonal seleccionado (cerezo)

El desarrollo de esta tarea ha sido prácticamente idéntico al descrito para nogal, excepto que se han utilizado microsatélites ya descritos de libre uso, accesibles a través de la literatura científica. Igual que en el caso de nogal, para la selección definitiva de *loci* microsatélite se han tenido muy presentes la consistencia de los resultados y el polimorfismo observado al analizar el material de cerezo.

Para los análisis genotípicos se ha utilizado una muestra de 27 ejemplares de cerezo, cuatro de ellos clones, de dos procedencias diferentes. La extracción de DNA genómico se realizó con el mismo kit que en el caso de nogal (DNAeasy Plant Mini Kit, de QIAGEN), obteniéndose resultados comparables en cantidad y calidad. Para las amplificaciones PCR se ensayaron distintas condiciones, obteniéndose los mejores resultados con el siguiente programa de 35 ciclos:

	previo:	94°C 1 min
35x	(paso 1)	94°C 30"
	(paso 2)	T _{opt} °C 30" (T _{opt} : temperatura óptima, según <i>primers</i>)
	(paso 3)	72°C 30"
	final:	72°C 20 min

La **Tabla VIII** presenta nuestro listado inicial de *primers* a evaluar. Para su elaboración tuvimos en cuenta sobre todo la capacidad de dichos *primers* para detectar polimorfismo, a tenor de la información disponible en la literatura científica. Hay muchos otros marcadores descritos para el género *Prunus*. Las referencias bibliográficas concretas son:

1. Cipriani G., Lot G., Huang W. G., Marrazzo M. T., Peterlunger E., Testolin R. 1999. AC/GT and AG/CT microsatellite repeats in peach [*Prunus persica* (L) Batsch]: isolation, characterization and cross-species amplification in *Prunus*. *Theor Appl Genet.* 99: 65-72.
2. Dirlewanger E., Cosson P., Tavaud M., Aranzana M.J., Poizat C., Zanetto A., Arús P., Laigret F. 2002. Development of microsatellite markers in peach [*Prunus persica* (L) Batsch] and their use in genetic diversity analysis in peach and sweet cherry (*Prunus avium* L.). *Theor Appl Genet.* 105: 127–138.

3. Sosinski B., Gannavarapu M., Hager L.D., Beck L.E., King G.J., Ryder C.D., Rajapakse S., Baird W.V., Ballard R.E., Abbott A.G. 2000. Characterization of microsatellite markers in peach [*Prunus persica* (L.) Batsch]. *Theor Appl Genet.* 101: 421–428.
4. Struss D., Ahmad R., Southwick S. M. 2003. Analysis of sweet cherry (*Prunus avium* L.) cultivars using SSR and AFLP markers. *J. Amer. Soc. Hort. Sci.* 128: 904-909.
5. Testolin R., Marrazzo T., Cipriani G., Quarta R., Verde I., Dettori M. T., Pancaldi M., Sansavini S. 2000. Microsatellite DNA in peach [*Prunus persica* (L.) Batsch] and its use in fingerprinting and testing the genetic origin of cultivars. *Genome* 43: 512–520.

Tabla VIII. Relación de *loci* microsatélite seleccionados inicialmente para el genotipado de cerezo. Se indica en cada caso la referencia bibliográfica y la especie de *Prunus* en que fueron descritos.

LOCUS	SECUENCIA	ESPECIE	REFERENCIA
UDP96-003 ^{NA}	F: TTG CTC AAA AGT GTC GTT GC R: ACA CGT AGT GCA ACA CTG GC	<i>P. persica</i>	Cipriani <i>et al.</i> (1999).
UDP96-013 ^{NA}	F: ATT CTT CAC TAC ACG TGC ACG R: CCC CAG ACA TAC TGT GGC TT	<i>P. persica</i>	Cipriani <i>et al.</i> (1999).
UDP96-015 ^{NA}	F: CCT TGA CCT ATT TGT TCG TCA R: ACT AGT CAA ACA ATC CCC CG	<i>P. persica</i>	Cipriani <i>et al.</i> (1999).
UDP96-018	F: TTC TAA TCT GGG CTA TGG CG R: GAA GTT CAC ATT TAC GAC AGG G	<i>P. persica</i>	Cipriani <i>et al.</i> (1999).
UDP98-407 ^{NA}	F: AGC GGC AGG CTA AAT ATC AA R: AAT CGC CGA TCA AAG CAA C	<i>P. persica</i>	Cipriani <i>et al.</i> (1999).
UDP96-001	F: AGT TTG ATT TTC TGA TGC ATC C R: TGC CAT AAG GAC CGG TAT GT	<i>P. persica</i>	Testolin <i>et al.</i> (2000)
UDP96-021	F: AAG CAG CAA TTG GCA GAA TC R: GAA TAT GAG ACG GTC CAG AAG C	<i>P. persica</i>	Testolin <i>et al.</i> (2000)
UDP98-410	F: AAT TTA CCT ATC AGC CTC AAA R: TTT ATG CAG TTT ACA GAC CG	<i>P. persica</i>	Testolin <i>et al.</i> (2000)
BPPCT 002	F: TCG ACA GCT TGA TCT TGA CC R: CAA TGC CTA CGG AGA TAA AAG AC	<i>P. persica</i>	Dirlewanger <i>et al.</i> (2002)
BPPCT 005	F: GCT AGC AGG GCA CTT GAT C R: ACG CGT GTA CGG TGG AT	<i>P. persica</i>	Dirlewanger <i>et al.</i> (2002)
BPPCT 026	F: ATA CCT TTG CCA CTT GCG R: TGA GTT GGA AGA AAA CGT AAC A	<i>P. persica</i>	Dirlewanger <i>et al.</i> (2002)
BPPCT 034 ^{NA}	F: CTA CCT GAA ATA AGC AGA GCC AT R: CAA TGG AGA ATG GGG TGC	<i>P. persica</i>	Dirlewanger <i>et al.</i> (2002)
BPPCT 037	F: CAT GGA AGA GGA TCA AGT GC R: CTT GAA GGT AGT GCC AAA GC	<i>P. persica</i>	Dirlewanger <i>et al.</i> (2002)
BPPCT 038	F: TAT ATT GTT GGC TTC TTG CAT G R: TGA AAG TGA AAC AAT GGA AGC	<i>P. persica</i>	Dirlewanger <i>et al.</i> (2002)

BPPCT 039	F: ATT ACG TAC CCT AAA GCT TCT GC R: GAT GTC ATG AAG ATT GGA GAG G	<i>P. pérsica</i>	Dirlewanger <i>et al.</i> (2002)
BPPCT 040 ^{NA}	F: ATG AGG ACG TGT CTG AAT GG R: AGC CAA ACC CCT CTT ATA CG	<i>P. pérsica</i>	Dirlewanger <i>et al.</i> (2002)
UCD-CH12	F: TTT CTG CCA CAA ACC TAA TGG R: AGA CAA AGG GAT TGT GGG C	<i>P. avium</i>	Struss <i>et al.</i> (2003)
UCD-CH18	F: ATG GAA GGC CAA GGC AAC R: AAT GTT CCC GGT TAT ATG C	<i>P. avium</i>	Struss <i>et al.</i> (2003)
PCHGMS1	F: GGG TAA ATA TGC CCA TTG TGC AAT C R: GGA TCA TTG AAC TAC GTC AAT CCT C	<i>P. avium</i>	Sosinski <i>et al.</i> (2000).

NA: no amplifica; F: *primer forward*; R: *primer reverse*

Tras eliminar las parejas de *primers* que no amplificaban o lo hacían de forma inconsistente, seleccionamos un total de 14 *loci* microsatélite para el genotipado definitivo de los cerezos de la empresa. La **Tabla IX** resume los resultados obtenidos al analizar con estos marcadores los 27 genotipos antes mencionados. Nótese que el número de alelos observados para la mayoría de los *loci* supera al número de alelos descrito en la literatura, lo que avala la idoneidad de nuestra selección. También confirma el elevado polimorfismo de estos marcadores, gracias al cual el objetivo último de la tarea se ha logrado: poder diferenciar ejemplares individuales con una fiabilidad muy elevada.

La heterogeneidad genética de la población en estudio ha resultado ser mucho mayor que la de nogal, por lo que no ha sido necesario utilizar marcadores adicionales. La probabilidad combinada de identidad con los 14 *loci* utilizados aquí es del orden de 10^{-14} , lo que significa que, en promedio, dos de cada cien billones de cerezos compartirán por azar la misma huella genética. En términos prácticos, esto significa que el sistema que hemos puesto a punto sirve perfectamente para diferenciar ejemplares individuales de cerezo propiedad de la empresa (**Tabla X**).

Tabla IX (página siguiente). Resultados del genotipado de cerezos de la empresa con los 14 *loci* microsatélite que hemos seleccionado. Se han reflejado en la tabla algunos datos tomados de la literatura, para facilitar el análisis comparado.

LOCUS	SECUENCIA DE LOS PRIMERS (5'→3')	REPETICIÓN	T _e TEOR. (°C)	T _e OBS. (°C)	TAMAÑO ALÉLICO TEOR. (bp)	TAMAÑO ALÉLICO OBS. (bp)	ALELOS CONOCIDOS	ALELOS OBSERVADOS
UDP96-001	F: AGT TTG ATT TTC TGA TGC ATC C R: TGC CAT AAG GAC CGG TAT GT	(CA)17	55-63	57	120	107-135	2	9
UDP96-018	F: TTC TAA TCT GGG CTA TGG CG R: GAA GTT CAC ATT TAC GAC AGG G	(AC)21	55-63	57	253	241	3	1
UDP97-402	F: TCCCATAACCAAAAAAACACC R: TGGAGAAGGGTGGGTA CTG	(AG)17	55-63	60	136	113-143	2	9
UDP96-021	F: AAG CAG CAA TTG GCA GAA TC R: GAA TAT GAG ACG GTC CAG AAG C	(GA)22(CA)11	57	57	146-157	101-117	6	6
UDP98-410	F: AAT TTA CCT ATC AGC CTC AAA R: TTT ATG CAG TTT ACA GAC CG	(AG)23	57	54	146-182	123-135	8	6
BPPCT 002	F: TCG ACA GCT TGA TCT TGA CC R: CAA TGC CTA CGG AGA TAA AAG AC	(AG)25	57	57	179-185	167-187	4	7
BPPCT 005	F: GCT AGC AGG GCA CTT GAT C R: ACG CGT GTA CGG TGG AT	(AG)10	57	57	157-199	140-198	6	10
BPPCT 026	F: ATA CCT TTG CCA CTT GCG R: TGA GTT GGA AGA AAA CGT AAC A	(AG)8GG(AG) 6	57	54	140-190	163-185	6	5
BPPCT 037	F: CAT GGA AGA GGA TCA AGT GC R: CTT GAA GGT AGT GCC AAA GC	(GA)25	57	57	142-156	133-170	5	10
BPPCT 038	F: TAT ATT GTT GGC TTC TTG CAT G R: TGA AAG TGA AAC AAT GGA AGC	(GA)25	57	57	101-133	98-133	5	10
BPPCT 039	F: ATT ACG TAC CCT AAA GCT TCT GC R: GAT GTC ATG AAG ATT GGA GAG G	(GA)20	57	57	134-150	130-147	5	5
UCD-CH12	F: TTT CTG CCA CAA ACC TAA TGG R: AGA CAA AGG GAT TGT GGG C	(CA)14	55-60	57	173-200	175-197	5	9
UCD-CH18	F: ATG GAA GGC CAA GGC AAC R: AAT GTT CCC GGT TAT ATG C	(CT)23	55-60	54	178-188	182-211	4	10
PCHGMS1	F: GGG TAA ATA TGC CCA TTG TGC AAT C R: GGA TCA TTG AAC TAC GTC AAT CCT C	(AC) ₁₂ (AT) ₆	50-60	57	194	121-151	4	6

Tabla IX. Relación de *loci* microsatélite seleccionados inicialmente para el genotipado de cerezo. Se indica en cada caso la referencia bibliográfica y la especie

Tabla X. Resumen de resultados del análisis de cerezo, desglosados por *locus* y alelo individual.

LOCUS	TAMAÑO ALELICO (bp)	FRECUENCIA ALÉLICA	PEP	PI
UDP96-001	107	0.019		
	109	0.241		
	110	0.037		
	114	0.019		
	115	0.019		
	118	0.056		
	124	0.278		
	126	0.315		
	135	0.019		
				0.541
UDP96-018	241	1.000		
			0.000	1.000
UDP97-402	113	0.023		
	117	0.023		
	121	0.409		
	123	0.205		
	125	0.023		
	133	0.068		
	137	0.023		
	139	0.023		
	143	0.205		
				0.529
UDP96-021	101	0.120		
	103	0.260		
	105	0.300		
	113	0.040		
	114	0.200		
	117	0.08		
				0.571
UDP98-410	123	0.037		
	125	0.056		
	127	0.519		
	129	0.352		
	135	0.019		
	137	0.019		
				0.336
BPPCT 002	167	0.056		
	168	0.111		

	179	0.204		
	181	0.463		
	183	0.037		
	185	0.111		
	187	0.019		
			0.503	0.115
BPPCT 005	140	0.056		
	146	0.093		
	153	0.019		
	155	0.278		
	159	0.148		
	161	0.019		
	163	0.315		
	183	0.019		
	196	0.019		
	198	0.037		
			0.596	0.074
BPPCT 026	163	0.407		
	169	0.315		
	177	0.056		
	183	0.204		
	185	0.019		
			0.429	0.153
BPPCT 037	133	0.037		
	135	0.167		
	139	0.259		
	141	0.056		
	145	0.204		
	152	0.120		
	154	0.019		
	156	0.019		
	162	0.056		
	170	0.056		
			0.676	0.047
BPPCT 038	98	0.222		
	108	0.185		
	115	0.148		
	117	0.019		
	119	0.074		
	121	0.037		
	123	0.130		
	129	0.111		
	131	0.056		
	133	0.019		
			0.709	0.038

BPPCT 039	130	0.063		
	134	0.417		
	140	0.125		
	145	0.250		
	147	0.146		
			0.492	0.119
UCD-CH12	175	0.259		
	179	0.148		
	181	0.241		
	187	0.148		
	189	0.093		
	191	0.019		
	193	0.019		
	195	0.056		
	197	0.019		
			0.642	0.057
UCD-CH18	182	0.288		
	184	0.058		
	186	0.173		
	188	0.077		
	190	0.058		
	192	0.212		
	194	0.038		
	196	0.019		
	201	0.019		
	211	0.058		
			0.659	0.052
PCHGMS1	121	0.060		
	133	0.080		
	141	0.480		
	143	0.060		
	145	0.300		
	151	0.020		
			0.424	0.162

Probabilidad de Exclusión de Paternidad (PEP): 0.99998

Probabilidad de Identidad (PI): 2.433×10^{-14}

HITOS - TAREA 4

HITOS PROPUESTOS EN EL CONVENIO (Anexo I)

- Desarrollar un sistema eficaz de tipificación molecular para cerezo basado en marcadores microsatélite (genotipado)
- Como en el caso anterior, se consideran prioritarios 1) simplificar al máximo el análisis y 2) minimizar el coste total por muestra

HITOS CONSEGUIDOS:

Como en el caso de nogal, esta tarea ha culminado con éxito. Hemos utilizado esencialmente la metodología puesta a punto para nogal (tareas 1 a 3), con pequeñas variaciones en las condiciones de PCR debidas a la naturaleza específica de los *primers*. El sistema de genotipado es, en ambos casos, sencillo, barato, reproducible y eficiente. Procesar un gran número de muestras no supone ningún problema.

El uso combinado de los 14 microsatélites seleccionados en esta tarea permite, como en el caso de nogal, genotipar ejemplares individuales de cerezo con una fiabilidad muy elevada: la probabilidad de que dos cerezos tengan la misma huella genética por azar es del orden de 10^{-14} . Es reseñable, en este sentido, el elevado polimorfismo que han detectado los marcadores seleccionados en los cerezos plantados por la empresa.

Validación metodológica

Todas las reacciones de amplificación se han repetido al menos cuatro veces, de forma independiente, para detectar y, en su caso, eliminar posibles errores de procedimiento: contaminaciones en las muestras, fallos durante la amplificación o el fraccionamiento u otros problemas. Previamente habíamos realizado experimentos encaminados a cuantificar la fiabilidad global del procedimiento en nogal. Para ello elegimos veinte muestras de DNA genómico, procedentes de otros tantos ejemplares individuales seleccionados al azar, y subdividimos cada una en cinco viales idénticos. A continuación genotipamos cada vial de forma independiente con nuestros marcadores –los de *cinerea* habían sido evaluados con anterioridad.

Realizamos en conjunto 2.100 reacciones de tipificación (21 *loci*; 20 x 5 genotipos). Los resultados se analizaron según el procedimiento estadístico descrito por Broquet y Petit en *Molecular Ecology* 13: 3601–3608 (2004). Las frecuencias determinadas para los parámetros ADO (*allele dropout*) y FA (*false allele*) oscilaron entre 0.000 y 0.018 para el primero (según el marcador concreto) y 0.000 y 0.021 para el segundo. ADO estima la probabilidad de no amplificar un alelo dado en un locus heterocigoto (nuestros *loci* lo son necesariamente) y FA la probabilidad de obtener falsos alelos debido a errores de la DNA polimerasa o el sistema de detección. Estos experimentos revelaron que la probabilidad de obtener un genotipo consenso es ya muy elevada con tres amplificaciones independientes (exactamente del 97.3%), incrementándose este valor de forma paulatina con el número de repeticiones. Hacer más de cuatro réplicas independientes de las reacciones de PCR (probabilidad >98.9%) es claramente innecesario. La fiabilidad de nuestro procedimiento global de genotipado es por tanto muy alta y comparable a la publicada en revistas científicas de calidad.

Resultados más relevantes

Los resultados más relevantes de este proyecto de I+D pueden resumirse así:

1.- Respecto a nogal, nuestros datos globales sugieren que ha intervenido un número reducido de progenitores para generar las semillas de las que derivan todos los híbridos de Bosques Naturales. Salvo excepciones, los *loci* seleccionados detectan sistemáticamente un nivel de polimorfismo moderado. La asignación de parentesco se ha contrastado con diversas líneas puras e híbridas de nogal facilitadas por la Dra. Neus Aletá (IRTA), una de las mayores expertas europeas en nogal.

2.- Los 24 marcadores desarrollados en este proyecto y los 7 marcadores procedentes de *J. cinerea* detectan niveles similares de polimorfismo en la muestra estudiada. Nótese que estos últimos han sido seleccionados precisamente por su elevada variabilidad (Ross-Davis y Woeste, 2008), lo que avala la calidad y el potencial aplicado de los nuevos marcadores. También refuerza las conclusiones del punto anterior respecto a la heterogeneidad genética de los nogales de la empresa.

3.- Mediante el análisis exhaustivo de un número inusualmente elevado de marcadores e innovaciones experimentales no previstas inicialmente, hemos podido soslayar la falta de heterogeneidad y alcanzar el objetivo principal del convenio: desarrollar un sistema para genotipar nogales a nivel individual con una fiabilidad muy elevada. Así, nuestros análisis exhaustivos de 50 genotipos diferentes seleccionados al azar permite afirmar que:

No hay dos ejemplares de nogal híbrido (es decir, dos genotipos) que muestren el mismo patrón alélico. Cada patrón observado, por tanto, tiene valor diagnóstico en la muestra analizada.

Utilizando conjuntamente todos los marcadores, incluidos los de *J. cinerea*, la probabilidad de identidad es extraordinariamente baja, del orden de 10^{-14} .

Téngase en cuenta que la población en estudio está compuesta por hermanos de una misma madre y, en una proporción elevada de casos, de un mismo padre (*full-siblings*). La probabilidad total de exclusión es del 99,99%.

- En resumen, el sistema de genotipado que hemos desarrollado permite distinguir entre genotipos muy relacionados, incluso *full-siblings*. Ya se ha señalado que no hay dos ejemplares con el mismo patrón alélico en la población analizada. Es importante recalcar que no hay diferencias significativas entre los *loci* de *cinerea* y los identificados en el proyecto. Nótese de nuevo que los primeros se han evaluado extensamente y se han seleccionado por su gran potencial discriminante.

4.- Hemos conseguido resultados comparables en el caso de cerezo, utilizando esencialmente la misma aproximación metodológica. La gran diferencia estriba en que partimos de marcadores descritos en la literatura científica. Tras una evaluación inicial con material de Bosques Naturales, seleccionamos una colección de 14 marcadores polimórficos. Su uso combinado también rinde una probabilidad de identidad en el rango de 10^{-14} . La probabilidad total de exclusión de paternidad es asimismo muy baja (99,99%). Ambos datos son perfectamente comparables a los obtenidos con nogal. Nuestros resultados avalan la transferibilidad de la metodología utilizada, incluso entre especies alejadas.

Conclusiones generales

Hemos desarrollado un sistema de tipificación molecular con un rendimiento excelente y también un protocolo experimental eficiente en todos sus aspectos, costes incluidos. La validez de este sistema de tipificación se ha puesto de manifiesto reiteradamente en este informe, tanto para nogal como para cerezo. No sólo lo avalan los índices cuantitativos calculados a partir de los datos experimentales, sino también la evidencia experimental que se muestra en figuras representativas. Además de reproducibilidad y especificidad, la capacidad discriminante alcanzada es más que suficiente para tipificar individuos –material clonal de élite, por ejemplo. La probabilidad de coincidencia por azar en los patrones genotípicos se ha cuantificado con precisión y es extremadamente remota en nogal y cerezo.

En este proyecto se han descrito y analizado por vez primera veinticuatro marcadores moleculares de nogal. El procedimiento experimental que hemos utilizado incluye innovaciones no previstas en el convenio inicial. Las hemos introducido por varios motivos, siendo uno de ellos la homogeneidad genética del material de nogal. Otro es su carácter innovador en el ámbito forestal. Es indudable que en la polinización del clon materno, MJ209, han intervenido pocas variedades, una conclusión avalada por nuestros análisis de nogales puros e híbridos del IRTA. Cuando se utilizan sólo los marcadores de *J. cinerea*, seleccionados por su gran capacidad discriminante, se alcanza la misma conclusión. Nuestra aproximación al genotipado de nogal es perfectamente transferible a cerezo, habiéndose alcanzado resultados comparables en cuanto a capacidad discriminante y robustez del análisis.

En suma, se han conseguido todos los objetivos planteados en la memoria técnica del proyecto.


* Pdo. Luis Gómez Fernández
Director del Proyecto

NOGAL

Genotecas enriquecidas más del 60%: Dos

Nivel real de enriquecimiento: 80% de insertos contienen microsatélites

Nivel real de redundancia: <5%

Tamaño promedio de inserto: 240 pares de bases

Secuenciación primaria: 2 x 680 lecturas (ambas hebras del DNA)

Diseño inicial de *primers*: 360 parejas, 180 por genoteca

Selección definitiva de *primers*: 24 parejas *de novo*; 7 parejas de *J. cinerea*

Cantidad de DNA necesaria: 5 ng por análisis y muestra

Tamaño muestral analizado: 50 ejemplares individuales de nogal híbrido

Máximo número de alelos por locus: 7

Ratio ADO (allele dropout): 0.000 - 0.018, según *locus*

Ratio FA (false allele): 0.000 - 0.021, según *locus*

Réplicas por reacción de genotipado: 4 independientes

Probabilidad de obtener genotipo consenso: >99%

Probabilidad total de exclusión: 0.999

Probabilidad de identidad: 6.258×10^{-14} (con todos los marcadores)

CEREZO

Diseño inicial de *primers*: Literatura científica

Selección definitiva de *primers*: 14 parejas (varias especies de *Prunus*)

Cantidad de DNA necesaria: 10 ng por análisis y muestra

Tamaño muestral analizado: 27 ejemplares individuales de cerezo

Máximo número de alelos por locus: 10

Probabilidad total de exclusión: 0.999

Probabilidad de identidad: 2.433×10^{-14} (con todos los marcadores)