

En esta memoria se presenta la Tesis Doctoral que lleva por título:

DESARROLLO DE UNA PLATAFORMA DE
TILLING EN MELÓN (*Cucumis melo* L.)

Escrita por Mireia González To

Director: Jordi Garcia Mas

Tutora: M^a Carme Espunya

Esta Tesis ha sido realizada en el Programa de Genòmica i Biotecnologia del IRTA que pertenece al Centre de Recerca en Agrigenòmica (CRAG) y dentro del programa de doctorado del Departamento de Bioquímica de la UAB.

A Enero de 2014



A la meva mare que m'ha ensenyat a estimar i a mantenir l'alegria

"Dijo Tennyson que si pudiéramos comprender una sola flor sabríamos quiénes somos y qué es el mundo"

Jorge L. Borges

Ara tot just intento pensar en el principi d'aquesta aventura i me n'adono que jo era una persona totalment diferent. Els agraïments moltes vegades es queden curts, les paraules tants cops no poden copsar tot el que ha significat un gest o la companyia d'una persona. Per això vull començar dient que, segurament, aquestes paraules mai diguin tot el que volen dir. Aquest només és un intent de reconèixer i sobretot, agrair-vos a tots els que m'heu acompanyat en aquesta aventura de 8 anys.

En primer lloc vull agrair a les persones que han fet possible aquest periple: a la Carme Espunya per ser la professora que em va descobrir el meu amor científic cap a les plantes. Al Jordi Garcia-Mas per confiar en una estudiant de pràctiques que arribava a l'IRTA de Cabrils amb moltes ganes d'aprendre i al Pere Arús per recolzar-lo en aquesta decisió. Vull agrair als que han col.laborat en aquest projecte perquè sense ells no s'hagués pogut dur a terme: a Fitó i al COMAV. En especial a la Belén Picó per aconseguir publicar un article d'aquest treball amb la seva dedicació i persistència.

L'època a l'IRTA ha estat un dels períodes més feliços de la meua vida. Tots els que hi éreu hi vau contribuir; cadascú a la seva manera i durada. Els becaris (Edu (i les guerres privades d'aigua al laboratori quan ja no hi quedava ningú), Ibo, Iria, Juan (¡Ánimo, ya no queda nada!), Julio, Alí, Claudio, Roger, Eudald...). A tots ells els hi agraeixo la companyia, el fet de poder compartir la mica de desesperació que tantes vegades pateix un becari i poder riure dels nostres errors. Si un no s'equivoca, no aprèn.

Des d'aquí, d'aquest petit espai meu i propi, vull fer un reconeixement a tots els tècnics, sense vosaltres els laboratoris no avançarien. És d'aquestes coses que tothom sap, però que crec que a vegades cal dir en veu alta i clara.

Els són: Fuensi (gràcies per ensenyar-me tant i per tenir tanta paciència amb el nostre estimat LL-COR), Àngel (gràcies per la teua disposició i la teua alegria incondicional) Joana, Ana, Patricia, Alicia, Dani, Manoli... En especial a la Vane, sempre al meu costat, inseparables companyes de "poyata". Gràcies per fer-me riure (i a vegades fer-me plorar), per fer-me ballar i treballar, per ensenyar-me les

tècniques i perquè fora de la feina m'has cuidat igual que cuides als teus. Gràcies per fer-me sentir tan estimada.

Als companys de caseta (Montse, Francesc, Olga, Ainoha, Jonatan...) perquè allà vam ser autodidactes, vam compartir cafès i coneixements. Potser la ciència no cal que sigui tan seria, a vegades. A la Puri, pel "cachondeo" que m'animava les tardes com a gairebé ningú, encara que li trepitgés els "fregaos".

A tots els investigadors (Werner, Jordi, Amparo, Momo, Montse M., Montse S., Wim, Txose, Gisela, Torben, Toni M.) i especial al Mourad per ensenyar-me que la ciència pot ser molt creativa, plena de colors inesperats (¡Esa eterna discusión con el salmón de los geranios!) i, sobretot, per acompanyar-me en tantes tardes i nits de petita solitud a Arenys en les que em cuinaves *cous-cous* i peix al forn com si fos la seva pròpia germana. Al Santi Vilanova, perquè sempre m'ha cuidat i perquè en els moments que queia ell sempre m'oferia la mà per ajudar-me a aixecar. A la Marta Pujol, no només per tot el que l'admiro com a professional i el que he après amb ella, sinó també com a persona. Marta, sense tu, segur, que no hagués acabat aquest projecte. Gràcies Marta.

Un dels aventatges que té la investigació és que et permet viatjar i fer estades a laboratoris d'altres països. A Evry vaig poder treballar al costat de l'Abdelhafid Bendahmane al qui vull agrair l'ensenyament de la tècnica del TILLING. També a la Christelle Troadec, amb qui vam compartir hores d'aprenentatge, llargues xerrades en *franco-català-anglès* i molts caramels. Merci beaucoup Christelle. La estancia en Evry fue como un "mini Erasmus" gracias a dos latinos: Nilo y Silvia. A ellos quiero agradecer que me acogieron en su piso como una más, me cocinaron a la italiana y me prepararon Pisco-Sour para olvidarnos entre todos de los diseños de cebadores y de las plantas de tomate. Nilo y Silvia siempre os recuerdo, gracias.

A pesar de haber acabado las prácticas de doctorado en Marzo de 2010 y de no haber presentado la tesis, conseguí una beca en Italia, más concretamente en Roma para trabajar con Giovanni Giuliano en el ENEA. Considero ese periodo como una ampliación del doctorado y de mi aprendizaje. Per questo non vorrei lasciare di ringraziare a Marco, Silvia, Killy, Sara, Patrizia, Elio (DT!) e Giulia l'incredibile esperienza di vivere e lavorare a Roma. Lavorare divertendomi così tanto: ballare e bere

la birra in laboratorio con Silvia al termine di lunghe giornate di lavoro tra i pomodori e PCR. Grazie a Giuliano per la fiducia in me e nelle mie capacità. E a tutti per rendere questo periodo uno dei giorni più felici della mia vita. Voglio tanto bene all'Italia e gli italiani!.

Finalment i no menys important, vull agrair als amics que m'acompanyeu des de sempre i que m'heu vist créixer i madurar en aquest projecte. Que heu aguantat les llargues i constants (i segurament pesades ;-)) converses sobre aquest tema i sempre heu confiat en mi. Crec que si confies en algú aquella persona també confia més amb ella mateixa per això vull dir-vos que és un regal preciós el que m'heu fet: Club Bikini (Marta, Anna P., Anna A.), PSFs!! (Martina, Jess, Raquel B. -gràcies també pel increïble disseny de la portada, sabia que faries alguna cosa que la sentiria pròpia, gràcies artista!- Raquel Ll., Nayibe, Isa, Nurieta, Alba, Tere...), Claudia sempre aprop estiguis on estiguis, Humanistes! (Marta, Eva, Laia) que m'heu acompanyat i juntes hem descobert tantes coses fantàstiques de la vida com per exemple que el coneixement és poder. Sou totes fantàstiques, dones increïbles. Us admiro a totes i de totes he après lliçons imprescindibles per créixer. Sou l'orgull del gènere femeni! ;-).

Al Jordi Carreras perquè sense ell probablement mai hagués iniciat aquesta aventura. Gràcies per ajudar-me tant a creure en mi. I a tants amics que sempre em feu riure, pensar, imaginar, crear, superar-me: Jorge, Aleix, Belén, La Peña!!!, Carme, Manolo, Chelo, Marisa, Alexis...Papa, gràcies per ensenyar-me la disciplina, la constància i el sacrifici que han estat realment indispensables per projectes com aquest.

Per acabar, vull agrair a la meva família tot l'esforç que han fet des de sempre i sobretot perquè sou els meus fans incondicionals i em cuideu com ningú. Mama no tinc paraules, ets fantàstica, la meva heroïna sense dubte. Marc, simplement t'adoro. Àvis (Eulàlia i Quimet) sou el meu exemple a seguir, no he conegut persones més bones que vosaltres. Patrick, tempiternament, gràcies per TOT, per la VIDA i per ensenyar-me que si un vol pot aconseguir els seus somnis. Nada por obligación, todo por ilusión. Gràcies, Gracias, Grazie, Thanks, Merci,

Mireia González To
Barcelona, 07/01/2014

ÍNDICE

LISTADO DE ABREVIACIONES.....	1
RESUMEN.....	3
1. INTRODUCCIÓN.....	5
1.1 EL MELÓN.....	5
1.1.1 TAXONOMÍA.....	5
1.1.2 MORFOLOGÍA.....	5
1.1.3 IMPORTANCIA ECONÓMICA DEL MELÓN.....	6
1.1.4 MEJORA GENÉTICA DEL MELÓN.....	8
1.1.5 HERRAMIENTAS GENÉTICAS Y GENÓMICAS DISPONIBLES.....	10
1.1.5.1 HERRAMIENTAS DISPONIBLES EN MELÓN.....	10
A) Mapas Genéticos.....	10
B) Transformación Genética.....	11
C) BACs.....	12
D) ESTs.....	12
E) Micro matrices de ADN.....	13
F) TILLING.....	13
1.1.5.2 EL GENOMA DEL MELÓN.....	14
1.1.5.3 LA ERA POST-GENÓMICA.....	14
1.2 MUTACIONES.....	16
1.2.1 TIPOS DE MUTACIONES.....	16
1.2.1.1 ALTERACIONES DE LAS BASES.....	16
A) ERRORES DE REPLICACIÓN.....	17
A.1 Debidos a las formas tautoméricas.....	17
A.2 Inserciones y deleciones.....	18
B) LESIONES ESPONTÁNEAS EN LAS BASES.....	18
B.1 Desaminación.....	18
B.2 Despurinización.....	19
B.3 Oxidación.....	19
1.3 MUTAGÉNESIS INDUCIDA.....	20
1.3.1 EL TÉRMINO MUTAGÉNESIS.....	20
1.3.2 LA MUTAGÉNESIS EN LA MEJORA GENÉTICA.....	20
1.3.3 TIPOS DE MUTÁGENOS.....	21

1.3.3.1 MUTAGÉNESIS QUÍMICA.....	22
A) Bases análogas.....	22
B) Agentes químicos modificadores de bases.....	23
C) Agentes intercalantes.....	24
1.3.3.2 MUTAGÉNESIS FÍSICA.....	24
A) Radiaciones electromagnéticas.....	24
B) Radiaciones ionizantes.....	24
1.3.3.3 OTROS MÉTODOS DE MUTAGÉNESIS.....	26
A) TRANSPOSONES.....	26
B) T-DNA (ADN de transferencia).....	26
C) VIRUS.....	26
D) Mutagénesis dirigida.....	27
1.4 TILLING.....	29
1.4.1 INTRODUCCIÓN.....	29
1.4.2 LA TÉCNICA DE TILLING.....	30
1.4.2.1 DESARROLLO DE LA POBLACIÓN MUTAGENIZADA.....	32
1.4.2.2 DESARROLLO DE LAS MEZCLAS DE ADN.....	35
1.4.2.3 DETECCIÓN DE MUTANTES.....	36
1.4.2.4 CARACTERIZACIÓN DE LOS MUTANTES.....	41
1.4.2.5 PLATAFORMAS DE TILLING EXISTENTES.....	43
1.5 GENES CANDIDATOS PARA TILLING USADOS EN EL MARCO DE ESTA TESIS DOCTORAL.....	46
1.5.1 <i>PDS</i> (Fitoeno Desaturasa).....	46
1.5.2 <i>eIF4E</i> (FACTOR DE INICIACIÓN DE LA TRADUCCIÓN 4E) y <i>eIF(iso)4E</i>	48
1.5.3 RECEPTOR DE ETILENO ETR1.....	55
OBJETIVOS.....	61
2. MATERIALES Y MÉTODOS.....	63
PRIMERA PARTE: PROTOCOLO TILLING.....	63
2.1.CREACIÓN Y DESARROLLO DE UNA POBLACIÓN MUTAGENIZADA.....	63
2.1.1 MUTACIÓN Y GERMINACIÓN DE LAS SEMILLAS (M0).....	63
2.1.2 CRECIMIENTO DE LAS GENERACIONES M1 Y M2.....	65
2.1.2.1 AUTOFECONDACIÓN DE LA GENERACIÓN M1 EN INVERNADERO.....	65
2.1.2.2 CRECIMIENTO DE LA GENERACIÓN M2.....	66
2.1.3 EXTRACCIÓN DE ADN.....	66
2.1.3.1 EXTRACCIÓN DE ADN MEDIANTE KIT COMERCIAL.....	66
2.1.3.2 MINI EXTRACCIÓN DE ADN.....	67

2.1.3.3 ELECTROFORESIS EN GEL DE AGAROSA	68
2.2 DETECCIÓN E IDENTIFICACIÓN DE LOS MUTANTES.....	69
2.2.1 ELECCIÓN Y SECUENCIACIÓN DEL GEN DIANA.....	70
2.2.1.1 AMPLIFICACIÓN DE FRAGMENTOS MEDIANTE (PCR).....	70
2.2.1.2 SECUENCIACIÓN DE ADN.....	71
2.2.1.3 ANÁLISIS DE LAS SECUENCIAS.....	72
2.2.2 DISEÑO DE CEBADORES PARA LA DETECCIÓN DE MUTANTES	72
2.2.3 AMPLIFICACIÓN DEL FRAGMENTO A ANALIZAR	72
2.2.3.1 <i>NESTED PCR</i>	73
2.2.4 FORMACIÓN DEL HETERODÚPLEX	73
2.2.5 DIGESTIÓN DEL HETERODÚPLEX.....	74
2.2.5.1 ELECCIÓN DE LA CONCENTRACIÓN DE LA ENZIMA <i>EndoI</i>	74
2.2.5.2 DIGESTIÓN DEL PRODUCTO DE LA PCR CON <i>EndoI</i>	74
2.2.5.3 PURIFICACIÓN DEL PRODUCTO DE LA DIGESTIÓN.....	74
2.2.6 SEPARACIÓN DE LAS MUESTRAS MEDIANTE GELES DE ACRILAMIDA VERTICALES Y DECONVOLUCIÓN	75
2.2.6.1 PREPARACIÓN DEL GEL DE ACRILAMIDA.....	75
2.2.6.2 CARGA DE LAS MUESTRAS EN EL GEL DE ACRILAMIDA.....	76
2.2.6.3 ANÁLISIS DEL GEL DE ACRILAMIDA: DECONVOLUCIÓN	77
2.2.7 SECUENCIACIÓN Y ANÁLISIS DE LA MUTACIÓN	79
2.2.8 CONFIRMACIÓN DE LA SEGREGACIÓN DE LA MUTACIÓN EN LA FAMILIA MUTANTE.....	79
2.3 ECOTILLING.....	80
2.3.1 PREPARACIÓN DEL ADN	80
2.3.2 DETERMINACIÓN DE LOS HAPLOTIPOS	80
SEGUNDA PARTE: PROTOCOLOS GENERALES	81
2.4 EXTRACCIÓN DE ARN	81
2.5 GENOTIPADO CON MARCADORES MOLECULARES.....	81
2.5.1 CAPS (<i>Cleaved Amplified Polymorphic Sequence</i>)	81
2.5.2 MICROSATÉLITES O SSRs (<i>Simple Sequence Repeats</i>)	82
2.5.3 DETECCIÓN DE SNPs MEDIANTE LA TÉCNICA DE SNAPSHOT	83
2.6 CLONAJE EN PLÁSMIDO	83
2.6.1 TRANSFORMACIÓN DE CÉLULAS COMPETENTES MEDIANTE CHOQUE TÉRMICO	84
2.6.2 EXTRACCIÓN DEL ADN PLASMÍDICO (MINIPREP).....	84
3. RESULTADOS	87
3.1 CONSTRUCCIÓN DE UNA PLATAFORMA DE MUTANTES TILLING EN MELÓN	87
3.1.1 LA POBLACIÓN DE TILLING DE MELÓN	87

3.1.1.1 COMPROBACIÓN DE LA HOMOCIGOSIS DE LA LÍNEA M62-113	87
3.1.1.2 DETERMINACIÓN DE LA DOSIS ÓPTIMA DE MUTÁGENO	90
3.1.1.3 PREPARACIÓN DE LAS MEZCLAS DE ADN DE LA POBLACIÓN DE TILLING	92
3.1.1.4 PUESTA A PUNTO DEL MÉTODO TILLING	93
3.1.1.4.1 POBLACIÓN DE TILLING DEL INRA.....	93
3.1.1.4.2 TAQ POLIMERASA EMPLEADA PARA TILLING	94
3.1.1.4.3 OPTIMIZACIÓN DE LA TÉCNICA TILLING.....	94
3.1.2 CONCLUSIONES	97
3.2 BÚSQUEDA DE MUTACIONES EN GENES DE IMPORTANCIA AGRONÓMICA Y ESTUDIO DE SU EFECTO FENOTÍPICO	99
3.2.1 <i>PDS</i> (FITOENO DESATURASA).....	100
3.2.1.1 DISEÑO DE CEBADORES PARA EL GEN <i>PDS</i>	101
3.2.1.2 DETECCIÓN, IDENTIFICACIÓN Y CARACTERIZACIÓN DE MUTANTES EN <i>PDS</i>	105
3.2.1.3 ESTUDIO GENOTÍPICO Y FENOTÍPICO EN LAS FAMILIAS C384 Y F1115	108
3.2.2 <i>eIF4E</i> (FACTOR DE INICIACIÓN DE LA TRADUCCIÓN).....	113
3.2.2.1 DETECCIÓN, IDENTIFICACIÓN Y CARACTERIZACIÓN DE LOS MUTANTES EN EL GEN <i>eIF4E</i>	114
3.2.2.2 PREDICCIONES DE LA ESTRUCTURA EN 3D DE LOS MUTANTES <i>eIF4E</i>	117
3.2.3 <i>eIF(ISO)4E</i> (ISOFORMA DEL FACTOR DE INICIACIÓN DE LA TRADUCCIÓN 4E).....	121
3.2.3.1 DETECCIÓN, IDENTIFICACIÓN Y CARACTERIZACIÓN DE LOS MUTANTES.....	123
3.2.4 <i>ETR1</i> (RECEPTOR DE ETILENO 1).....	123
3.2.4.1 DETECCIÓN, IDENTIFICACIÓN Y CARACTERIZACIÓN DE LOS MUTANTES.....	124
3.2.5 EFICIENCIA DE LA MUTAGÉNESIS EN LA POBLACIÓN DE TILLING	125
3.3 EcoTILLING	128
3.3.1 ANÁLISIS MEDIANTE EcoTILLING EN EL GEN <i>eIF4E</i>	128
3.3.2 ANÁLISIS MEDIANTE EcoTILLING EN EL GEN <i>ETR1</i>	136
3.4 CATÁLOGO DE MUTANTES FENOTÍPICOS.....	140
3.4.1 MUTANTES EN ESTADIO DE PLÁNTULA	141
3.4.2 MUTANTES EN ESTADIO DE PLANTA ADULTA.....	142
3.4.3 MUTANTES EN FRUTO	143
3.4.4 CATÁLOGO DE IMÁGENES DE LOS MUTANTES	145
4. DISCUSIÓN.....	155
4.1 POBLACIÓN MUTAGENIZADA DE MELÓN Y PUESTA A PUNTO DEL MÉTODO TILLING	155
4.2 EFECTOS DE LAS MUTACIONES IDENTIFICADAS EN LOS GENES ANALIZADOS	161
4.2.1 <i>PDS</i> (FITOENO DESATURASA).....	161
4.2.2 <i>eIF4E</i> (FACTOR DE INICIACIÓN DE LA TRADUCCIÓN 4E).....	162

4.2.3 <i>eIF(iso)4E</i>	170
4.2.4 <i>ETR1</i> (RECEPTOR DE ETILENO 1)	171
4.2.5 FUTUROS EXPERIMENTOS	172
4.3 EcoTILLING	174
4.3.1 <i>eIF4E</i> (FACTOR DE INICIACIÓN DE LA TRADUCCIÓN)	174
4.3.2 <i>ETR1</i> (RECEPTOR DE ETILENO 1)	178
4.4 ESTUDIO FENOTÍPICO DE LA POBLACIÓN DE TILLING	183
CONCLUSIONES	189
Bibliografía	192
ANEXO I	211
ANEXO II	249
ANEXO III	257

ÍNDICE DE TABLAS Y FIGURAS

Figura 1.1: Gráfico con los principales productores de melón a nivel mundial.....	7
Figura 1.2: Gráfico de los países con más exportación de melón a nivel mundial	8
Figura 1.3: Efecto del EMS a nivel molecular.....	23
Figura 1.4: Protocolo empleado para el método de TILLING.....	32
Figura 1.5: Detección de una mutación en una población de TILLING en melón en un gel de acrilamida vertical de LI-COR 4300.....	38
Figura 1.6: Ejemplo de curva de disociación de fragmentos amplificados y renaturalizados. .	40
Figura 1.7: Ruta de síntesis de carotenos en plantas.....	47
Figura 1.8: Representación del sistema de resistencia en plantas basado en la interacción gen- a-gen (Flor, 1971).	49
Figura 1.9: Funcionamiento del inicio de la traducción	52
Figura 1.10: Funciones postuladas para el <i>eIF(iso)4E</i>	55
Figura 1.11: Esquema del ciclo de Yang	56
Figura 1.12: Esquema del <i>ETR1</i>	57
Figura 2.1: Procedimiento para obtener el ADN de la población de TILLING.	66
Figura 2.2: Diseño de las placas pool.	67
Figura 2.3: Esquema del protocolo de TILLING	73
Figura 2.4: Montaje gel LI-COR4300	76
Figura 2.5: Visualización de las muestras en el gel de TILLING.	77
Figura 2.6: Plantilla para la deconvolución.....	78
Figura 3.1: Posición de los SSRs analizados en la línea M62-113.....	90
Figura 3.2: Porcentajes de germinación.	91
Figura 3.3: Esquema del fragmento secuenciado del gen <i>eIF4E</i>	95
Figura 3.4: Esquema de la región de <i>PDS</i> de melón.....	102
Figura 3.5: Secuencia proteica del gen de la <i>PDS</i> en melón.....	103
Figura 3.6: Árbol filogenético del gen de la <i>PDS</i> de melón.....	104

Figura 3.7: Esquema del fragmento de la <i>PDS</i> analizado mediante TILLING.....	105
Figura 3.8: Cromatogramas de las dos familias con mutaciones en la zona exónica de la <i>PDS</i>	106
Figura 3.9: Representación de la parte secuenciada de la <i>PDS</i>	107
Figura 3.10: Representación de los 20 aminoácidos.....	108
Figura 3.11: Imagen de tres frutos de plantas M2	109
Figura 3.12: Plántulas M2 de la familia mutante C384.....	111
Figura 3.13: Cromatogramas de la <i>PDS</i>	112
Figura 3.14: Representación del gen <i>eIF4E</i> en melón (<i>nsv</i>).....	113
Figura 3.15: Árbol filogenético del gen <i>eIF4E</i>	114
Figura 3.16: Alineamiento entre las proteínas de <i>eIF4E</i> y el <i>eIF(iso)4E</i> de meló	116
Figura 3.17: Modelaje de la proteína <i>eIF4E</i> del mutante 2.....	118
Figura 3.18: Modelaje de la proteína <i>eIF4E</i> en 3D con posible ligando (molécula 5'CAP del ARNm).....	118
Figura 3.19: Modelaje de la proteína <i>eIF4E</i> en 3D mediante el <i>Output</i> del programa PyMOL del mutante 5.....	119
Figura 3.20: <i>Output</i> del programa PyMOL del mutante 5	119
Figura 3.21: Modelaje de la proteína <i>eIF4E</i> en 3D mediante el <i>Output</i> del programa PyMOL del mutante 8	120
Figura 3.22: Modelaje de la proteína <i>eIF4E</i> en 3D mediante el <i>Output</i> del programa PyMOL del mutante 9.....	120
Figura 3.23: Modelaje de la proteína en 3D del gen <i>eIF4E</i>	121
Figura 3.24: Esquema del gen <i>eIF(iso)4E</i>	122
Figura 3.25: Árbol filogenético del gen <i>eIF(iso)4E</i>	123
Figura 3.26: Esquema del gen <i>ETR1</i>	123
Figura 3.27: Árbol filogenético del gen <i>ETR1</i>	124
Figura 3.28: Gel de EcoTILLING para el primer amplicón de <i>eIF4E</i>	129
Figura 3.29: Cromatograma de las variedades CUM474 y M62-113.....	130

Figura 3.30: Gel de EcoTILLING para el segundo amplicón de <i>eIF4E</i>	133
Figura 3.31: Cromatogramas de M62-113 y PI 505602.....	136
Figura 3.32: Gel de EcoTILLING para el amplicón del gen <i>ETR1</i>	137
Figura 3.33: Cromatograma del SNP 1 del primer exón del gen <i>ETR1</i>	138
Figura 3.34: Procedimiento para la selección de melones con fenotipos	144
Figura.4.1: Plántula con fenotipo albino.....	162
Figura 4.2: Alineamiento mediante <i>ClustalW</i> de la proteína eIF4E	164
Figura 4.3: <i>Output</i> del programa PMUT del gen eIF4E.....	167
Figura 4.4: Alineamiento mediante <i>ClustalW</i> de <i>eIF4E</i> y <i>eIF(iso)4E</i>	171
Figura 4.5: Traducción a proteína del ARNm del gen eIF4E.....	177
Figura 4.6: Estructura de los dominios identificados en <i>ETR1</i> de <i>Arabidopsis</i>	178
Figura 4.7: Alineamiento múltiple de los genes <i>ETR1</i> en <i>Arabidopsis</i> , tomate y melón y los genes <i>ETR4</i> , <i>ETR5</i> y <i>NR</i> de tomate.	181
Figura 4.8: Acciones en las que se han detectado el SNP 1	182
Figura 4.9: Información disponible en la base de datos UTILLdb.....	184
Figura 4.10: Fenotipos de fruto interesantes a nivel comercial	185
Figura 4.11: Fenotipos identificados en planta en otras poblaciones de TILLING.	187
Figura 4.12: Fenotipos identificados en fruto en otras poblaciones de TILLING	188
Tabla 1.1: Representación de las formas tautoméricas	17
Tabla 1.2: Resumen de los mutágenos más habituales.....	22
Tabla 1.3: Poblaciones de TILLING desarrolladas en animales y plantas.....	34
Tabla 2.1: Dosis de EMS utilizadas para mutar la población	64
Tabla 3.1: Localización de SSRs.....	88
Tabla 3.2: Dosis de EMS usadas para mutar las semillas.....	92
Tabla 3.3: Tabla comparativa.....	94
Tabla 3.4: Datos utilizados para obtener las secuencias de los genes candidatos	100
Tabla 3.5: Mutantes encontrados en el fragmento de 1188 pb del gen <i>PDS</i>	107

Tabla 3.6: Germinación de la semillas M2 en las familias C384 y F1115.	109
Tabla 3.7: Resultados de la segregación de la mutación en las familias mutantes de la <i>PDS</i> ..	112
Tabla 3.8: Mutantes del primer del gen <i>eIF4E</i>	115
Tabla 3.9: Mutantes en el segundo amplicón del gen <i>eIF4E</i>	115
Tabla 3.10: Características del mutante encontrado en el gen <i>eIF(iso)4E</i>	123
Tabla 3.11: Eficiencia de mutación de poblaciones de TILLING melón.	125
Tabla 3.12: Tasas de mutación de TILLING del IRTA y el INRA.....	127
Tabla 3.13: SNPs detectados en el primer amplicón del gen <i>eIF4E</i>	131
Tabla 3.14: SNPs detectados en el segundo amplicón de <i>eIF4E</i>	134
Tabla 3.15: Haplotipos encontrados mediante el análisis por EcoTILLING.....	135
Tabla 3.16: Polimorfismos detectados en el primer exón del gen <i>ETR1</i>	138
Tabla 3.17: Tabla con los haplotipos detectados para el gen <i>ETR1</i>	139
Tabla 3.18: Porcentajes de germinación en las 400 familias M2 observadas.....	141
Tabla 3.19: Fenotipos encontrados en estadio de plántula	142
Tabla 3.20: Fenotipos encontrados en estadio de planta adulta	143
Tabla 3.21: Fenotipos estudiados en frutos	145
Tabla 4.1: Porcentaje de bases susceptibles al mutágeno EMS	157
Tabla 4.2: Tabla resumen de los mutantes encontrados en cada población.....	159
Tabla 4.3: Resultados del estudio de segregación del gen <i>eIF4E</i>	169
Tabla 4.4: Tabla comparativa de los polimorfismos EcoTILLING.	175
Tabla 4.5: Fenotipos de las accesiones que presentan el SNP1	182

LISTADO DE ABREVIACIONES

ADN: Ácido Desoxirribonucleico
AFLP: Amplified Fragment Length Polymorphism
BAC: Bacterial Artificial Chromosome
BLAST: Basic Local Alignment Search Tool
CODDLE: Codons Optimized to Discover Deleterious Lesions
COMAV: Centro de Conservación y Mejora de la Agrodiversidad Valenciana
CRAG: Centre de Recerca en Agrigenómica, Barcelona
eIF4E: Factor de iniciación de la traducción 4E
EMS: Etil Metano Sulfonato
EST: Expressed Sequence Tag
ETR: Ethylene Receptor
FAO: Food and Agriculture Organization
LDH: Líneas Doble Haploides
MAPA: Ministerio de Agricultura, Pesca y Alimentación.
MNSV: Melon Necrotic Spot Virus
M0: Semillas mutageneizadas para germinar y producir la generación M1
M1: Primera generación de plantas crecidas a partir de las semillas M0
M2: Segunda generación de plantas crecidas a partir de las semillas de las plantas M1
NCBI: National Center for Biotechnology Information
NILs: Líneas casi Isogénicas (Near Isogenic Lines)
pb: pares de bases
PCR: Reacción en cadena de la polimerasa (Polymerase Chain Reaction)
PDS: Fitoeno Desaturasa
PI: PI 161375 Shongwan Charmi
PS: Piel de Sapo T111
QTL: Quantitative Trait Locus
RAPD: Random Amplified Polymorphic DNA
RFLP: Restriction Fragment Length Polymorphism
SIFT: Sorting Intolerant From Tolerant
SNP: Single Nucleotide Polymorphism
SSR: Simple Sequence Repeat
TILLING: Targeting Induced Local Lesions IN Genomes
URGV: Unité de Recherche en Génomique Végétale, Evry (Francia)

RESUMEN

El trabajo descrito en esta tesis doctoral se enmarca dentro del proyecto MELOGEN (GEN2013-20237, 2004-2007) uno de cuyos objetivos era la construcción de una plataforma de TILLING en melón.

La disponibilidad de recursos genéticos y genómicos para el melón se ha incrementado a lo largo de los últimos años e incluso la secuencia del genoma está disponible. Sin embargo, las herramientas genómicas para determinar la funcionalidad de los genes en esta especie son todavía limitadas. El TILLING (Targeting Induced Local Lesions IN Genomes) es un método de genética inversa mediante el cual se pueden crear e identificar nuevas mutaciones en genes candidatos y observar su efecto fenotípico. Además, también se pueden detectar fenotipos de interés para posteriormente identificar la mutación responsable de los mismos.

En este trabajo se ha desarrollado una población de TILLING en *Cucumis melo* a partir de la línea doble haploide del tipo "Piel de Sapo" M62-113. Aplicando inicialmente el mutágeno químico EMS (Etil Metano Sulfonato) sobre 17.000 semillas se ha desarrollado una colección de mutantes compuesta por 3.268 familias M2. A partir de 400 familias M2 se observaron diversos fenotipos mutantes en las diferentes etapas de crecimiento de la planta. Algunos de estos mutantes se han descrito con detalle generando un catálogo de los fenotipos observados en cada familia, algunos de los cuales pueden tener interés comercial.

Para medir la tasa de mutación de la población se han analizado cuatro genes, la fitoeno desaturasa *PDS*, los factores de iniciación de la traducción *eIF4E* y *eIF(iso)4E* y el receptor de etileno *ETR1*. Los mutantes detectados han permitido calcular la tasa de mutación de la población, una mutación cada 1,5 Mb. En el caso de *PDS* se ha podido obtener un mutante con fenotipo albino, demostrando así la efectividad del método TILLING en melón. Además, paralelamente se analizaron los mismos genes en otra población mutagenizada independiente.

También se realizó un estudio de EcoTILLING con 113 accesiones de melón y se identificaron 19 haplotipos distintos en el gen *eIF4E*, uno de ellos, con un SNP que provoca un cambio de aminoácido y que sólo se encuentra en la accesión PI 505602 y dos haplotipos en el gen *ETR1*, uno de los cuales también contiene un SNP que provoca un cambio de aminoácido.

El TILLING facilitará los estudios de genética inversa en melón y abre la puerta a futuros estudios de genómica funcional en un genoma recientemente secuenciado. Este trabajo ha representado también el primer paso para poder desarrollar nuevas poblaciones de TILLING en melón con mayores tasas de mutación y en otros fondos genéticos.

1. INTRODUCCIÓN

1.1 EL MELÓN

1.1.1 TAXONOMÍA

La especie *Cucumis melo* L. pertenece a la familia de las Cucurbitáceas, a la tribu Melothriaceae, subtribu Cucumerinae, género *Cucumis*. A parte de este género que incluye también el pepino (*Cucumis sativus*) (Jeffrey, 1990), encontramos otros dos géneros que agrupan especies de importancia económica a nivel mundial como *Citrullus lanatus* (sandía) y *Cucurbita pepo* (calabaza y calabacín) (Jeffrey, 1990). El género *Cucumis* se divide en dos subgéneros que separan básicamente las especies africanas de las asiáticas. El primer subgénero, denominado *Melo*, incluye, entre otros, al melón. El segundo subgénero, *Cucumis*, incluye, entre otros, al pepino (Jeffrey, 1990). Existen distintas clasificaciones del melón, sin embargo, una de las más completas es la propuesta por Pitrat y colaboradores (2000) que divide la especie en 16 grupos, 5 de los cuales se agrupan en la subespecie *agrestis* y 11 en la subespecie *melo*.

El melón es una especie con una gran variabilidad genética y extensamente cultivada en todo el mundo. Un reciente estudio de secuencias de ADN nuclear y plasmídicas en 100 accesiones de *Cucumis* de distintas zonas geográficas (África, Australia y Asia) ha sugerido que el melón tiene un origen asiático y que probablemente su antecesor silvestre se originó en la India (Sebastian, 2010).

1.1.2 MORFOLOGÍA

La planta del melón es anual y herbácea y puede ser de porte rastrero o trepador dependiendo de las condiciones de cultivo a las cuales se someta. Su sistema radicular es abundante, muy ramificado y de rápido desarrollo. Los tallos principales están recubiertos de formaciones pilosas y presentan nudos en los que se desarrollan hojas, zarcillos y flores. Los nuevos tallos brotan de las axilas de las hojas. La hoja tiene forma de limbo ovado, habitualmente pentagonal y con los márgenes dentados. Las hojas también son vellosas por el envés. Las flores son de color amarillo y pueden ser masculinas, femeninas o hermafroditas. Las primeras en

desarrollarse son las flores masculinas en los entrenudos primarios mientras que las femeninas y hermafroditas aparecen más tarde en las ramificaciones posteriores. La polinización es entomófila) y mayoritariamente realizada por abejas (Handel, 1982; Monforte y col., 2004).

La forma del fruto es muy variable (esférica, elíptica u ovada). Se han podido identificar genes que parecen estar relacionados con la forma de fruto por su implicación en la determinación de la forma de la flor o del ovario, entre ellos, el gen *a* (monoico) y el gen *p* (pentámero) (Perin y col, 2002a).

Para la corteza del fruto existe también una gran variabilidad tanto de colores (verde, amarillo, anaranjado o blanco) como de texturas. Encontramos pieles lisas, reticuladas o estriadas. La pulpa puede ser blanca, amarilla, anaranjada o verdosa. El fruto del melón puede madurar de manera climatérica o no climatérica según la variedad. La maduración climatérica se caracteriza por un aumento de la tasa de respiración y es iniciada por un pico en la concentración de etileno. Sin embargo, la maduración no climatérica se caracteriza por una disminución continua en la tasa de respiración y en la producción de etileno (Moore y col, 2002). La identificación de algunos *Quantitative Trait Locus* (QTLs) responsables de este carácter junto con la localización de genes candidatos de firmeza del fruto en líneas de introgresión (NILs) obtenidas de los parentales PI 161375 y T111, entre otros estudios, han permitido comenzar a comprender cómo funciona este mecanismo (Moreno y col., 2008). Estas características han hecho que el melón sea considerado como una de las especies modelo para poder comprender el proceso de maduración del fruto así como su regulación a nivel genético y molecular.

1.1.3 IMPORTANCIA ECONÓMICA DEL MELÓN

El melón se cultiva en zonas tropicales y subtropicales secas. En Europa, se cultiva estacionalmente en verano y primavera. Según los últimos datos de la FAO, en 2011 China era el primer productor mundial con 13.085.500 toneladas, seguido por Irán con 1.801.830 y por Turquía con 1.647.990, mientras que España queda en séptimo lugar con una producción de 869.475 toneladas (Figura 1.1).

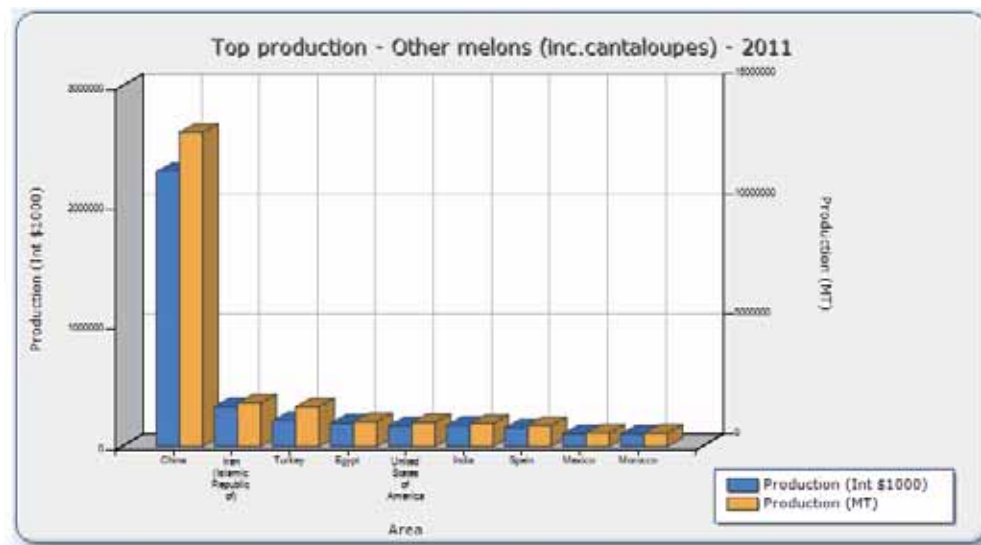


Figura 1.1: Gráfico con los principales productores de melón a nivel mundial. España se encuentra en el séptimo lugar (Int 1000\$: precio medio de todas las monedas internacionales tomando el dólar americano como referencia; MT: Tonelada metrica). Fuente FAO, <http://faostat.fao.org/site/342/default.aspx> (2013).

Si analizamos las exportaciones, España es el primer exportador mundial con 376.103 toneladas, seguido de Guatemala y Honduras (Figura 1.2). Este dato corrobora el hecho de que en España el melón es un producto no sólo muy demandado por parte del consumidor, si no que además, tiene mucha demanda a nivel internacional, y por lo tanto, un importante papel en la economía del país. Cabe decir que en España el melón es la segunda especie hortícola más cultivada después del tomate. En el 2011 se cultivaron 32.000 Ha y se produjeron 869.475 toneladas de melón. La mayor parte de la superficie cultivada de melón se encuentra en las regiones de Murcia, Castilla la Mancha y Andalucía (<http://faostat.fao.org>).

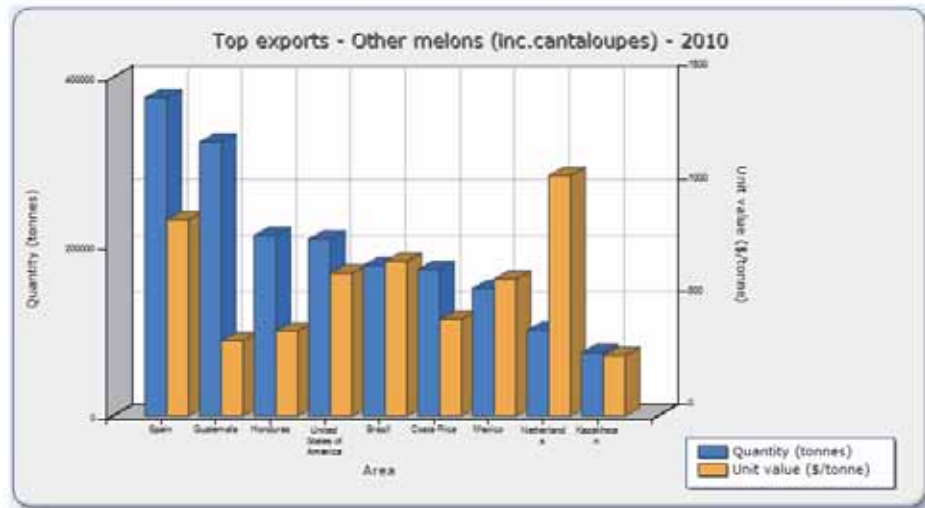


Figura 1.2: Gráfico de los países con más exportación de melón a nivel mundial. España se encuentra en primer lugar. Fuente FAO, <http://faostat.fao.org/site/342/default.aspx> (2013).

1.1.4 MEJORA GENÉTICA DEL MELÓN

La importancia económica del melón en nuestro país justifica que existan empresas productoras de semillas interesadas en mejorar esta especie y adaptarla a las necesidades de los productores y los consumidores. Algunos de los objetivos de esta mejora genética son obtener producciones con un alto rendimiento, incorporar resistencias a plagas o enfermedades y la mejora del fruto tanto en caracteres morfológicos (por ejemplo la forma del fruto) como organolépticos (por ejemplo el aroma o la dulzura). Las resistencias buscadas son principalmente frente a virus como por ejemplo, el virus del mosaico del pepino, *Cucumber Mosaic Virus* (CMV) o el virus de las manchas necróticas el melón, *Melon Necrotic Spot Virus* (MNSV) y frente a hongos como el oídio (*Podosphaera xanthii*, *Golovinomyces cichoracearum*) o a *Fusarium* (*Fusarium oxysporum* f. sp. *melonis*). Asimismo, se procura optimizar la producción con variedades que ofrezcan un alto rendimiento en condiciones adversas de sequía y elevada salinidad o que permitan una adecuada calidad en postcosecha.

Antes del descubrimiento de los marcadores moleculares este tipo de mejora dependía directamente de la evaluación mediante la observación directa del fenotipo. Sin embargo, la expresión fenotípica es producto de la interacción del

genotipo con el ambiente y no todos los factores ambientales pueden ser controlados, hecho que dificulta el proceso de selección (Allard, 1999).

Gracias a las nuevas tecnologías en biología molecular se han desarrollado herramientas como los marcadores moleculares que permiten la selección asistida. En líneas generales, los marcadores genéticos permiten rastrear con rapidez el ADN de la planta que se está seleccionando permitiéndonos saber si el alelo de un gen se encuentra o no en la planta.

Los marcadores genéticos son polimorfismos heredables de forma mendeliana sencilla (Strachan y Read, 1999). Su variabilidad puede ser detectada y utilizada en el estudio de diversos caracteres genéticos. En la mejora clásica se utilizaban los marcadores morfológicos que normalmente se producían a partir de una mutación de un gen dando como resultado el fenotipo diferenciado. Posteriormente se usaron los marcadores basados en isoenzimas o variantes moleculares de una misma proteína (Stebbins y Moody, 1989; McMillin y Scandalios, 1983). Dado que cada isoenzima tiene una distinta composición en aminoácidos también resulta distinta su carga eléctrica. Esta característica se usaba para separarlas mediante electroforesis visualizando el patrón diferencial de cada individuo. Posteriormente, se desarrollaron los marcadores moleculares basados en los polimorfismos de ADN. Los *Restriction Fragment Length Polymorphism* (RFLPs) son marcadores basados en las diferencias de longitud de fragmentos de ADN obtenidos a partir de la digestión del genoma. Los fragmentos se digieren con distintas enzimas de restricción que cortan el ADN por lugares específicos de la secuencia (dianas de restricción). La presencia de polimorfismos puede afectar a las dianas de restricción impidiendo que la enzima digiera correctamente y por lo tanto que cada genotipo muestre un patrón diferencial (Baudracco-Arnas, 1996; Oliver, 2001).

Asimismo, los marcadores basados en la técnica de la Reacción en Cadena de la Polimerasa (PCR) han servido, entre otras utilidades, para seguir construyendo y saturando los mapas genéticos. Hay muchos tipos de marcadores derivados de la técnica de la PCR, como los *Amplified Fragment Length Polymorphism* (AFLPs), los *Random Amplified Polymorphic DNA* (RAPDs), los microsatélites o *Simple Sequence Repeats* (SSRs) o los SNPs (*Single Nucleotide Polymorphism*). Los AFLPs se basan en la digestión con enzimas de restricción del ADN genómico y posteriormente su amplificación mediante PCR (Vos y col., 1995). Los RAPDs se generan por la amplificación aleatoria de segmentos de ADN genómico a partir de un solo cebador de secuencia corta y arbitraria (Williams, 1990). Los SSRs son fragmentos de secuencias donde hay una repetición en tándem de 2-6 nucleótidos, aunque esta

longitud puede variar en función de la variedad. Gracias a esta variación se pueden desarrollar marcadores basados en la detección del número de repeticiones presentes en los genotipos analizados (Hamada y col., 1982; Gonzalo y col., 2005).

A pesar de existir una gran variedad de marcadores moleculares, en los últimos años se ha hecho un uso más exhaustivo de los marcadores SNPs. Este tipo de polimorfismo es muy abundante en el ADN (Brookes, 1999) y existen varias tecnologías rápidas y eficientes que permiten su descubrimiento y su genotipado (Smith y col., 2012). Algunos ejemplos de técnicas para su descubrimiento son la hibridación mediante micromatrices (Borevitz y col., 2003), el análisis *in silico* a partir de secuencias de ESTs (*Expressed Sequence Tag*) (Iehisa y col., 2012) o la resecuenciación de amplicones (Choi y col., 2007). Actualmente, en especies con el genoma secuenciado se pueden obtener cientos de miles de SNPs a partir de resecuenciación de variedades. Para su genotipado se han desarrollado técnicas como el uso de sondas TaqMan, entre muchas otras. Recientemente se ha saturado el mapa genético de melón tras el desarrollo de marcadores moleculares tipo SNPs (García-Mas y col., 2012).

1.1.5 HERRAMIENTAS GENÉTICAS Y GENÓMICAS DISPONIBLES

1.1.5.1 HERRAMIENTAS DISPONIBLES EN MELÓN

En el 2005 se realizó la primera reunión internacional de genómica de Cucurbitáceas en Barcelona, donde se estableció el *International Cucurbit Genomics Initiative* (ICuGI), que tomó decisiones de cierta relevancia para la investigación en esta especie. Una de ellas fue obtener una exhaustiva colección de secuencias de ESTs. Otra decisión relevante fue unificar los mapas genéticos que hasta entonces se habían desarrollado de un modo independiente en distintos laboratorios. De este modo se obtendría un mapa genético consenso mucho más saturado (Díaz y col., 2011). Finalmente, se decidió elaborar una única página web donde todos los datos genéticos y genómicos fueran de acceso público (<http://www.icugi.org>).

A) MAPAS GENÉTICOS

Un mapa genético o mapa de ligamiento es un mapa de marcadores polimórficos donde se definen las distancias entre los marcadores en función de la frecuencia de recombinación que ha habido entre ellos. Reproducen la estructura de los cromosomas mostrando las posiciones de los genes (en caso de que el marcador esté

anclado en un gen) o de marcadores anónimos basados en fragmentos de secuencias no codificantes.

Un mapa genético se obtiene generalmente a partir de poblaciones segregantes obtenidas de cruzamientos entre dos líneas parentales. Cuanto más alejados filogenéticamente sean los parentales, más informativo podrá ser el cruzamiento, ya que el grado de polimorfismo será más elevado. Existen distintos mapas genéticos de melón (Baudracco-Arnas, 1996; Périn y col., 2002). Pero cada mapa genético ha sido elaborado con distintas líneas parentales o distintos tipos de poblaciones. Los tipos de marcadores que se han posicionado en cada mapa genético también son distintos. Oliver y col. (2001) elaboraron un mapa genético a partir de una población F2 obtenida del cruzamiento entre dos líneas parentales, una el tipo *Inodorus* "Piel de sapo" T111 y la accesión coreana Songwhan Charmi, PI 161375. Este mapa se generó inicialmente con 411 marcadores RFLPs. Posteriormente Gonzalo y col. (2005) generaron una población de líneas doble haploides (LDHs), a partir del mismo cruzamiento anterior. La ventaja de las LDHs es que constituyen una población inmortal ya que son totalmente homocigóticas y por lo tanto, en las generaciones posteriores obtenidas por autofecundación, se obtienen plantas idénticas a las parentales (Gonzalo y col., 2005). A partir de esta población se generó un nuevo mapa genético basado en el anterior, y finalmente se fusionó la información de ambos mapas. El resultante contenía aproximadamente 479 marcadores con una densidad de 2.5 cM/marcador (Gonzalo y col., 2005), al cual se añadieron nuevos marcadores tipo SNPs y SSRs en posteriores trabajos (Deleu y col. 2009, Fernández-Silva y col., 2008). Durante el 2011 y dentro del ICuGI se publicó un nuevo mapa consenso basado en la fusión de la información que habían aportado ocho mapas genéticos publicados anteriormente. Este nuevo mapa consenso contiene 1592 marcadores (640 SSRs, 330 SNPs, 252 AFLPs, 239 RFLPs, 89 RAPDs y 11 caracteres morfológicos, entre otros) con una densidad media de 0,72 cm/marcador (Diaz y col., 2011). En este mapa consenso también se han posicionado todos los QTLs descritos en melón hasta la fecha..

B) TRANSFORMACIÓN GENÉTICA

Otra herramienta frecuentemente usada en biotecnología es la transformación genética para la obtención de plantas transgénicas. En el caso del melón, aunque se ha conseguido la transformación con *Agrobacterium tumefaciens*, es un método ineficiente. La transformación en melón se utilizó por primera vez con éxito por Fang y Grumet en 1990 pero la tasa de eficiencia fue muy baja. Uno de los

inconvenientes en la transformación de las plantas de melón fue el descubrimiento de que ciertos explantes producían niveles más altos de etileno durante la inoculación con *Agrobacterium* (Ezura y col., 2000). Esta producción de etileno inhibía la expresión del gen *vir* y consecuentemente la transformación (Nonaka y col. 2008). Actualmente todavía se están identificado los factores por los cuales la eficiencia de transformación en melón es baja (Chovelon y col., 2011).

C) BACs

Para poder realizar experimentos como el clonaje posicional de genes y probar la funcionalidad de ciertos genes, ha sido indispensable disponer de colecciones de BACs (Cromosomas Artificiales Bacterianos). Los BACs son también de gran utilidad para poder desarrollar mapas físicos, o hacer estudios de microsintenia con otras especies, entre otras aplicaciones (Budiman y col., 2000; Deleu y col., 2007; González y col., 2010a). Las primeras genotecas de BACs en melón fueron construidas a partir de la línea MR-1 (Luo y col., 2001). Posteriormente se han generado otras genotecas que se han usado para detectar genes de interés que confieren resistencias a patógenos (Nieto y col., 2006) o genes que determinan el sexo en las flores de las plantas de melón (Boualem y col., 2008). Además, los BACs permiten localizar marcadores moleculares de un mapa genético y de este modo realizar mapas físicos de la especie, proporcionando anclajes para la localización de genes candidatos y para el ensamblaje de genomas (González y col., 2010b). Mediante clonaje posicional se han aislado los genes de resistencia *Vat*, *nsv* y *Fom-2* frente a algunas enfermedades que sufre el cultivo del melón (transmisión de virus por áfidos, MNSV (*Melon Necrotic Spot Virus*) y *Fusarium oxysporum f. sp. melonis* raza 0,1, respectivamente) (Pauquet y col., 2004; Nieto y col., 2006; Joobeur y col., 2004).

D) ESTs

El conocimiento de secuencias de ESTs es clave para la selección de genes candidatos para caracteres de interés agronómico. Gonzalez-Ibeas y colaboradores en 2007 (Anexo Ia) publicaron 30.000 nuevas secuencias de ESTs obtenidas a partir de ocho genotecas de cDNA elaboradas a partir de ARNm extraído de muestras de fruto en distintos estadios de desarrollo y de tejidos infectados con patógenos. Con la obtención de secuencias de ESTs se pudieron detectar nuevos polimorfismos y a partir de ellos desarrollar nuevos marcadores moleculares aumentando la saturación del mapa genético (Deleu y col., 2009) (Anexo Ib). En 2011 otro estudio a partir de 15 genotecas de cDNA de frutos, flores, hojas, raíces, cotiledones, y callos de cuatro

genotipos de melón, permitió obtener 71.577 ESTs a partir de genotecas enriquecidas y 22.179 ESTs a partir de genotecas de cDNA estándar. La integración de estos ESTs junto con los ESTs disponibles anteriormente permitió obtener un total de 24.444 unigenes (Clepet y col., 2011). Recientemente se han generado también datos de RNA-seq a partir de distintos genotipos y tejidos de melón, utilizando pirosecuenciación 454 (Blanca y col., 2011; Corbacho y col., 2013; Portnoy y col., 2011) o secuenciación mediante SOLiD (Blanca y col., 2012).

E) MICRO MATRICES DE ADN

La elaboración de micro matrices de ADN (*microarrays*) ha sido otra herramienta que se ha desarrollado en melón y que aporta información a nivel transcriptómico o de expresión de genes en tejidos y condiciones específicas. La micro matriz de melón se diseñó con 17.510 unigenes derivados de 33.418 ESTs (Mascarell-Creus y col., 2009). Estos ESTs fueron aislados a partir de distintos estadios de desarrollo de la planta o en situaciones ambientales específicas (González-Ibeas y col., 2007). Para validar la micro matriz se realizaron tres experimentos con distintas muestras de melón. En la primera, se comparó la expresión en fruto inmaduro (15 días después de la polinización) con la expresión de una muestra de fruto maduro (46 días después de la polinización) en la variedad T111 que permitió comparar los cambios transcripcionales durante el desarrollo de fruto. En una segunda hibridación se comparó tejido de raíz después de la inoculación con el hongo *Monosporascus cannonballus* con tejido de raíz no infectado en la variedad resistente *Cucumis melo ssp agrestis* Pat81. Finalmente en la tercera hibridación, se comparó una muestra de tejido de cotiledones infectados por el virus del mosaico del pepino (CMV) con una muestra de tejido no infectado de la variedad Tendral. (Mascarell-Creus y col., 2009). Recientemente la micro matriz de ADN se ha utilizado para descubrir cambios transcripcionales asociados a la resistencia al virus del mosaico de la sandía (WMV) en la accesión TGR-1551 (Gonzalez-Ibeas y col., 2012) o la resistencia a *Monosporascus cannonballus* (Roig et al 2013).

F) TILLING

En los últimos años se han desarrollado un conjunto de técnicas aplicables a la genómica funcional. Una de las más importantes es el desarrollo de plataformas para la caracterización de mutantes mediante TILLING. Las poblaciones de TILLING son poblaciones de mutantes elaboradas a partir de la mutagenización de las semillas o el polen, con un agente mutágeno. La primera prueba de mutageneización a gran escala en melón se hizo en la línea Noy Yizre-el, un melón tipo Galia, de la que se

obtuvieron 3000 familias M2. Los estudios fenotípicos detectaron mutaciones en distintos estadios de desarrollo y tejidos, incluidos cotiledones, hojas, flores y frutos (Tadmor y col., 2007). Sin embargo, la primera población de TILLING generada en esta especie es la que se presenta en este trabajo, obtenida a partir de una línea doble haploide M62-113 de melón tipo “Piel de Sapo” con un total de 3268 familias M2 (Gonzalez-To y col., 2011). En 2010 se publicó una nueva población en la cual se utilizó una variedad de melón tipo *Charentais* (CharMono) para desarrollar una plataforma de TILLING de 4023 familias M2. En esta población se buscaron mutantes en 11 genes, implicados en el proceso de maduración de fruto y en la determinación del sexo de las flores (Dahmani y col., 2010)

1.1.5.2 EL GENOMA DEL MELÓN

El melón es una especie diploide con una número básico de cromosomas $n=12$ ($2n=2x=24$) (Whitaker, 1976). El tamaño del genoma es de aproximadamente 450 Mb (Arumuganathan, 1991), tamaño similar al genoma del arroz (419 Mb) y casi cuatro veces el genoma de *Arabidopsis thaliana* (125 Mb). En el año 2009 se inició el proyecto MELONOMICS con el objetivo de obtener un borrador de la secuencia del genoma del melón. Actualmente, el genoma del melón ya ha sido publicado, habiéndose podido ensamblar aproximadamente un 83,3% del genoma, que contiene 27.427 genes (García-Mas y col., 2012).

1.1.5.3 LA ERA POST-GENÓMICA

En la última década, se ha podido obtener la secuencia del genoma de distintos organismos ([http:// www.ncbi.nlm.nih.gov/Genomes](http://www.ncbi.nlm.nih.gov/Genomes)). Sin embargo, todavía hoy se están desarrollando herramientas para poder explotar esta información. Los genomas están secuenciados pero todavía no se sabe con exactitud cómo identificar la funcionalidad de todos los genes que han sido anotados. Por lo tanto, en la mayoría de los genomas secuenciados existe un alto porcentaje de genes anotados bioinformáticamente sin función demostrada, denominados “putative” o “unknown” (López-Lara y col., 2006). Para poder anotar un genoma se combinan básicamente tres métodos: 1) Métodos *ab initio* de predicción de genes que permiten predecir la presencia de un gen. 2) Mediante homología con ESTs de la misma especie. 3) Mediante homología con proteínas de las cuales ya se ha determinado su función y que están presentes en las bases de datos. Para llegar a descifrar la función de cada gen anotado en el genoma hace falta desarrollar otros métodos como pueden ser las técnicas de genética inversa.

Además, con la secuenciación de genomas se ha podido observar que una gran parte del genoma contiene secuencia no codificante. Por ejemplo, en el genoma humano sólo el 1,5% es codificante para aproximadamente 25.000 genes (Wolfsberg, 2011). Esto llevó hace unos años a definir el “ADN basura” o el “junk DNA”. Se denominó así a todas aquellas secuencias del genoma que no codifican para una proteína o secuencias reguladoras, mayoritariamente transposones o elementos repetitivos. Inicialmente, se pensaba que esta parte del genoma no tenía ninguna función, sin embargo, se ha demostrado que parte de estas secuencias se transcriben y algunas de ellas regulan la expresión de los genes (The ENCODE (ENCyclopedia Of DNA Elements) Project Consortium, 2004) o que son importantes para estabilizar ciertas estructuras como son los telómeros y los centrómeros (Seffernick, 2009).

1.2 MUTACIONES

1.2.1 TIPOS DE MUTACIONES

Definimos mutación como un cambio en la información genética (ADN). Las mutaciones tienen importancia porque aportan variantes alélicas y variantes funcionales provocando fenotipos distintos del fenotipo silvestre. Las mutaciones son heredables, aunque en organismos multicelulares hay que distinguir entre mutaciones producidas en los gametos y las producidas en las células somáticas. Las mutaciones producidas en los gametos se transferirán a la siguiente generación mientras que las producidas en las células somáticas sólo afectarán al individuo que las padece y no tendrán un efecto hereditario.

La frecuencia de mutación en los genes depende de varios factores, aunque los más importantes sean la longitud del gen y el hecho de contener zonas con más probabilidad de ser mutadas (por tener una alta probabilidad de mutación debido a su composición química), como son las regiones CpG (Gruenbaum y col., 1981, Montelone, 1998). La frecuencia con la que la ADN polimerasa inserta una base equivocada en mamíferos es de 1×10^{-9} (Kumar y col., 2007) mientras que la frecuencia en los virus de ARN es de 1×10^{-6} (Tomas y Elena, 2010). En plantas no parece haber un estudio preciso de estos valores.

Las mutaciones naturales aportan variantes alélicas que pueden aportar nuevas propiedades a una especie. Un ejemplo ampliamente estudiado en plantas son las variantes alélicas del gen *eIF4E*. En melón, así como en pimiento (Ruffel y col., 2002), guisante (Gao y col., 2004) y otras especies se han detectado mutaciones naturales en el gen *eIF4E* responsables de la resistencia a distintos virus, como veremos más adelante. En todos los casos el virus no puede interactuar con el factor de iniciación de la traducción 4E y por lo tanto no puede replicarse e infectar la planta (Nieto y col., 2006).

1.2.1.1 ALTERACIONES DE LAS BASES

Las bases nitrogenadas son susceptibles a sufrir lesiones espontáneas en su conformación química. Estas alteraciones pueden producir varios efectos en la cadena molde afectando a la complementariedad de la misma o provocando otros tipos de alteraciones como las inserciones o las deleciones.

A) ERRORES DE REPLICACIÓN

A.1 Debidos a las formas tautoméricas

Cada una de las bases del ADN puede existir en la naturaleza en dos formas tautoméricas. Las diferencias químicas entre las dos formas tautoméricas consisten en el desplazamiento de átomos de hidrógeno que migran hacia posiciones más inestables. Normalmente, en el ADN se encuentra la forma *ceto*, mientras que las formas *imino* o *enol* son hasta 10.000 veces más inestables (Montelone, 1998). Existen dos tipos de reacciones tautoméricas, la Ceto-enólica (donde un grupo ceto (=O) se convierte en un grupo enol (-OH)) o Imina-amina (donde el grupo amina (=NH) se convierte en un grupo amino (-NH₂)). Las formas tautoméricas inestables tienen específicamente alteradas las propiedades de complementariedad en la cadena de ADN afectando en su replicación. En la Tabla 1.1 puede observarse como cambia la complementariedad de las bases.

Tabla 1.1: Representación de las formas tautoméricas de cada base que forma el ADN y de la complementariedad que adquieren.

Forma Tautomérica	Adopta propiedades de	Complementa con
A*	G	C
G*	A	T
C*	T	A
T*	C	G

Si en la cadena molde que se está replicando hay una forma tautomérica de una base se puede producir un error de complementariedad provocando la incorporación de una base que no es la correspondiente a la secuencia original y en consecuencia producir una mutación. Estos cambios pueden ser transiciones basados en el cambio de una purina a otra purina (A,G) o de una pirimidina a otra pirimidina (T, C). Las consecuencias de estos cambios dependen de la posición y del tipo de cambio que se produzca así como de su localización (en región exónica o intrónica). Si la base cambiada se encuentra en región exónica y no provoca un cambio de aminoácido en la traducción del codón, éste será un cambio silencioso. Normalmente, este tipo de cambio se produce cuando muta la tercera base del codón puesto que el código

genético es degenerado. Si el cambio de base provoca un cambio de aminoácido se denomina cambio sustitutivo o no sinónimo. Finalmente, si el cambio de base produce un codón stop se denomina cambio sin sentido dado que podría detenerse la traducción allí donde se de la mutación y en consecuencia, no poder sintetizar la proteína en su totalidad. Los cambios sustitutivos y sin sentido son los más graves ya que pueden afectar seriamente a la estructura y por lo tanto a la funcionalidad de la proteína.

A.2 Inserciones y deleciones

Una de las mutaciones espontáneas más habituales en secuencias repetidas largas son las deleciones de fragmentos. Streisinger y Owen (1985) formularon un modelo que explica los cambios de fase durante la síntesis de ADN, los cuales cambian la pauta de la lectura de la proteína. Si durante la replicación en la cadena molde tenemos una secuencia con una repetición de nucleótidos pueden ocurrir dos casos. El primero es que durante la replicación se cree un lazo en la cadena replicada. La replicación seguirá su curso y finalmente, puesto que el lazo ha provocado un corrimiento de la cadena replicada, se añadirá un nucleótido más de los que hay en la cadena molde. El segundo caso es si el lazo ocurre en la cadena molde. Al crearse un corrimiento en la cadena molde, la cadena que se está replicando tendrá un nucleótido menos puesto que donde se ha producido el lazo no se ha añadido la base complementaria. En el primer caso se obtendrá una cadena con un nucleótido de más, por lo tanto se habrá producido una inserción. En el segundo caso la cadena tendrá un nucleótido de menos y por lo tanto se habrá producido una deleción.

B) LESIONES ESPONTÁNEAS EN LAS BASES

B.1 Desaminación

Los nucleótidos pueden perder el grupo amino por desaminación oxidativa provocando una complementariedad errónea en la replicación del ADN. Dado que la desaminación de citosina produce uracilo, la célula tiene sistemas de reparación para reparar este error. Las glucosilasas de uracilo de ADN reconocerán los uracilos en la cadena, los escindirán y el hueco será rellenado por la ADN polimerasa I y finalmente ligado por una ligasa. Si el sistema de reparación no detecta y repara este error, en la cadena molde aparecerá un uracilo que posteriormente emparejará con una adenina en vez de la guanina que le correspondería (Alberts y col., 2006).

B.2 Despurinización

La despurinización, implica la interrupción del enlace glucosídico entre la base y la desoxiribosa y la pérdida posterior de un residuo de guanina o adenina del ADN. Si esta lesión persiste, conduce a un daño genético importante porque los sitios sin purina que se originan no pueden enlazar con su base complementaria durante la replicación. Sin embargo, sistemas de reparación eficaces eliminan los sitios sin purinas. Bajo ciertas circunstancias se puede insertar una base en un sitio sin purina; con frecuencia esto da como resultado una mutación (Alberts y col., 2006).

B.3 Oxidación

El metabolismo aerobio genera, como subproducto, ciertas especies activas de oxígeno, como radicales superóxido (O_2^-), peróxido de hidrógeno (H_2O_2) y radicales hidroxilo (OH^-). Todas ellas causan lesiones oxidativas en el ADN que dan lugar a mutaciones, las más frecuentes son G a T por oxidación de la guanina (Alberts y col., 2006).

1.3 MUTAGÉNESIS INDUCIDA

1.3.1 EL TÉRMINO MUTAGÉNESIS

Se denomina mutagénesis a la producción intencionada de variabilidad genética a través de métodos físicos o químicos. La utilización de estos métodos provoca cambios al azar en el ADN, y en consecuencia, también en los genes y en el organismo. En plantas, la mutagénesis generalmente se aplica en las semillas aunque en algunas especies como el maíz se ha aplicado en el polen. Este método implica que hace falta esperar, al menos, a la segunda generación para observar efectos fenotípicos (Kodym y Afza, 2003).

La mutagénesis puede ser dirigida o inducida. El primer término hace referencia a aquellas mutaciones precisas que se provocan en un gen deseado. El segundo, en cambio, hace referencia a la mutagénesis inducida por mutágenos químicos o físicos que provocan cambios al azar y de una manera generalizada en todo el genoma. Los primeros investigadores que produjeron artificialmente mutaciones en el ADN fueron Muller y Stadler en el 1927 y 1928 respectivamente con la demostración del efecto mutagénico de los rayos X en maíz, cebada y *Drosophila melanogaster* (Alberts y col., 2006).

1.3.2 LA MUTAGÉNESIS EN LA MEJORA GENÉTICA

La mejora genética se basa en observar y seleccionar los mejores individuos en función de los caracteres de interés para posteriormente fijar ese carácter en una población. De este modo se desarrollan poblaciones con caracteres óptimos para un objetivo comercial. La mejora genética usa como herramienta la variabilidad genética existente entre poblaciones y entre individuos. Esta fuente de variabilidad se genera espontáneamente cómo resultado de errores provocados en la división celular y en el sistema de reparación del ADN (la radiación solar tiene un importante papel en estos errores). Las mutaciones naturales no son abundantes y menos aun cuando llegan a afectar a la funcionalidad de las proteínas (en humanos la tasa de mutación es del orden de 10^{-6} mutaciones/locus/generación, aunque puede variar dentro del rango de 10^{-4} hasta 10^{-7} dependiendo del tamaño del gen (Dolgin, 2009)). Es por ello que a principios del siglo XX se empezaron a aplicar técnicas de mutagénesis para producir más variabilidad y de un modo más rápido. A principios de los años 50 se

empezaron a usar comercialmente cultivos mejorados a través de la mutagénesis. La FAO (*Food and Agriculture Organization*) empezó a documentar todos los cultivos obtenidos por mutagénesis y en 1994 la lista contenía más de 1800 cultivares (www.fao.org). Actualmente la mayoría de cultivos explotados en el mundo son y/o provienen de cruzamientos de los primeros cultivares mutados. Por ejemplo, en 1994 la FAO detalló que el 70% de trigo duro (*Triticum aestivum*) cultivado en Italia provenía de cultivos mutados.

La mutagénesis química y la radiación ionizante han sido utilizadas eficientemente para generar variabilidad genética que posteriormente ha sido utilizada en la mejora genética. Estas metodologías tienen la ventaja de ser más económicas y al mismo tiempo, más aceptadas socialmente que la tecnología transgénica. Actualmente, a través de estos métodos existen 2.428 variedades de cultivos mutagenizados en el mundo, 501 de ellos en arroz (<http://www-mvd.iaea.org/MVD>).

1.3.3 TIPOS DE MUTÁGENOS

Un mutágeno es un producto natural o elaborado por el hombre que puede alterar la estructura o la secuencia del ADN. Podemos clasificar estos agentes mutágenos en función de cómo producen las lesiones ya sea mediante un mecanismo físico o químico. La mutagénesis mediante mutantes químicos y radiaciones ionizantes se ha usado en plantas en genética directa (Shirley y col., 1992). Este tipo de métodos tienen ciertas ventajas en comparación con la mutagénesis insercional, ya que generan diversas series alélicas, inducen mutaciones con alta frecuencia y son aplicables a diversas especies (Upadhyaya y col, 2011). Por lo general, la mutagénesis química produce mutaciones a distintos niveles, ya sean mutaciones sin sentido o mutaciones sustitutivas. En cambio, la mutagénesis mediante radiación ionizante acostumbra a producir deleciones cromosómicas. Por lo tanto, la selección del mutágeno debe basarse en la eficiencia y la especificidad para inducir mutaciones y en la alta probabilidad de encontrar el tipo de mutación deseada (Tabla 1.2). También es importante conocer el tipo de mutación que produce el mutágeno para poder seleccionar el método de detección más eficiente (Koornneef y col, 1982).

Tabla 1.2: Resumen de los mutágenos más habituales con su mecanismo de acción y el tipo de mutación que producen.

Causa/mutágeno	Mecanismo	Tipo de mutación producida
Espontánea	Errores de replicación y de los mecanismos de reparación del ADN	Todo tipo de mutaciones
2-aminopurina (2AP)	Base análoga	A-T a G-C y G-C a A-T transiciones
5-bromouracilo	Base análoga	A-T a G-C y G-C a A-T transiciones
Etil Metano Sulfonato (EMS)	Agente alquilante, genera O ⁶ -metilguanina	G-C a A-T transiciones
Radiación UV	Dímeros de pirimidina	Mayormente cambios G-C a A-T transiciones, deleciones, cambios de pauta de lectura (menos frecuente)
Rayos X	Radiación ionizante	Deleciones cromosómicas
Bromuro de etidio	Agente intercalante	Cambios de pauta de lectura, duplicaciones y deleciones

1.3.3.1 MUTAGÉNESIS QUÍMICA

Dependiendo del modo de acción o mecanismos de los agentes químicos mutagénicos, podemos distinguir tres grupos distintos:

A) BASES ANÁLOGAS

Las bases análogas son estructuras químicas muy semejantes a las bases nitrogenadas ya sean purinas o pirimidinas y que por lo tanto pueden sustituir a éstas e incorporarse en el ADN durante la replicación y posteriormente provocar un mal emparejamiento de bases, perpetuando una mutación en la secuencia. Dos ejemplos de estos agentes son:

-5-bromouracilo: compuesto artificial usado en investigación. Su semejanza a la base nitrogenada timina hace que empareje con adenina.

-2-aminopurina: molécula semejante a la adenina y por lo tanto provoca emparejamiento con timina o a veces, menos frecuentemente, con citosina (Alberts y col., 2006).

B) AGENTES QUÍMICOS MODIFICADORES DE BASES

Existen varios agentes que provocan alteraciones en la estructura y/o en el emparejamiento de las bases nitrogenadas. Citaremos los ejemplos de aquellos más usados:

-Ácido nitroso: provoca desaminación de citosina a uracilo, de metilcitosina a timina y de guanina a otra base nitrogenada lo que provoca transiciones .

-Agentes alquilantes: los agentes alquilantes fueron la primera clase de mutágenos en ser descubiertos cuando Auerbach y Robson (1946) descubrieron los efectos mutagénicos del “gas mostaza” y sus componentes durante la segunda Guerra Mundial. Existen varios agentes alquilantes como el metilmetano sulfato (MMS), Etil Metano Sulfonato (EMS), y la nitrosoguanidina que afectan al ADN. Estos agentes son eficientes mutágenos y de fácil utilización. Es por ello que son habitualmente usados en plantas, en especial el EMS. El EMS alquila las guaninas, haciendo que el emparejamiento sea con la timina y no con la citosina y por lo tanto provoca transiciones del tipo G/C-A/T (Sega, 1984) (Figura 1.3).

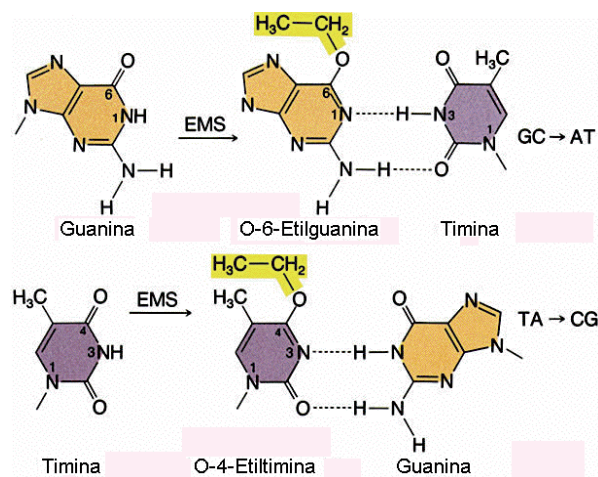


Figura 1.3: Efecto del EMS a nivel molecular. Las guaninas y las timinas se metilan y ese cambio químico hace que se comporten como una adenina y una citosina respectivamente y en consecuencia complementan con bases distintas a las que deberían complementarse (fuente: <http://www.ucm.es/info/genetica/grupod/Mutacion/mutacion.htm>).

C) AGENTES INTERCALANTES

Los agentes intercalantes son moléculas que tienen analogía con los nucleótidos y por lo tanto, son capaces de deslizarse entre éstos e intercalarse en la doble hélice haciendo que la polimerasa se equivoque y añada otra base produciendo deleciones o corrimientos de lectura. Un ejemplo de estos agentes son la proflavina y la naranja de acridina, así como el bromuro de etidio (Alberts y col., 2006).

1.3.3.2 MUTAGÉNESIS FÍSICA

A) RADIACIONES ELECTROMAGNÉTICAS

Las radiaciones electromagnéticas están formadas por la combinación de campos magnéticos y eléctricos y se propagan a través del espacio transportando energía. La radiación electromagnética, cuya longitud de onda está entre los 400 y los 15 nm se denomina UV (ultravioleta). La longitud de onda del UV cercano está entre los 200-400 nm. Esta es la misma longitud de onda que absorbe el ADN, con un óptimo en 254 nm que es la absorción máxima. Esta característica hace que el UV pueda provocar lesiones en el ADN. Las mutaciones producidas más comúnmente son dímeros de iminas y estas duplicaciones provocan que en la replicación la ADN polimerasa inserte un nucleótido incorrecto en esa posición.

B) RADIACIONES IONIZANTES

Las radiaciones ionizantes son aquellas radiaciones capaces de ionizar la materia, extrayendo los electrones de sus estados electrónicos que les mantenían ligados al átomo. Entre las potentes especies químicas formadas por la radiación ionizante se encuentran los radicales libres, siendo el más importante el radical hidroxilo (OH·), que reacciona con macromoléculas, como el ADN, y las inactiva produciendo roturas que pueden dar lugar a cambios estructurales importantes, como son las mutaciones cromosómicas (Shirley, 1992).

La radiación ionizante ha sido ampliamente utilizada para inducir mutaciones en la mejora genética (Bruggemann y col., 1996; Cecchini y col., 1998; Shikazono y col., 1998 y 2001). En plantas se ha observado que el uso de radiaciones ionizantes produce mayormente deleciones y modificaciones cromosómicas. Las radiaciones ionizantes pueden provenir de sustancias que emiten estas radiaciones de forma espontánea o bien de generadores artificiales como por ejemplo los generadores de rayos X o los aceleradores de partículas.

-Radiaciones Gamma y rayos-X: Las radiaciones con rayos gamma-(γ) y rayos-X producen deleciones y otras mutaciones cromómicas, aunque sólo algunas de ellas han sido caracterizadas a nivel molecular (Oppenheimer y col., 1991; Shirley y col., 1992; Nambara y col., 1994). Los rayos gamma-(γ) y los rayos-X han sido usados para producir mutantes en *Arabidopsis* (Rédei y Koncz, 1992) y en arroz (Manzoor y Cheema, 2003). Wu y colaboradores (2005) usaron dos dosis diferentes de rayos gamma-(γ) (250 y 500 Gy) para producir una colección de mutantes en IR64 (variedad *indica* del arroz) en los cuales midieron el tamaño de las deleciones (en Kilobases) comprobando las alteraciones cromosómicas (Wu y col., 2005).

-Neutrones rápidos: El método de genética inversa basado en el *fast neutron* (inductor de deleciones) ha sido desarrollado para identificar y aislar genes de interés en plantas (Li y col., 2001; Li y Zhang, 2002; Wu y col., 2005). De acuerdo con Li y col. (2001), este método produce deleciones con eficiencia y también se pueden producir líneas mutantes con una cobertura de mutación más eficiente que con la mutagénesis insercional que ofrece la tecnología del T-DNA. Este método ha sido usado en arroz, produciendo 8,000 líneas mutantes en la variedad *indica* IR64 (Wu y col., 2005).

-Iones acelerados: Usando iones de carbón ionizados, se han conseguido aislar nuevos mutantes en *Arabidopsis* (*ast*, *frl1*, *uvi1*, *suw1*, *tt18*, y *tt19*) (Tanaka y col., 1997; Hase y col., 2000; Sakamoto y col., 2003; Shikazono y col., 2003). Analizando la secuencia nucleotídica de los mutantes se descubrió que las mutaciones producidas eran inversiones, translocaciones, y pequeñas deleciones (Shikazono y col., 1998, 2003, Sakamoto y col., 2003). Estudiando con detalle la frecuencia de estos cambios, se obtuvo que la frecuencia de las deleciones era de 6.1×10^{-5} , frecuencia comparable a la inducida por los neutrones rápidos (Li y col., 2001). Estos resultados indicaron que la mutagénesis mediante iones acelerados podía ser usada tanto para la genética directa como para la genética inversa de plantas.

1.3.3.3 OTROS MÉTODOS DE MUTAGÉNESIS

A) TRANSPOSONES

Los elementos transponibles tienen la capacidad de desplazarse por el genoma e integrarse en una región. Si la región donde se integran es un gen, éste puede perder su funcionalidad debido a que se altera su secuencia. Los transposones son moléculas de ADN con capacidad móvil. Tienen la capacidad de saltar de una posición a otra del genoma. Los transposones se han clasificado en función de su mecanismo de transposición, dependiendo de si este proceso se da con un paso intermedio de transcripción (retrotransposones) o si simplemente se desprenden y se acoplan en regiones diversas del genoma (transposones de ADN). La utilización de transposones como método de mutagénesis insercional ha sido usada como tecnología de genética inversa (Hirochika y col., 2004; McCarty y col., 2005).

B) T-DNA (ADN DE TRANSFERENCIA)

Agrobacterium tumefaciens es una bacteria que puede infectar a más de 330 géneros de dicotiledóneas (Izquierdo y col., 2001). La bacteria puede infectar a través de una herida o una incisión transfiriendo en ella el plásmido *Ti* (inductor de tumores) y provocándole a la planta un tumor calloso denominado "tumor de la corona". El plásmido *Ti* mide entre 150-200 Kb de las cuales 20-23 se integran aleatoriamente en varios lugares del genoma huésped. A este ADN se le denomina T-ADN (ADN transferido) (Izquierdo y col., 2001). Esta propiedad de integración en el genoma se ha usado para el aislamiento de genes en plantas. El T-ADN introduce mutaciones al azar llegando a inactivar la función de un gen. Mediante esta tecnología es posible aislar el gen que es responsable de la falta de función con sondas específicas para el T-ADN y al mismo tiempo, relacionarlo con un fenotipo mutante (Izquierdo y col., 2001). Esta tecnología de mutagénesis insercional se ha utilizado tanto en *Arabidopsis thaliana*, como en arroz (<http://signal.edu/cgi-bin/tdnaexpress> y <http://signal.salk.edu/cgi-bin/RiceGE>) y en muchas otras especies (An y col., 2003), entre ellas en sandía, para obtener resistencias a virus como el virus del mosaico amarillo del calabacín (ZYMV) (Yu y col., 2011).

C) VIRUS

Los virus son moléculas infectivas que pueden clasificarse en función de la naturaleza de su material genético (mono o bicadena de ARN o ADN según la clasificación de Baltimore (Baltimore y Smore, 1971)). Los virus infectan a las plantas a través de plagas (insectos) u hongos. La utilización de virus en plantas como

vectores puede ofrecer grandes ventajas. Se multiplican rápidamente y pueden producir un gran número de copias del transgén (insertado previamente en el virus) alcanzando valores máximos en una o dos semanas después de la infección (Izquierdo y col., 2001). Aunque se han descubierto varios virus en plantas que pueden usarse como vectores (por ejemplo el virus del jaspeado del tabaco (TEV) (Agudelo-Romero y col., 2008)) su utilización no es muy frecuente. El mayor problema que presentan los virus en plantas es que si el gen insertado supera la Kilobase, su inestabilidad es muy alta y la planta lo elimina (Izquierdo y col., 2001), (Wright y Voytas, 1998).

Gracias al mecanismo infectivo de los virus combinado con los mecanismos de defensa de las plantas se ha podido desarrollar la técnica denominada Silenciamiento Génico Inducido por Virus (VIGS) (Lu y col., 2003). Cuando el virus infecta a la planta ésta desencadena un mecanismo de defensa basada en la interacción con el ARNm del virus y en consecuencia se bloquea su transcripción. Sin embargo, si se usa el virus como vector insertándole un gen de la planta hospedadora el proceso irá dirigido contra los ARNm del mismo gen silenciando así su expresión (Fu y col., 2005).

D) MUTAGÉNESIS DIRIGIDA

Una de las recientes tecnologías descritas de mutagénesis dirigida en plantas es la denominada tecnología TALENs (*Transcription Activator-Like Effector Nucleases*). Existen patógenos bacterianos del género *Xanthomonas* que infectan a una amplia gama de especies y que durante la infección, proporcionan a las células de la planta una serie de proteínas conocidas como *Transcription Activator-Like Effectors* (Boch y Bonas, 2010). Estas proteínas se unen a promotores de genes diana de la planta y de este modo modifican el transcriptoma de la planta huésped. Para unirse al ADN de la planta, TALE contiene un dominio de unión compuesto por varias repeticiones de monómeros de 34 aminoácidos cada uno de los cuales sólo dos residuos, que están localizados en las posiciones 12 y 13, son hipervariables. Cada motivo hipervariable TALE reconoce un par de bases en una secuencia de ADN.

TALENs se ha creado mediante la fusión de un dominio de unión (TALE) con una nucleasa FokI. De este modo el complejo puede reconocer una secuencia concreta y producir un corte de doble cadena en un sitio específico para inducir mutagénesis dirigida, como por ejemplo la inserción de un fragmento exógeno de doble cadena induciendo a la represión o activación del gen dependiendo de la función del

dominio donde se ha localizado la inserción de TALENs. Esta tecnología de mutagenesis dirigida ya se ha usado con éxito en varias especies como por ejemplo en *Arabidopsis thaliana* (Cermak y col., 2011).

1.4 TILLING

1.4.1 INTRODUCCIÓN

El TILLING (*Targeting Induced Local Lesions IN Genomes*) es una técnica de genética inversa que combina la generación de mutaciones puntiformes (o SNPs) inducidas a través de un agente mutágeno con una eficiente técnica de rastreo molecular, que permite identificar los polimorfismos a nivel de secuencia nucleotídica en diversas especies vegetales (McCallum y col., 2000 a, b). Gracias a esta técnica es posible localizar, en una población constituida por miles de individuos, una serie alélica de mutaciones puntiformes inducidas en genes de interés (Till y col., 2003).

La técnica del TILLING fue utilizada por primera vez por McCallum en la planta modelo *Arabidopsis thaliana* (McCallum y col., 2000a) y desde entonces ha sido una técnica utilizada con éxito en varios organismos ya sean vegetales (trigo duro (Slade y col., 2004); cebada (Talamè y col., 2008); arroz (Till y col., 2007); maíz (Till y col., 2004a); soja (Cooper y col., 2008); guisante (Dalmais y col., 2008); melón (Dahmani-Mardas y col., 2010; Gonzalez-To y col., 2011); *Lotus japonicus* (Perry y col., 2003; Gilchrist y Haughn, 2005)) o animales (pez-zebra (Wienholds y col., 2003); *Drosophila melanogaster* (Winkler y col., 2005); *Caenorhabditis elegans* (Gilchrist y col., 2006)), independientemente de las dimensiones de los genomas.

Esta técnica tiene algunas características que la hacen peculiar y ventajosa. Destacaremos tres por su relevancia. La primera, es que no se requiere el uso de transformación genética y por lo tanto los mutantes obtenidos no son transgénicos. Esta característica es una ventaja puesto que los transgénicos no están aceptados socialmente en Europa y esto perjudica la comercialización del producto de origen transgénico. La segunda es que es una técnica que puede ser automatizada y optimizada permitiendo analizar un gran número de muestras en un tiempo mínimo. Finalmente, gracias a este método se pueden obtener no sólo mutaciones sin sentido, sino también mutaciones del tipo sustitutivo las cuales pueden afectar a la función del gen sin llegar a inactivarlo, es decir, creando series alélicas.

Desde un punto de vista teórico, la técnica TILLING consiste en individualizar las variantes alélicas que contienen una mutación significativa (por ejemplo, que produzcan cambios de aminoácido) mediante un rastreo molecular. Además, se

pueden seleccionar nuevos genotipos que provocan fenotipos que pueden ser útiles a nivel comercial. El TILLING es una técnica aplicable a todos los organismos ya que su eficiencia no depende del tamaño del genoma, ni del sistema reproductivo, ni del ciclo biológico (Gilchrist y Haughn, 2005). Otra ventaja es que las mutaciones se producen a través de mutagénesis química y por lo tanto, son mutaciones irreversibles y distribuidas aleatoriamente a lo largo de todo el genoma. Las poblaciones ideales son aquellas en las que se pueden identificar mutantes para cada gen diana analizado. Además, mediante esta técnica se pueden analizar también diferencias genotípicas naturales entre variedades de una misma especie lo que se denomina, más específicamente, EcoTILLING (de *Ecotype* TILLING) (Comai y col., 2004).

1.4.2 LA TÉCNICA DE TILLING

La técnica del TILLING se divide en tres fases: 1) Producir la población mutagenizada, 2) Desarrollar la plataforma: extracción y preparación del ADN para el rastreo, y

3) Analizar y detectar los mutantes para los genes de interés en la población.

En primer lugar se desarrolla la población mutando las semillas M0 y creciendo la generación M1 (Figura 1.4). Posteriormente, a través de la autofecundación de los individuos M1 se obtiene la generación M2. Las semillas de la generación M2 se recogen y se almacenan de modo que una vez identificadas las familias mutantes se puedan volver a plantar individuos M2 para confirmar el fenotipo y la segregación de la mutación. Una vez obtenida la generación M2 también se procede a la extracción del material genético. Este material genético es el resultado de la extracción de ADN de la mezcla de las muestras de hoja de diez individuos de cada familia M2. Después de extraer el ADN de todas las familias que componen la población, se organiza en placas mezclando el ADN de cada cuatro a ocho familias para poder agilizar el proceso de detección de mutantes (mezcla de ADNs o pools). Para el desarrollo de algunas poblaciones este primer paso en el que se obtiene el ADN de la población ha variado ligeramente. Algunos autores han extraído el ADN de tan solo una planta de la generación M2 y no de la mezcla de diez individuos, reduciendo así considerablemente el tiempo de la preparación de la población, aunque en este caso se suele autofecundar el individuo M2 para perpetuar la mutación en la siguiente generación.

A continuación se inicia el proceso de detección de mutantes en genes de interés. Para ello, se diseñan cebadores de modo que amplifiquen las zonas del gen que se quieren analizar. Mediante el diseño de cebadores específicos se puede proceder a la amplificación del gen (o de una región del gen) mediante PCR. La PCR se realiza sobre la mezcla de ADNs mencionada anteriormente, por lo tanto, amplificará todos los alelos que se encuentren en esa mezcla (tanto el alelo silvestre como el mutado). Una vez desnaturalizado el producto de la PCR se procede a la renaturalización. En el caso de la renaturalización entre una cadena mutada con una silvestre la complementación no será total y por lo tanto se producirá un heterodúplex en el nucleótido mutado. Posteriormente, una endonucleasa que digiere exclusivamente heterodúplex (*Cel I*) será la responsable de cortar las cadenas donde haya una mutación. Después se visualizan los fragmentos digeridos en un LI-COR 4300. La detección de las mutaciones o SNPs puede hacerse mediante otros métodos que se comentarán posteriormente, como por ejemplo, el HRM (*High Resolution Melting*), aunque el método más habitual es el uso del LI-COR 4300, utilizado en este trabajo. Gracias a esta tecnología se pueden visualizar dos canales distintos en los cuales se detectan los fragmentos amplificados desde cada cebador 5' o 3' de la cadena hasta el punto de digestión. De este modo, se puede confirmar la mutación puesto que la suma del tamaño de los fragmentos digeridos en cada canal debe coincidir con la totalidad del tamaño del fragmento amplificado sin digerir (la explicación de la visualización de los fragmentos digeridos en el LI-COR está detallada en la Figura 1.5).

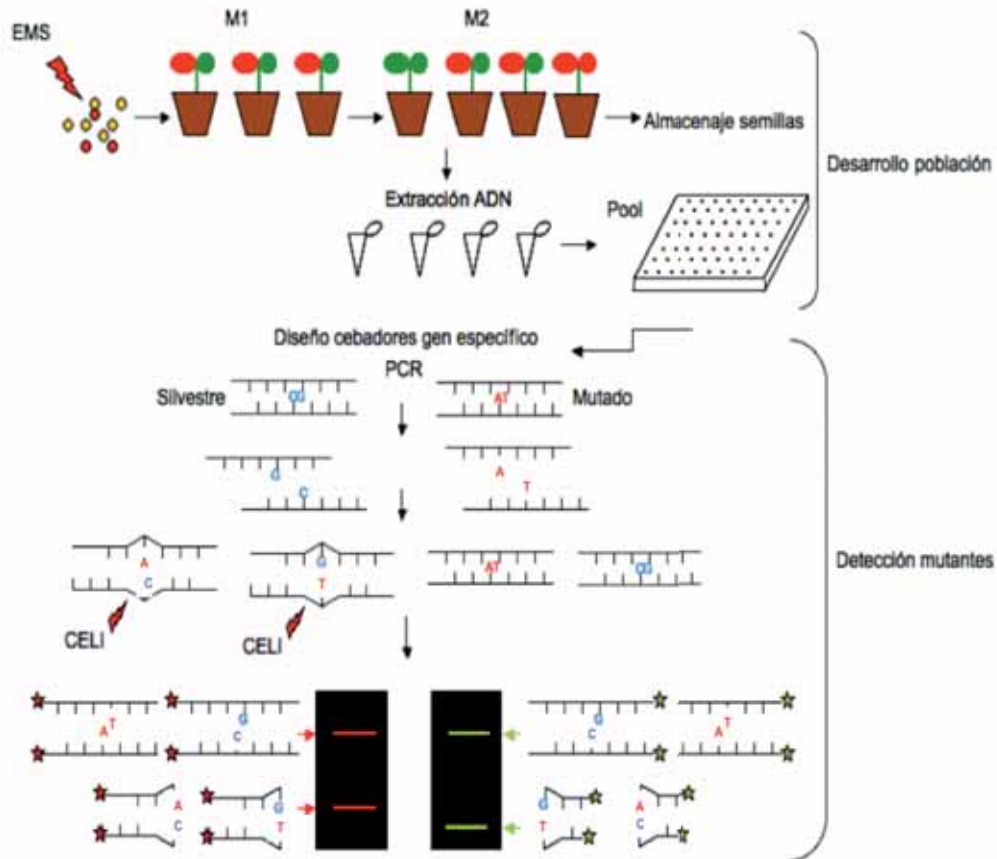


Figura 1.4: Protocolo empleado para el método de TILLING. Una vez mutadas las semillas (M0) con el mutágeno EMS, se crece la generación M1. Mediante la autofecundación de esta generación se obtiene la generación M2. De cada familia M2 se crecen 10 individuos y se les toma una muestra de hoja. Posteriormente, se extrae el ADN de esa mezcla y de este modo se obtiene el ADN de cada familia M2. De la generación M2 también se almacenan las semillas para experimentos futuros. Los ADNs de todas las familias se organizan en placas mezclándolas de cuatro en cuatro, en nuestro caso. Para la detección de los mutantes, primero se amplifica el gen de interés. Después se desnaturalizan las cadenas y de este modo se provocan los heterodúplex que digerirá la endonucleasa *Cell*. Finalmente, la detección de los fragmentos de digestión se realiza usando el LI-COR 4300.

1.4.2.1 DESARROLLO DE LA POBLACIÓN MUTAGENIZADA

Las plantas son organismos ideales para mutagenizar ya que tienen la capacidad de mantener al organismo en un estado latente con la semilla. El mutágeno ideal para usar en el método de TILLING es aquél que induce mutaciones aleatoriamente

abarcando regiones codificantes y en una alta frecuencia ya sean puntuales, pequeñas deleciones o inserciones. El mutágeno químico EMS (Étil Metano Sulfonato) genera SNPs (transiciones G:C-A:T (Greene y col., 2003, Slade y col., 2005)) y se puede aplicar, optimizando la dosis en cada especie, para producir mutaciones sin producir letalidad de las semillas o de las plantas (Greene y col., 2003, Koornneef y col., 1982). El procedimiento habitual usado hasta ahora para producir plataformas de TILLING ha sido introducir las semillas en una solución con el mutágeno diluido durante aproximadamente 10-24 horas y en oscuridad, así se consigue mutar las células del embrión que más tarde producirán los tejidos reproductivos. En la mayoría de los casos los individuos de la generación M1 tendrán la mutación en heterocigosis y probablemente sólo en una parte de la planta. Para asegurar la probabilidad de encontrar la mutación en homocigosis hace falta autofecundar la generación M1 y obtener la generación M2. De este modo la probabilidad de encontrar un individuo homocigoto sigue la proporción de herencia mendeliana 1:2:1 (Henikoff y Comai, 2003). También en algunas especies como en maíz se ha aplicado el mutágeno directamente en el polen para luego introducirlo en la flor femenina silvestre de la misma variedad. Las semillas de las mazorcas resultantes se germinan para obtener la generación M1. Posteriormente mediante autofecundación se obtiene la M2 (Till y col., 2003).

Desde el año 2000 se han desarrollado poblaciones de TILLING en distintas especies. Uno de los factores decisivos ha sido optimizar el protocolo de mutagenización de las semillas con el objetivo de obtener una densidad de mutación alta y la vez, que no afecte a la viabilidad de las plantas mutageneizadas. Para determinar la dosis del mutágeno (mayormente el EMS) en cada especie, ha sido necesario hacer experimentos previos para poder determinar la dosis óptima. Es por ello que encontramos distintos rangos de dosis en función de la especie (Tabla 1.3). Para *Drosophila melanogaster* se determinó que la concentración óptima era de 50 mM (Winkler y col., 2005), mientras que en *Arabidopsis thaliana* está dentro del rango de 20-40 mM (Greene y col., 2003). Recientemente se han publicado poblaciones de TILLING en melón que han permitido ajustar la dosis de mutágeno para esta especie (Dahmani-Mardas y col., 2010; Gonzalez-To y col., 2011). Para el melón se ha comprobado que la dosis de hasta el 3% de mutágeno proporciona una población con una frecuencia de mutación de 1/573 Kb (Dahmani-Mardas y col., 2010).

Tabla 1.3: Poblaciones de TILLING desarrolladas en animales y plantas por orden cronológico. Se pueden observar las distintas dosis de mutágeno (en algunos casos expresadas en mM, en otras en % de v/v dependiendo de la publicación) utilizadas en genomas de diferente tamaño y con diferentes ploidías. En sombreado gris las poblaciones de melón desarrolladas hasta el momento, una de ellas la elaborada en este trabajo (Gonzalez-To y col., 2011). *ENU: N-etil-N-nitrosourea, **Az-MNU: Ácido de sodio y metil-nitrosourea, ***NMU: N-nitroso-N-metilurea.

Especie	Mutágeno	Dosis de mutágeno	Tamaño del genoma	Ploidía	Tasa de mutación	Referencia
<i>Drosophila melanogaster</i>	EMS	50 mM	180 Mb	2X	1/209 kb	Bentley y MacLennan, 2000
<i>Danio rerio</i>	ENU*	3.0 mM	1,700 Mb	2X	1/235 kb	Wienholds y col., 2003
<i>Arabidopsis thaliana</i>	EMS	20-40 mM	125 Mb	2X	1/300 kb	Greene y col., 2003
<i>Hordeum vulgare</i>	EMS	20-30 mM	5,300 Mb	2X	1/Mb	Caldwell y col., 2004
<i>Zea mays</i>	EMS	1 %	2,500 Mb	2X	1/500 kb	Till y col., 2004,a
<i>Oryza sativa</i>	EMS	0.8-1 %	430 Mb	2X	0.5/Mb	Wu y col., 2005
<i>Oryza sativa</i>	EMS	1.6 %	430 Mb	2X	1/Mb	Wu y col., 2005
<i>Triticum turgidum subsp. durum</i>	EMS	0.75-1%	12,000 Mb	4X	1/40 kb	Slade y col., 2005
<i>Triticum aestivum</i>	EMS	0.75-1.2%	17,000 Mb	6X	1/24 Kb	Slade y col., 2005
<i>Drosophila melanogaster</i>	EMS	50 mM	180 Mb	2X	1/156 kb	Winkler y col., 2005
<i>Drosophila melanogaster</i>	EMS	125 mM	180 Mb	2X	1/90.5 kb	Winkler y col., 2005
<i>C. elegans</i>	EMS	25 mM	100 Mb	2X	1/293 kb	Gilchrist y col., 2006
<i>Zea mays</i>	EMS	0.0625 %	2,500 Mb	2X	0,93/kb B73	Weil y Monde, 2007
<i>Zea mays</i>	EMS	0.0625 %	2,500 Mb	2X	2.10/kb W22	Weil y Monde, 2007

<i>Pisum sativum</i>	EMS	4 mM	4,300 Mb	2X	1/669 kb	Triques y col., 2007
<i>Oryza sativa</i>	EMS	1.5 %	430 Mb	2X	1/294 kb	Till y col., 2007
<i>Oryza sativa</i>	Az-MNU**	1 mM Az- 15 mM MNU	430 Mb	2X	1/265 kb	Till y col., 2007
<i>Glycine max</i>	NMU***	2.5 mM	1,115 Mb	2X	1/140 kb	Cooper, 2008
<i>Glycine max</i>	EMS	50 mM	1,115 Mb	2X	1/250 kb	Cooper, 2008
<i>Glycine max</i>	EMS	40 mM	1,115 Mb	2X	1/550 kb	Cooper, 2008
<i>Arabidopsis thaliana</i>	EMS	0,2 %	125 Mb	2X	1/415 Kb	Bush y Krysan, 2010
<i>Avena sativa</i>	EMS	0,9 %	13,000 Mb	6X	1/30 Kb	Chawade y col., 2010
<i>Brassica rapa</i>	EMS	0,3 y 0,4 %	283,8 Mb	2X	1/60 Kb	Stephenson y col., 2010
<i>Solanum lycopersicum</i>	EMS	1%	950 Mb	2X	1/322 kb	Minoia y col., 2010
<i>Solanum lycopersicum</i>	EMS	0,7%	950 Mb	2X	1/574 kb	Minoia y col., 2010
<i>Cucumis melo</i>	EMS	1-3%	450 Mb	2X	1/573 kb	Dahmani-Mardas, 2010
<i>Arachis hypogaea</i>	EMS	0,4 y 1,2 %	-	4X	1/967 Kb	Knoll y col., 2011
<i>Helianthus annuus</i>	EMS	0,7 %	3189 Mb	2X	1/475 Kb	Sabetta y col., 2011
<i>Cucumis melo</i>	EMS	0,5-1,5%	450 Mb	2X	1/ 1,5 Mb	Gonzalez-To y col., 2011

1.4.2.2 DESARROLLO DE LAS MEZCLAS DE ADN

Otro factor que puede limitar la identificación de mutantes es el diseño de las mezclas de ADNs (pools). La diferencia entre hacer mezclas de cuatro familias o mezclas de ocho es el coste y el tiempo necesarios para la detección. Los factores que condicionan la eficiencia en la detección de mutantes en una mezcla son la calidad del ADN genómico, la precisión de su cuantificación y el método usado para realizar la detección de SNPs. En las diversas plataformas publicadas se han diseñado habitualmente mezclas de dos hasta ocho familias, aunque la óptima es la mezcla de ocho ya que reduce hasta la mitad el coste y el tiempo de rastreo de la población

respecto al uso de mezclas de cuatro familias. Una vez detectada una mutación en una mezcla se debe repetir el análisis en cada ADN individual que compone la mezcla. Este proceso se denomina deconvolución y es necesario para individualizar la familia responsable de la mutación (Colbert y col., 2001). Una alternativa para la detección de mutantes es realizar mezclas en dos dimensiones de manera que cada muestra está presente en dos únicos pools. Esta estrategia ya se ha usado a gran escala en el *MaizeTILLING service* (<http://genome.purdue.edu/maizetilling/>). Aunque duplicar las mezclas reduce la velocidad de la detección de los mutantes también reduce la incertidumbre (descarta falsos positivos o artefactos). Además, el paso de la detección individual de la familia o deconvolución no es necesario.

1.4.2.3 DETECCIÓN DE MUTANTES

Son varias las tecnologías que existen para la detección de SNPs y que han sido usadas para detectar mutaciones en plataformas de TILLING. Desde la electroforesis con tinción de bromuro de etidio (Comai y Henikoff, 2006), el HPLC (*High performance liquid chromatography*) (McCallum y col., 2000b), la espectroscopía de masas, la electroforesis capilar en gradiente en condiciones desnaturizantes o la digestión enzimática a nivel nucleotídico (Comai y Henikoff, 2006), hasta las últimas tecnologías como la *High Resolution Melting* (HRM) (Dong y col., 2009), la secuenciación masiva *Next Generation Sequencing* (NGS) (Rigola y col., 2009) o la utilización de MALDI-TOF (*Matrix-Assisted Laser Desorption/Ionization-Time Of Flight*) (Chawade A y col., 2010). Cualquiera de estas técnicas se puede usar en el TILLING para la detección de mutantes, sin embargo, el método seleccionado debe permitir un análisis masivo y rápido, así como de máxima sensibilidad.

Todos los métodos citados anteriormente han sido usados para la detección de SNPs (McCallum y col., 2000a, Colbert y col., 2001, Wienholds y col., 2003) pero el método más común para el TILLING ha sido la detección de las mutaciones a partir de la digestión con nucleasas específicas de heterodúplex (Figura 1.5). El procedimiento común es amplificar una región del gen de interés mediante PCR con cebadores específicos y marcados con un fluoróforo distinto para cada extremo 5' (IRDye 700) y 3' (IRDye 800). El producto de la PCR es desnaturizado y renaturalizado de modo que los alelos mutados se mezclan con los alelos silvestres y forman los heterodúplex. Al incubarse conjuntamente con la endonucleasa específica de heterodúplex se digieren sólo aquellos fragmentos que contienen una mutación.

La visualización se realiza mediante geles verticales de acrilamida en plataformas como el LI-COR DNA analyzer tal y como puede verse en la Figura 1.5. En primer lugar el fragmento del gen de interés seleccionado es amplificado mediante un cebador directo marcado fluorescentemente con IRDye 700, y con un cebador reverso marcado fluorescentemente con IRDye 800. Después de la formación del heterodúplex las muestras son incubadas con *Cell* para, posteriormente, ser desnaturalizadas (detallado en la Figura 1.4) y cargadas en un gel vertical de acrilamida para visualizarlas mediante el LI-COR 4300. Para cada gel se producen dos imágenes: una producida por la señal que emite el fluoróforo incorporado en el cebador directo (IRDye 700, rojo) y la otra la señal producida por el fluoróforo incorporado en el cebador reverso (IRDye 800, verde). Por lo tanto, en el canal del IRDye 700 podremos ver el fragmento que abarca la longitud desde el extremo 5' hasta el punto de la digestión (que es el mismo lugar donde se ha formado el heterodúplex y por lo tanto también donde está la mutación). Mientras que en el canal IRDye 800 podremos ver el fragmento que abarca la longitud desde el extremo 3' hasta el punto de la digestión. La suma del tamaño del fragmento digerido en el canal IRDye 700 y el tamaño del fragmento digerido en el canal IRDye 800 tiene que coincidir con el tamaño total del fragmento amplificado tal como puede observarse en el gel de la Figura 1.5 ($600+400=1000$ pb).

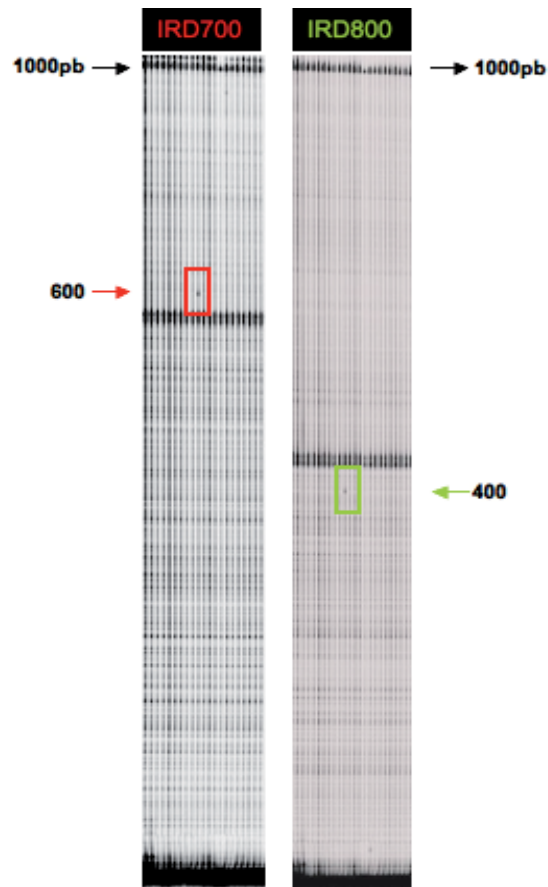


Figura 1.5: Detección de una mutación en una población de TILLING en melón en un gel de acrilamida vertical de LI-COR 4300. A la izquierda se observa el gel visualizado en el canal IRD700, a la derecha en el canal IRD800. En cada canal se visualiza el fragmento digerido desde uno de los extremos, ya que el cebador directo (5', rojo) lleva el marcaje IRD700 mientras que el cebador reverso (3', verde) lleva el marcaje para el IRD800. La suma del tamaño de los fragmentos detectados en cada canal (IRD700: 600 pb + IRD800: 400 pb) debe coincidir con el tamaño total del fragmento sin digerir visualizado en el extremo superior del gel (1000 pb).

Para el proceso de digestión del heterodúplex se han descrito algunas endonucleasas (Fuhrmann y col. 2005, Youil y col., 1995, Mashal y col., 1995, Till y col., 2004b). La endonucleasa más comúnmente utilizada para el TILLING es *CelI* extraída de apio (*celery*) (Oleykowski y col., 1998). Es una endonucleasa que actúa digiriendo el ADN de cadena sencilla. Se ha demostrado también que otras endonucleasas como las extraídas de judía son eficientes para detectar mutaciones mediante el método TILLING (Till y col., 2004b). La elección de la enzima también puede ser un factor

limitante para la sensibilidad del método, y se demostró que la utilización de extracto de *Cell* juntamente con el método de visualización de LI-COR 4300 era el más sensible para la detección de SNPs en mezclas de 8 familias (dilución 1/16 del alelo mutado en una especie diploide) (Greene y col., 2003). La enzima *Cell* reconoce los heterodúplex con la preferencia $C/C \geq A/C \sim T/C \geq G/G$, y reconoce con menor especificidad $A/C \sim A/A \sim T/C > T/G \sim G/A$ y todavía con menos T/T (Oleykowski y col., 1998; Yang y col., 2000). Al mismo tiempo, esta enzima tiene una buena especificidad en el reconocimiento de heteroduplex en una dilución alélica de 1:16. Es por ello, que se utiliza para la detección de SNPs en mezclas de 8 familias en especies diploides. En cambio, la enzima *EndoI* (clonada a partir de un gen que codifica por una endonucleasa identificada en *Arabidopsis*), reconoce todos los tipos de mutaciones con la misma especificidad en diluciones alélicas de 1:64 (Triques y col., 2007). En cuanto al aspecto económico, la obtención comercial de ambas endonucleasas tienen un coste alto, sin embargo, para la obtención de *Cell* se han publicado protocolos para obtener un extracto de apio que funciona eficientemente y reduce el coste comparado con la obtención de la enzima purificada (Yang y col., 2000). Durante los últimos años se han hecho estudios comparando la eficiencia de ambas endonucleasas en la detección de polimorfismos mediante EcoTILLING. Los resultados obtenidos son idénticos demostrando que la eficiencia es prácticamente igual (Bono y col., 2011). Por lo tanto, la utilización de una u otra endonucleasa dependerá del tipo de SNP esperado según el mutágeno utilizado, el coste de la misma y de la región a estudiar (Yeung y col., 2005; Dodgson y Wells, 1977).

Una ventaja importante del método enzimático, confrontado con las otras técnicas citadas anteriormente, es que a la vez que se detecta la mutación, también se detecta su localización dentro del gen. Posteriormente, mediante la secuenciación del fragmento en la familia mutante se puede visualizar en el cromatograma si la mutación se encuentra en homocigosis o heterocigosis.

A partir del año 2009 se empezó a usar otro método de detección de los fragmentos mutados: la tecnología de *High Resolution Melting* o Análisis de curvas de disociación del ADN (Dong y col., 2009; Gady y col., 2009). Esta técnica se basa en monitorizar la cinética de disociación de los fragmentos amplificados y renaturalizados previamente para provocar la formación de heterodúplex. Mediante esta aplicación se puede determinar la T_m (*Temperature melting*) de los amplicones para comprobar si hay presencia de mutación o no. La complementariedad entre dos cadenas de un individuo silvestre será absoluta y, por lo tanto, la estabilidad será alta, por lo que su temperatura de disociación será mayor. Sin embargo, si las cadenas renaturalizadas

poseen un heterodúplex, su estabilidad será menor y su temperatura de disociación será menor también (ejemplo en la Figura 1.6). Esta técnica ya ha sido utilizada en trigo (Dong y col., 2009), tomate (Gady y col., 2009) y *Arabidopsis* (Bush y Krysan, 2010) y puesta a punto recientemente en nuestro laboratorio (Bergareche, comunicación personal).

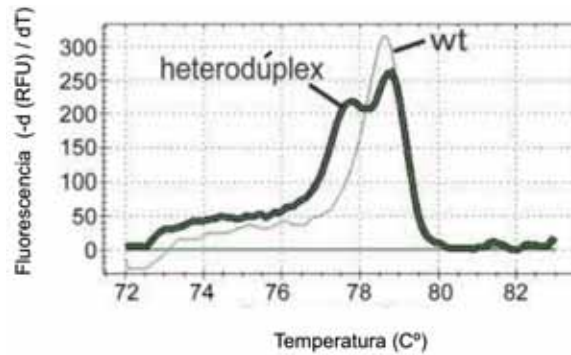


Figura 1.6: Ejemplo de curva de disociación de fragmentos amplificados y renaturalizados. Se puede determinar la presencia de un heterodúplex porque aparece un pico a una temperatura menor (77,6°C) a la determinada para la mezcla silvestre (79,5°C). Las unidades de señal de fluorescencia se miden en función de la velocidad de cambio de las unidades de fluorescencia relativa (RFU) con el tiempo (T) $(-d(RFU) / dT)$ (fuente: Bush y Krysan, 2010).

A pesar de la eficiencia probada de los métodos mencionados anteriormente, la tendencia actual es de usar otros dos métodos no menos eficientes: MALDI-TOF y NGS. La técnica de MALDI-TOF parece estar dando muy buenos resultados en especies de plantas como *Avena sativa* (Chawade y col., 2010). En cuanto a las nuevas tecnologías de secuenciación, las más usadas en TILLING son 454 *Genome Sequencer FLX Ti* (Roche Applied Science) e *Illumina* (Solexa) *Genome Analyzer*. En tomate se han utilizado simultáneamente más de 12.000 muestras para analizar distintos genes en una sola placa 454 PicoTiter Plate (PTP) (Rigola y col., 2009). La última innovación en este sentido ha sido utilizar las tecnologías de secuenciación masiva combinadas con la elaboración de mezclas en distintas dimensiones (Tsai y col., 2011). Para confirmar una mutación mediante secuenciación hace falta una cobertura del genoma de 20-30X en *Arabidopsis* (Ossowsky y col., 2010). En arroz (*Oryza sativa*) y en trigo (*Triticum durum*) se han detectado eficientemente mutaciones mediante la secuenciación de fragmentos de PCR con *Illumina* en un total de 768 individuos combinándolos en mezclas de hasta 94 individuos (Tsai y col., 2011).

El método TILLING combinado con la secuenciación masiva y la aplicación de mezclas en distintas dimensiones parece presentar varias ventajas. En primer lugar la aplicación de varias mezclas permite la confirmación del mutante y disminuye la probabilidad de error. Además, con la identificación de la familia mutante se obtiene también la información del tipo de cambio nucleotídico y de su posición, así como de su efecto en la secuencia aminoacídica. Otra ventaja es que se disminuye la necesidad de uso de cebadores marcados con fluoróforos ni tampoco endonucleasas. Finalmente, es un método escalable a distintos niveles pudiéndose aplicar para tantos genes como sea necesario obteniendo una gran precisión en la detección de mutaciones y en un período de tiempo muy reducido (Tsai y col., 2011).

1.4.2.4 CARACTERIZACIÓN DE LOS MUTANTES

La tasa de mutación obtenida ha sido variable en las distintas plataformas de TILLING desarrolladas hasta ahora (Tabla 1.3), en *Arabidopsis* ha sido identificada una mutación cada 300 Kb (Greene y col., 2003), en maíz una cada 500 kb (Till y col., 2004,a) y en arroz una cada 294 Kb (Till y col., 2007), mientras que en cebada la tasa de mutación es de 1 cada Mb (Caldwell y col., 2004). Actualmente la densidad de mutación más alta ha sido descrita en trigo tetraploide (1/40 kb) y trigo hexaploide (1/24 kb) (Slade y col., 2005). En cuanto al melón se ha desarrollado una población con una tasa de mutación de 1/573 Kb (Dahmani-Mardas y col., 2010).

El método TILLING permite la identificación de las familias mutantes. Una vez identificadas las familias hace falta secuenciar el fragmento donde se encuentra la mutación para poder predecir sus efectos. Debemos analizar si se trata de una mutación en región intrónica o exónica. En caso de localizarse en un exón, examinaremos si el cambio de nucleótido produce también un cambio de aminoácido o si, por el contrario, es silenciosa. En caso de provocar un cambio de aminoácido también deberemos comprobar si el cambio provocado puede afectar a la estructura tridimensional de la proteína. El caso más extremo es aquél en el que el cambio de aminoácido provoca un codón "stop" a la proteína produciendo una interrupción de la traducción. Por lo tanto, las mutaciones identificadas en región codificante de la secuencia nucleotídica pueden clasificarse en:

- mutaciones sin sentido: en las cuales el cambio nucleotídico provoca una conversión de un aminoácido a un codón stop, interrumpiendo así la correcta traducción de la proteína.

- mutaciones sustitutivas o no sinónimas: en las cuales el cambio nucleotídico asociado a la mutación provoca un cambio de aminoácido. Este cambio queda también condicionado por las características químicas del nuevo aminoácido, que puede ocasionar cambios conformacionales y/o funcionales de la proteína.
- mutaciones silenciosas o sinónimas: en las cuales la mutación nucleotídica no provoca un cambio en el codón y por lo tanto no hay un cambio aminoacídico y la proteína mutante será idéntica a la silvestre.

Para una tecnología basada en la genética inversa como el TILLING, las mutaciones sin sentido y sustitutivas son las consideradas más importantes ya que pueden afectar potencialmente a la funcionalidad de la proteína y por lo tanto, dar lugar a un nuevo fenotipo. Durante la construcción de la población resulta esencial conocer la probabilidad de que esto pueda suceder en la región analizada (McCallum y col., 2000b). Es importante analizar una región del gen en la cual podamos encontrar potencialmente más mutaciones tipo sin sentido y sustitutivas. Para escoger una región óptima se dispone de herramientas como el programa CODDLE (*Codon Optimized to Detect Deleterious Lesions*), que permite calcular la probabilidad de obtener mutaciones puntuales en una determinada región y que afecten a la estructura de la proteína (<http://www.proweb.org/coddle>) (Till y col., 2003). CODDLE es un programa que muestra todas las mutaciones que podrían ocurrir en una secuencia determinada en función del tipo de mutágeno utilizado y al mismo tiempo, analiza en cada posición la probabilidad de afectar a la funcionalidad de la proteína en caso de cambio.

Una vez identificada una mutación sin sentido puede estimarse cómo afectará a la funcionalidad de la proteína. Para poder realizar una predicción del efecto de la mutación existen programas informáticos específicos basados en la homología con secuencias proteicas de otras especies. Un ejemplo es el PARSESNP (*Project Aligned Related Sequences and Evaluate SNPs*) (<http://www.proweb.org/parsesnp/>; Taylor y Greene, 2003; Till y col., 2003) y el programa SIFT (*Sorting Intolerant From Tolerant*) que efectúan la predicción en base a un PSI-BLAST (*Position-Specific Iterative Basic Local Alignment Search Tool*) y en base a las características físico-químicas de los aminoácidos (<http://blocks.fhcrc.org/sift/SIFT.html>; Ng y Henikoff, 2001).

El TILLING por lo tanto, proporciona:

- Mutaciones individuales que pueden producir series alélicas que provoquen nuevos fenotipos. Esta gradación fenotípica ligada a una serie alélica puede aproximarnos a conocer la funcionalidad del gen.
- Permite crear y caracterizar una nueva variedad genética en el germoplasma de especies cultivadas de interés.

En el caso de los mutantes identificados mediante TILLING es necesario utilizar retrocruzamientos para obtener la mutación aislada en un fondo genético deseado. Debe tenerse en cuenta que los mutantes detectados mediante TILLING llevan incorporadas no sólo la mutación detectada en el gen de interés sino también muchas otras mutaciones distribuidas al azar en el resto del genoma (Slade y col., 2005). Para poder identificar con mayor exactitud la funcionalidad de un gen no basta con encontrar el mutante y observar el fenotipo sino que además, es necesario obtener el mutante aislado de otras posibles mutaciones adicionales. Sólo de este modo se podría confirmar con certeza que la mutación identificada es la responsable de cierto fenotipo. Por lo tanto en la fase final del protocolo de TILLING se deben efectuar varios retrocruzamientos, entre el mutante y el silvestre de la misma variedad, para poder aislar la mutación de interés (Slade y col., 2005).

1.4.2.5 PLATAFORMAS DE TILLING EXISTENTES

Existen plataformas de TILLING en numerosas especies vegetales y animales. La gran ventaja del TILLING es que puede utilizarse en especies diploides o poliploides, con diferentes tamaños de genomas y que posean cualquier tipo de reproducción (Barkley y Wang, 2008). El único requerimiento es poder mutar las semillas o el polen para posteriormente crecer las generaciones siguientes y extraer el ADN para analizarlo. En la Tabla 1.3 se muestra un resumen de algunas de las poblaciones de TILLING desarrolladas hasta el momento demostrando que esta técnica es aplicable en diferentes especies. Cabe destacar también la variabilidad del tamaño de los genomas en los cuales se ha aplicado la técnica. Los tamaños en organismos diploides abarcan desde las 100 Mb en *C. elegans* (Gilchrist y col., 2006) hasta las 5,300 Mb de *Hordeum vulgare* (Caldwell y col., 2004). La técnica también ha sido eficiente en organismos poliploides como *Triticum turgidum* subsp. *Durum*, tetraploide (Slade y col., 2005) y *Triticum aestivum*, hexaploide (Slade y col., 2005), en este caso los tamaños de genomas abarcan hasta las 17.000 Mb de *Triticum aestivum*. La tasa de mutación es uno de los valores más importantes para determinar si una población de TILLING es eficiente. Cuánto mayor sea la probabilidad de encontrar

mutantes en una plataforma de TILLING más efectiva será. Las tasas de mutación en poblaciones de TILLING de plantas abarcan un amplio rango. En diploides desde 1/60 kb en *Brassica rapa* (Stephenson y col., 2010) hasta 0,5/1 Mb en *Oryza sativa* (Wu y col., 2005) mientras que en especies poliploides encontramos tasas desde 1/24 Kb en *Triticum aestivum* (Slade y col., 2005) hasta 1/967 Kb en *Arachis hypogaea* (Knoll y col., 2011). Debe tenerse en cuenta que en el procedimiento de cálculo no siempre se han usado las mismas variables y por lo tanto, resulta difícil hacer comparaciones entre especies en términos generales. Para algunas poblaciones de TILLING se han desarrollado estudios fenotípicos de los mutantes en distintos tejidos y diferentes fases de crecimiento. Las bases de datos de estas poblaciones permiten observar fenotipos mutantes en el desarrollo de la planta, del fruto, de la flor o de los cotiledones, entre otros. El URGV (*Unité de Recherche en Génomique Végétale*) dispone de una base de datos (UTILLdb) que permite solicitar semillas de mutantes de distintas especies como tomate, colza, guisante o melón (<http://www.versailles.inra.fr/urgv/tilling.htm>). También existen otras como la del centro *Rothamsted Research Institute* que proporciona mutantes de una población de TILLING de trigo. Para *Arabidopsis* existe un servicio web (<http://tilling.fhcrc.org>) donde se puede encontrar toda la información acerca de los mutantes ya caracterizados y un servicio de identificación para genes de interés. También existen algunas empresas privadas que han utilizado el TILLING como método para desarrollar productos comercializables. Este es el caso de *Arcadia Biosciences* (<http://www.arcadiabio.com/>), una empresa norteamericana que ha aislado mediante TILLING un tomate mutante de fruto de larga vida, capaz de conservarse en óptimas condiciones durante muchos más días que un tomate normal, y un tipo de trigo apto para celíacos (<http://www.arcadiabio.com/celiacsafewheat>). Finalmente mencionar la base de datos de soja (<http://www.soybeantilling.org/index.jsp>) realizada por el departamento de genómica de plantas de la Universidad de Illinois (*Southern Illinois University at Carbondale*) la cual es una plataforma de TILLING eficiente por sus altas tasas de mutación y por ofrecer la posibilidad de buscar mutantes en función de un fenotipo determinado o de un gen de interés (Cooper y col., 2008). En algunas poblaciones se han podido identificar mutaciones que han proporcionado un nuevo fenotipo. Por ejemplo, en tomate la mutación G1485A en el gen *eIF4E1* provoca un cambio en el *splicing* del gen proporcionando un ARNm incompleto de modo que la proteína no puede sintetizarse correctamente. Esta

mutación confiere resistencia a dos potivirus: al virus Y de la patata (*Potato Virus Y*, PVY) y al virus del moteado del pimiento (*Pepper mottle virus*, PeMV) (Piron, 2010).

En melón la técnica no solo ha podido proporcionar mutantes fenotípicos (Dahmani-Mardas, 2010; Gonzalez-To y col., 2011) sino que además, se ha podido utilizar para confirmar la función de algunos genes. Un ejemplo es la confirmación de la función del gen CMWIP1 (Martin y col., 2009). Se seleccionaron tres individuos mutantes para este gen dentro de la población de TILLING. Las mutaciones (Ser306Phe, Gly242Arg y Phe193Leu) en homocigosis producían la feminización de las flores en toda la planta y por lo tanto, la conversión de plantas monoicas a ginoicas (Martin y col., 2009).

Durante los últimos años han aparecido estudios metodológicos que han permitido perfeccionar la técnica y optimizarla. Una de estas nuevas metodologías es la llamada NEATILL (*Nucleic acid Extraction from Arrayed Tissue for TILLING*), que ha optimizado el modo de preparación de las mezclas de ADN. En la técnica clásica se mezclaban los ADNs extraídos a partir de la generación M2. En cambio, esta nueva técnica mezcla directamente los cotiledones de plantas M2. Además, la preparación de las placas de mezcla es distinta. Se denomina "pool en 2D" y consiste en realizar dos tipos de placas de mezcla: unas basadas en la mezclas de los ADNs de las filas de las placas de la población y las otras con las mezclas de los ADNs de las columnas de las placas de la población. La ventaja es que una vez detectado un mutante no hace falta hacer la deconvolución ya que con esta tecnología se confirma directamente la familia mutada en cuestión (Sreelakshmi y col., 2009).

Otra reciente publicación optimiza todavía más el protocolo clásico: iTILLING (Bush y Krysan, 2010). En este método el proceso para obtener la generación M2 es distinto. Después de autopolinizar la generación M1 y crecerla, se recogen las semillas M2. Sin embargo, en este caso se mezclan las semillas y posteriormente se hacen germinar directamente en placas *Ice-Cap* (Krysan, 2004; Clark y Krysan, 2007). Gracias a este tipo de placas, una vez germinadas las plantas y aparecidas las primeras hojas se puede realizar la extracción directamente en placa. La detección de mutantes se hace mediante la tecnología de análisis de la curva de disociación (HMR). Estos cambios permiten reducir el tiempo de espera para realizar la detección de mutantes y, a la vez, permiten seleccionar las plantas mutantes antes de ser trasladadas al invernadero con lo que se reduce considerablemente el coste del experimento (Bush y Krysan, 2010).

1.5 GENES CANDIDATOS PARA TILLING USADOS EN EL MARCO DE ESTA TESIS DOCTORAL

La elección del gen diana para el análisis y por lo tanto, la detección de mutantes en una plataforma de TILLING depende del objetivo de la investigación que se esté realizando. Para este trabajo la elección de los genes ha sido, por un lado, comprobar la eficiencia de la plataforma utilizando el gen Fitoeno desaturasa (*PDS*) y, por otro lado, poder detectar mutantes en genes de interés agronómico como es el caso de *ETR1* relacionado con la maduración del fruto y de los factores de iniciación de la traducción 4E: *eIF4E* y su isoforma *eIF(iso)4E* involucrados en resistencias recesivas a virus. En el siguiente apartado se pretende introducir estos genes así como dar la información molecular y funcional que se tiene de ellos y poder entender su elección para el análisis mediante TILLING en esta tesis doctoral.

1.5.1 *PDS* (Fitoeno Desaturasa)

La *PDS* es una enzima clave para la producción de carotenoides (Giuliano y col., 1993; Li y col., 1996; Isaacson y col., 2002). Los carotenoides son pigmentos orgánicos que se encuentran en plantas y otros organismos fotosintéticos. El color de estos pigmentos puede variar desde el amarillo hasta el rojo pasando por una gran variedad de naranjas. El color es provocado por su estructura molecular compuesta por varios dobles enlaces de carbono que producen el proceso denominado conjugación (Giuliano, 1993). Durante este proceso aumentan los dobles enlaces y, por lo tanto, también aumenta la longitud de onda a la que percibe la luz la molécula, dando cada vez un color más rojizo. Los carotenoides tienen varias funciones dependiendo del orgánulo dónde estén acumulados (Linden y col., 1991). En el cloroplasto, los carotenoides absorben la luz en el espectro azul (entre 400 y 600 nm) y la energía que reciben la transfieren a la clorofila (Li y col., 1996). Asimismo, los carotenoides son protectores de la oxidación en los procesos fotoreactivos en organismos fotosintéticos (Giuliano, 1993; Bartley y col., 1999). En los cromoplastos, donde la clorofila no está presente, se acumulan en tejidos como los pétalos de las flores o los frutos en estadios avanzados desarrollando una función básicamente estructural y de atracción para la polinización (Grotewold, 2006). La investigación con cloroplastos de plantas superiores ha ayudado a conocer con más detalle la localización de la *PDS*. En *Nicotiana tabacum* se pudo localizar la *PDS* en un

75% en las membranas de tilacoides. Un porcentaje mucho menor fue detectado tanto en la membrana externa (5%) como en el estroma (16%) (Linden y col., 1991; Al-Babili y Hartung, 1999). La regulación de esta enzima ha sido investigada en distintas plantas como el narciso (Al-Babili y Hartung, 1999) y el tomate (Pecker y col., 1992; Giuliano y col., 1993). La concentración de carotenoides así como la expresión de la *PDS* y la Fitoeno Sintasa (*PSY*) aumentan en condiciones de intensidad de luz en plántulas, sin embargo, el aumento en plantas adultas no parece significativo. En la raíz, estas concentraciones disminuyen considerablemente mientras que en flores aumentan en pétalos y anteras y disminuyen en sépalos y carpelos. Durante el desarrollo de la flor la expresión de ambas enzimas aumenta diez veces inmediatamente antes de la anthesis (Giuliano y col., 1993).

La biosíntesis de carotenos (Figura 1.7) se inicia con tres condensaciones realizadas por la Geranylgeranyl-Difosfato sintasa obteniendo Geranylgeranyl-Difosfato (GGDP). Posteriormente la enzima Fitoeno sintasa (*PSY*) cataliza la reacción de GGDP a fitoeno. Luego, una serie de desaturaciones convierten el fitoeno en licopeno. Los productos intermedios tienen un rango de colores característico de los carotenos, del rojo del licopeno hasta el amarillo de la zeaxantina (Giuliano y col., 1993) (Figura 1.7).

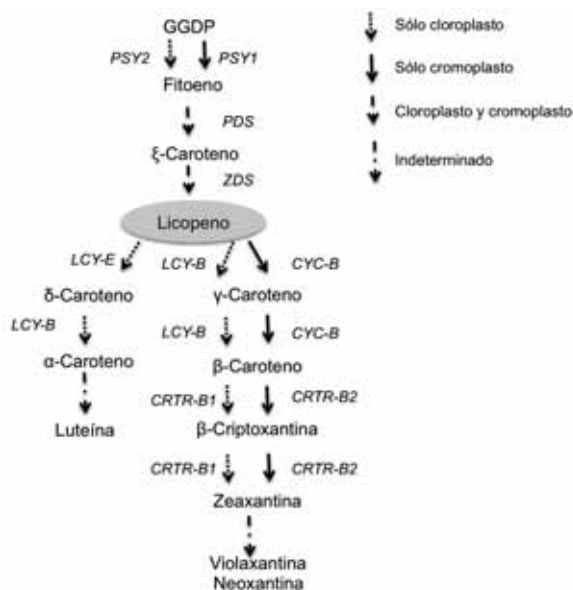


Figura 1.7: Ruta de síntesis de carotenos en plantas. A partir del Geranylgeranyl-Difosfato (GGDP) hasta la obtención de luteína y zeaxantina, pasando por la formación de licopeno (figura traducida y adaptada a partir

de la original de Rachel L. Skelton (Skelton y col., 2006)).

La conversión de fitoeno a licopeno viene mediada en gran proporción por la *PDS*. Si esta enzima no funciona correctamente o no está presente, la síntesis de carotenos se detiene por completo (Hirschberg, 2001). La ausencia de estos productos hace que el tejido no tenga pigmentación alguna dando como resultado una planta albina. La vía de carotenos prosigue en la formación de xantofilas. Éstas son las precursoras de la hormona de crecimiento y desarrollo de la planta, el ácido abscísico. Por consiguiente, la interrupción de la expresión de la *PDS* también puede proporcionar fenotipos enanos (Qin y col., 2007).

En resumen, un mutante total o parcial en este gen puede provocar fenotipos fácilmente detectables sin necesidad de hacer una caracterización o ensayos exhaustivos a nivel metabólico o bioquímico. La visualización de estos fenotipos resulta un control preciso y rápido para experimentos como por ejemplo de transformación o de silenciamiento de genes (Zheng, 2010). En algunos estudios previos con mutantes de la *PDS* se habían obtenido plantas de *Arabidopsis* albinas y enanas (Qin y col., 2007) mediante el silenciamiento o la mutación dirigida por T-DNA. Recientemente, en tomate se ha silenciado la *PDS* mediante VIGS observando también una disminución de la pigmentación en hojas mientras que el silenciamiento de *PDS* no ha afectado a la pigmentación en el fruto (Quadrana y col., 2011).

Por este motivo la *PDS* resulta un buen candidato para la comprobación de la eficiencia de mutación de una población de TILLING. Gracias a la secuenciación del genoma del melón (García-Mas y col., 2012) se ha determinado que esta especie sólo presenta una copia del gen de la *PDS*, dato relevante para el presente estudio de TILLING.

1.5.2 *eIF4E* (FACTOR DE INICIACIÓN DE LA TRADUCCIÓN 4E) y *eIF(iso)4E*

Las plantas están en contacto constante con otros organismos exponiéndose a infecciones: ataques de insectos, hongos, bacterias y virus, entre otros patógenos. La interacción entre una planta y un microorganismo puede ser de varios tipos (Keen, 1990), desde relaciones parasitarias (Greenberg y Yao, 2004) hasta simbióticas (López-Ráez y col., 2010; Fernández-Aparicio, 2010). Como resultado de esta constante interacción, las plantas han ido desarrollado mecanismos de defensa constitutivos e inducibles (Keen, 1990). Como ejemplo de la defensa constitutiva, podemos citar las características físico-químicas de la pared celular (Denny, 1995). Sin embargo, existe otro mecanismo de defensa más complejo e inducible basado en la síntesis de compuestos químicos (Denny, 1995). Para la descripción de este tipo de defensa se ha

sugerido la existencia de un mecanismo por parte de la planta basado en receptores capaces de reconocer la infección y activar el sistema de defensa.

H. Flor (Flor, 1971) mediante estudios entre la interacción del lino cultivado (*Linum usitatissimum*) y el hongo *Melampsora lini* desarrolló la teoría del gen-a-gen (Figura 1.8). En esta teoría se evidenciaba la existencia de genes en la planta que confieren resistencia a un patógeno específico. Actualmente, se sabe que las resistencias en plantas pueden estar generadas por un sólo gen (resistencia monogénica) o por el efecto de varios genes (resistencia poligénica). Posteriormente se ha propuesto que los inductores de la infección son genes llamados *Avr* (factores de avirulencia) presentes en el patógeno, dándole a la hipótesis de Flor un fundamento genético (Parker y Coleman, 1997). Los genes *Avr* del patógeno serían capaces de inducir una respuesta de defensa en aquellos huéspedes que contengan los genes de resistencia específicos (*R*) y que reconocen al gen *Avr* (Parker y Coleman, 1997). Si esta interacción no se puede dar por alteración de uno de los dos genes (*R* o *Avr*) entonces la planta resulta infectada. Aunque el modelo fue desarrollado para la interacción hongo-planta, también ha sido válido para el entendimiento de otras interacciones frente a otros patógenos como bacterias, virus, nemátodos o insectos (Keen, 1990)

Planta Patógeno	Rr/RR	rr
AA/Aa	Resistente	Susceptible
aa	Susceptible	Susceptible

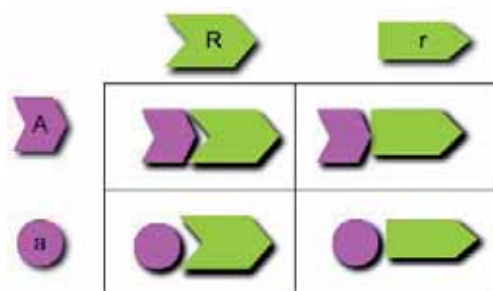


Figura 1.8: Representación del sistema de resistencia en plantas basado en la interacción gen-a-gen (Flor, 1971). En la tabla superior se representa la interacción en base genética. Para simplificar, suponemos que la

planta contiene el gen de resistencia *R* mientras que, por el contrario, el genotipo *r* significa que la planta no tiene la resistencia. De igual modo el patógeno es portador de la avirulencia (*A*) mientras que el genotipo *a* no es capaz de interactuar con *R*.

Actualmente se utiliza el término virulento para aquellos patógenos que son capaces de infectar al huésped mientras que los que no tienen éxito se denominan avirulentos. Con estas bases se describió un modelo donde se podían dar infecciones compatibles e incompatibles (Van del Plank y col., 1965). Una infección compatible consiste en la interacción entre un patógeno virulento y un huésped susceptible. Una infección incompatible se da entre un patógeno avirulento y un huésped resistente, de modo que si el huésped contiene el gen de resistencia capaz de reconocer el factor de avirulencia del patógeno, desencadena un proceso de señalización para la activación de la respuesta de defensa. A esta respuesta se la denomina respuesta de hipersensibilidad (HR) (Bent, 1996). Sin embargo, si la planta no contiene el complejo *R* capaz de reconocer el factor *Avr* entonces el patógeno infecta la planta (Martin y col., 2003), y podemos considerar ese factor de avirulencia como un factor virulento.

Por lo tanto los genes *R* desempeñan un papel fundamental en el reconocimiento del gen *Avr* y en consecuencia en la resistencia de la planta frente a patógenos. La rápida evolución de las plantas en este sentido ha llevado a una gran diversificación de genes *R* (Chisholm y col., 2006). Éstos se han podido clasificar en función de los dominios estructurales que contienen (Bent, 1996; Van Ooijen, 2007). Se han descrito cinco grupos: la clase CNL que comprende los genes de resistencia que codifican para proteínas con al menos un dominio en hélice super enrollada, un sitio de unión a nucleótidos y un dominio rico en repetición de Leucina (Leu) (CN-NB-LRR). La segunda clase TNL contienen al menos un dominio de receptor tipo Toll y los otros dos dominios iguales a la clase CNL (TIR-NB-LRR). La clase RLP (*receptor-like protein*) tienen un receptor de quinasa serina-treonina y un dominio extracelular compuesto por una región rica en repeticiones de leucina (Ser/Thr-TUR). La clase RLK son aquellos genes con un dominio quinasa y una región extracelular rica en repeticiones de leucina (LRR-Kin). Finalmente, el último grupo incluye todos aquellos genes descritos en que su funcionamiento es a través de distintos mecanismos moleculares como por ejemplo *mlo* y *asc-1* (Buschges, 1997; Brandwagt, 2002). Los genes *R* han sido aislados en varias especies de plantas, muchos de ellos en la familia de las solanáceas, donde se han aislado 33 genes de resistencia (Van Ooijen, 2007; Romer y col., 2007). Sin embargo, también se han clonado muchos otros en otras especies como en *Arabidopsis thaliana* (21 *R*-genes) (Meyers y col.,

1999), *Oryza sativa* (arroz, 4 R-genes) (Song y col., 1995), *Cucumis melo* (1 R-gen) (Joobeur y col., 2004) y *Lactuca sativa* (lechuga, 1 R-gen) (Shen y col., 2002).

Los virus son capaces de desarrollar mecanismos para asegurar su supervivencia y tienen la capacidad de mutar y conseguir evadirse de los sistemas inmunitarios. Acostumbran a constituirse a partir de tres o cuatro proteínas. Esta limitación les obliga, en muchas ocasiones, a utilizar los recursos del propio huésped para poder sobrevivir y hacer efectiva la infección. Fraser (Fraser, 1986) propuso que la pérdida de componentes del huésped necesarios para el ciclo de la vida de los virus podía ser también un tipo de resistencia. A esta resistencia se le llama resistencia recesiva. Esta hipótesis ha sido confirmada, sobretodo, con proteínas relacionadas con la maquinaria de traducción celular (Robaglia y Caranta, 2006).

Algunos sistemas de resistencia recesiva en plantas han sido descubiertos recientemente. Un ejemplo son los factores de iniciación de la traducción y en particular la familia de proteínas *eIF4E* y *eIF4G* (Figura 1.9) (Nieto y col., 2006, Robaglia y Caranta, 2006, Truniger, 2008).

El proceso de traducción de mRNA a proteína consta básicamente de tres fases: iniciación, elongación y terminación (Alberts y col., 2006). La fase de iniciación consiste en la interacción entre las dos subunidades del ribosoma y el codón de inicio del mRNA. Durante la elongación se da la síntesis del péptido en función de la lectura de los codones del mensajero dentro del ribosoma. Finalmente, la terminación implica la interacción del ribosoma con el codón *stop* y la disociación completa del mRNA y el complejo ribosomal.

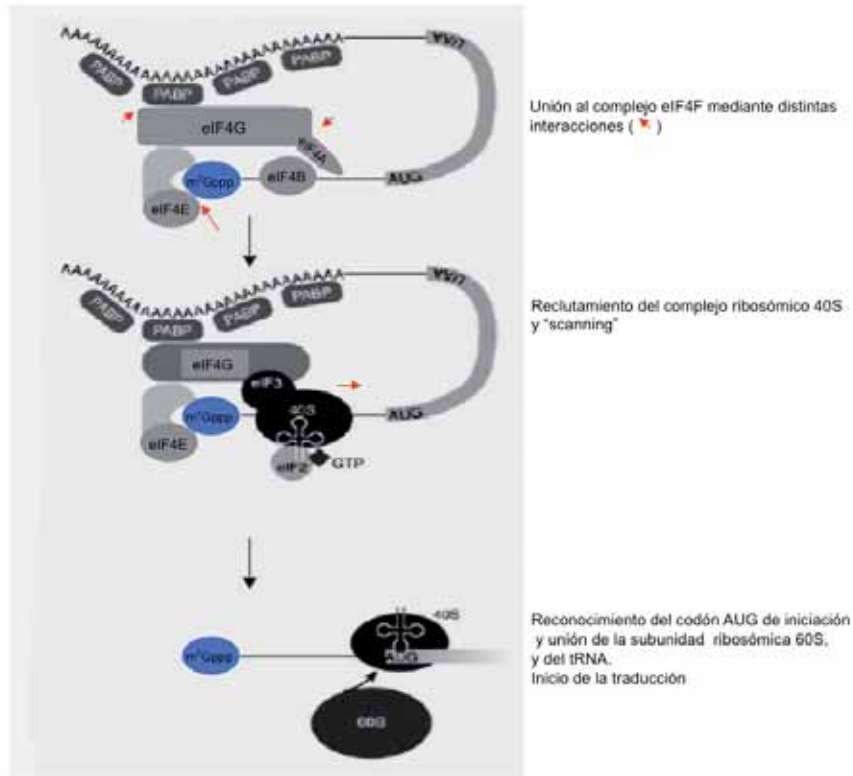


Figura 1.9: Esquema del funcionamiento del inicio de la traducción dónde los factores de iniciación juegan un papel importante en la estabilización del mRNA (fuente: Lellis y col., 2002).

En la fase de iniciación de la traducción contribuyen una serie de factores eucariotas de iniciación de la traducción (*eIF*). Inicialmente la subunidad ribosómica 40s se acerca al 5' del mRNA. Después, un movimiento lateral o *scanning* conduce al reconocimiento del primer codón AUG (el codón iniciador de la traducción). Posteriormente se da la formación de la subunidad ribosómica 60S y la fase de elongación. El factor *eIF4F* une la parte 5' con la parte 3' del mRNA. El complejo *eIF4F* está compuesto por dos subunidades: *eIF4E* y dos proteínas más: *eIF4G* y *eIF4A*. La proteína *eIF4E* se une al extremo 5' del mRNA a través de una proteína de unión PABP (*Poly A Binding Protein*) (Figura 1.9). Por otro lado, *eIF4G* es un complejo de proteínas que tiene la capacidad de interactuar con otras regiones del sistema de iniciación de la traducción como por ejemplo el extremo poli A (a través de la proteína PABP) o con el factor *eIF4A* (helicasa que desenreda estructuras secundarias) y también con el factor *eIF3* (complejo de varias subunidades que interactúa con la subunidad ribosómica 40S) (Lellis y col., 2002).

Las funciones de estos dos factores (*eIF4E* y *eIF4G*) son básicamente interactuar con el mRNA y dirigirlo de modo que el resto de proteínas puedan actuar sobre él. Así por ejemplo *eIF4G* unido a PABP se encarga de interactuar con el extremo 3' para permitir la recircularización sobre si mismo del mRNA mientras que la proteína *eIF4E* interactúa con el extremo 5' (Robaglia y Caranta, 2006).

Más de la mitad de los genes naturales de resistencia a potivirus en plantas se basan en un mecanismo de resistencia recesiva mientras que en otros tipos de virus sólo se encuentra este tipo de resistencia en un 20% (Provvidenti y Hampton, 1992). Los potivirus son virus de cadena simple positiva de mRNA. Están compuestos por una proteína viral (VPg) en el extremo 5', una cadena de poliadeninas en el extremo 3' y otras proteínas virales que después de la infección pasan a ser funcionales mediante las proteasas virales. La proteína VPg es la clave para la interacción con el factor *eIF4E*, o su isoforma *eIF(iso)4E*, tal y como se ha demostrado en experimentos de doble híbrido en levaduras (Wittmann, y col., 1997). Además otros estudios más precisos definen que mutaciones en la proteína VPg impiden la interacción con el factor *eIF(iso)4E* y, por lo tanto, la incapacidad de infectar por parte del virus (Léonard y col., 2000). Se ha demostrado también, que mutantes en el gen *eIF(iso)4E* producidos mediante EMS en *Arabidopsis* pierden la susceptibilidad a TEV (*Tobacco etch virus*) (Lellis y col., 2002).

Se ha podido aislar al gen *eIF4E* como responsable de algunas resistencias naturales en distintas especies. Algunos ejemplos son el gen *mo1* en lechuga que da resistencia al virus LMV (*Lettuce mosaic virus*) (Nicaise y col, 2003), el gen *sbm1* en guisante que da resistencia a PSbMV (*Pea seed-borne mosaic virus*) (Gao y col , 2004), el gen *pot-1* en tomate que da resistencia a PVY (*Potato virus Y*) (Ruffel y col, 2005) y finalmente también en melón el gen *nsv* que da resistencia al *Carmovirus* MNSV (*Melon necrotic spot virus*) (Nieto y col., 2006). Todos ellos son homólogos de *eIF4E* y contienen mutaciones específicas que determinan la resistencia. Todas las mutaciones descritas en estos genes están localizadas en zonas no conservadas del gen. Gracias a la predicción de la estructura 3D se ha podido describir la localización de las mutaciones cerca de la proteína de unión Vgp del mRNA o bien en zonas externas de esta (Robaglia y Caranta, 2006). Nieto y colaboradores en el 2006 desarrollaron un trabajo en melón donde demostraron mediante clonaje posicional y microsintenia con *Arabidopsis* que el locus del gen *nsv* que daba resistencia al virus MNSV correspondía al gen que codificaba por el factor de iniciación de la traducción (*eIF4E*). Complementaron el trabajo con experimentos de expresión confirmando que el gen *nsv* era *eIF4E* y además se identificó la mutación responsable de la resistencia a MNSV que provocaba un cambio en la posición 228 (His-Leu) de la secuencia proteica. Otro estudio de EcoTILLING, (Nieto y

col., 2007) confirmó que la mutación descrita es la que confería resistencia a MNSV en varias de las variedades analizadas. Posteriormente, esta información fue ampliada investigando el mecanismo por el cual el factor *eIF4E* regulaba la resistencia a MNSV. Por un lado se mapó la región en el 3'UTR que determinaba la (a)virulencia mediante virus quiméricos (virulentos y avirulentos) y por otra parte, se identificó el fragmento llamado 3'-CITE (*Cap Independent Translational Enhancer*) en el 3'UTR comprobando que era decisivo para la virulencia del virus. El virus no podía infectar porque no se daba una correcta interacción entre esta parte del RNA viral y el complejo de iniciación de la traducción y por lo tanto, la planta resultaba resistente (Truniger y col., 2008).

En *Arabidopsis* se confirmó que *eIF(iso)4E* y *eIF4E* tienen un 46% de semejanza en la secuencia aminoacídica (Rodríguez y col., 1998). Aunque se ha demostrado que ambas proteínas se encuentran en la misma concentración (Rodríguez y col., 1998) no se sabe exactamente cómo funciona el mecanismo mediante el cual se decide cuál de las dos será la responsable de interaccionar con la proteína VPg del RNA viral. Han sido propuestas tres posibles funciones para *eIF(iso)4E* (Figura 1.10) gracias al estudio en mutantes con pérdida de susceptibilidad en *Arabidopsis* (Lellis y col., 2002).

Respecto al gen *eIF4E* debe comentarse que ya se han encontrado mutantes mediante TILLING en tomate y cebada (Talamé y col., 2008). En tomate se identificó un mutante que perdía la capacidad de *splicing* entre el segundo y tercer exón y la proteína no podía expresarse correctamente. Se caracterizó la inmunidad de este mutante exponiéndolo a distintos virus demostrándose que era resistente a dos de ellos: al virus Y de la patata (*Potato Virus Y*, PVY) y al virus del moteado del pimiento (*Pepper mottle virus*, PeMV) (Piron y col., 2010).

Gracias a la reciente publicación del genoma del melón (García-Mas y col., 2012) se ha podido comprobar que existe sólo una copia del gen *eIF4E* así como de su isoforma *eIF(iso)4E*.

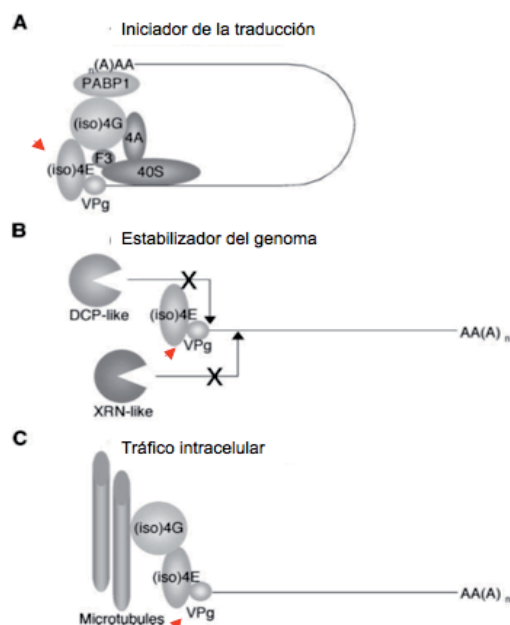


Figura 1.10: Funciones postuladas para el *eIF(iso)4E*. **A)** Iniciador de la traducción por su función de unión a la proteína VPg del mRNA viral. **B)** Estabilizador del genoma. La unión entre *eIF(iso)4E*- VPg puede proteger al mRNA viral de la acción de proteínas “decapadoras” como la DCP-like o exonucleasas que degradan el RNA como la XRN-like. **C)** Tráfico intracelular. *eIF(iso)4E*, interacciona con *eIF(iso)4G* que lleva adheridos microtúbulos con actividad de adhesión. De este modo *eIF(iso)4E* une las proteínas virales con otras proteínas con actividad transportadora (fuente: Lellis y col., 2002).

1.5.3 RECEPTOR DE ETILENO ETR1

El etileno (C₂H₄) es una hormona que entre otras muchas funciones en plantas actúa como activadora de la maduración de los frutos climatéricos. La producción de etileno comporta un cambio en la consistencia del fruto (Brummell y Harpster, 2001), en el color de la piel (Flores y col., 2002) y en la producción de aromas, entre otros efectos (Seymour y McGlasson, 1993). La síntesis de etileno viene regulada por un mecanismo de retroalimentación constituido básicamente por el ciclo de Yang (Figura 1.11) juntamente con la producción de etileno a partir de SAM (S-adenosylmethionine) y de la difusión del etileno a través del fruto (Genard y Gouble, 2005). Pero la regulación del etileno puede ser tanto a nivel biosintético como a nivel receptivo. Se conoce de la acción limitante sobre la producción de etileno de algunas enzimas como la ACC oxidasa y la ACC sintasa (Figura 1.11). Pero la regulación a nivel de recepción es más

desconocida. El hecho de que el etileno sólo actúe en tejidos específicos y que sea una molécula gaseosa nos indica que su recepción es un proceso complejo.

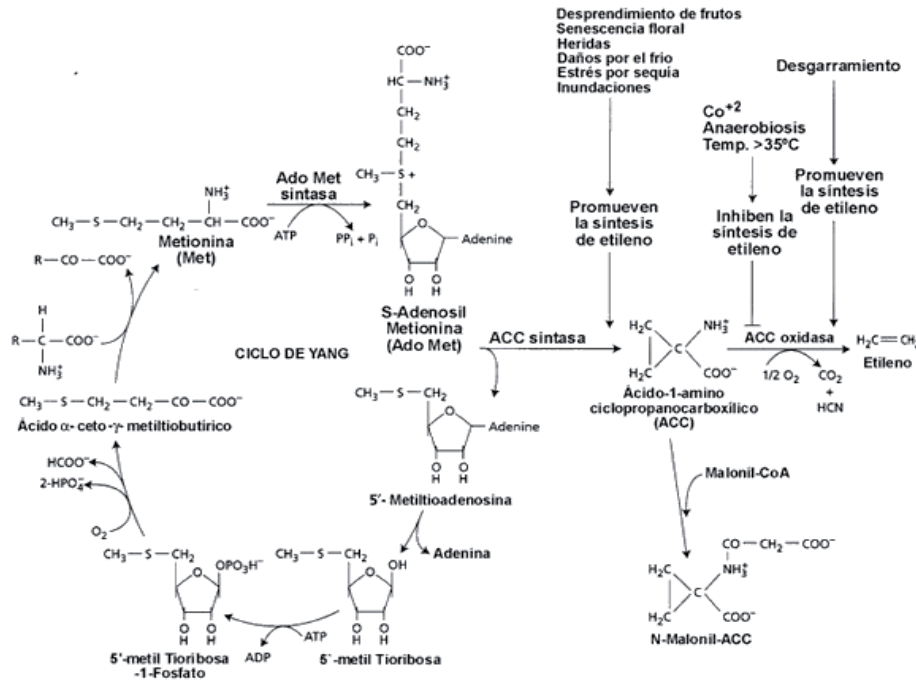


Figura 1.11: Esquema del ciclo de Yang y la vía de SAM- C₂H₄ hasta la formación de etileno. Varios factores externos regulan en parte la síntesis de etileno ya sea a nivel positivo como a nivel negativo (fuente: Soberón y col., 2005)

La triple respuesta al etileno se basa en tres cambios morfológicos durante la germinación en condiciones de oscuridad y presencia de etileno: hipocotilos y raíces más cortas y finas y un crecimiento característico en forma de gancho (Bleecker, 1999; Guzman y Ecker, 1990). Han sido identificados varios mutantes en *Arabidopsis* que no muestran la triple respuesta al etileno y que por lo tanto, presentan insensibilidad a esta hormona. El aislamiento de los genes responsables de estos fenotipos han sido usados para entender el mecanismo de recepción del etileno así como su señal de transducción (Chang y col., 1993; Lehman y col.; 1996). Uno de estos genes, que codifica para un receptor del etileno, es el *ETR1*.

El *ETR1* forma parte de una familia de genes que en *Arabidopsis* está compuesta por cinco receptores: *ETR1*, *ERS1*, *EIN4*, *ETR2*, y *ERS2* (Hua y col., 1995). Todos ellos son receptores del etileno y, a la vez, reguladores negativos de éste. En *Arabidopsis* se ha demostrado que mutantes deficientes en estas proteínas son insensibles al etileno (Wilkinson y col., 1995). Estructuralmente todos son muy similares siendo proteínas transmembranales aunque en el caso de *ETR1*, *ERS1* y *EIN4* contienen además un dominio receptor C-terminal a continuación del dominio Histidina quinasa (Chen y Gallie, 2010).

La proteína *ETR1* contiene una región de alta afinidad y de unión al etileno (Figura 1.12). En *Arabidopsis* es un dímero de 2 proteínas integrales de membrana con actividad histidina quinasa y capacidad autofosforilante. La unión del etileno a su receptor induce su autofosforilación a nivel de residuos de histidina y luego la transferencia de estos fosfatos hacia residuos de aspartato. El receptor así activado inicia una cascada de señalizaciones hacia otras proteínas efectoras (cascada del tipo MAP quinasa con destino final a nivel del ADN) (Soberón y col., 2005) (Figura 1.12).

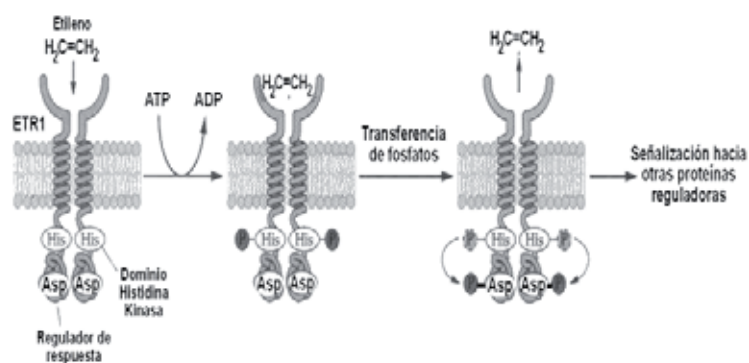


Figura 1.12: Esquema del *ETR1* y del mecanismo de recepción del etileno en *Arabidopsis*. (fuente: Soberón y col., 2005).

Existen mutaciones descritas en *Arabidopsis* en el extremo amino hidrofóbico que reducen o anulan la unión al etileno (Schaller y Bleecker, 1995). Basándose en estas pruebas varios autores han concluido que el gen *ETR1* tiene un cometido importante en la recepción del etileno (Yen y col., 1995, Zhong y col., 2008, Castagna y col., 2007). Asimismo, en *Arabidopsis* está descrita la mutación correspondiente al aminoácido Cys45Tyr de la región N-terminal que provoca un fenotipo con una intensa

insensibilidad al etileno (Bleecker, 1999). Por otro lado, otros estudios (Rick y Bulter, 1956) demostraron que el mutante *Nr* (*Never ripe*) de tomate tenía frutos que no maduraban. Ahora se sabe que este mutante *Nr* es insensible al etileno (Lanahan y col., 1994) y que el locus que provoca el fenotipo *Nr* tiene homología con el gen *ETR1* de *Arabidopsis*. La mutación responsable en tomate es un cambio de base que provoca un cambio aminoacídico de prolina a leucina. Estos antecedentes indican que *ETR1* es un buen candidato para limitar la recepción de etileno y en consecuencia, un gen diana interesante para ser analizado dentro de una población de TILLING, especialmente en melón, una especie modelo para estudiar el proceso de maduración puesto que se encuentran variedades climatéricas y no climatéricas. Por ejemplo, las variedades del tipo *Cantalupensis* presentan un pico en la producción de etileno que coincide con la madurez del fruto y con su abscisión (Shellie y Salveit, 1993). Sin embargo, las variedades del tipo *Inodorus* no presentan un pico de etileno y no se produce abscisión del fruto, dando frutos de larga vida (Kendall, 1988). Un aumento de la concentración de etileno en fruto de melón *Charentais* provoca una disminución de la dureza en la piel (Rose y Hadfield, 1998). En otro estudio se ha demostrado que un tratamiento del fruto de melón con 1-metilciclopropeno (1-MCP), un inhibidor de la acción de etileno, provoca una reducción del ablandamiento de la piel demostrando la implicación del etileno en los cambios de la pared celular durante el proceso de maduración en variedades climatéricas (Nishiyama y col., 2007). El amarilleo y el contenido de carotenoides en la piel, la formación de la zona de abscisión peduncular y la respiración climatérica en el melón *Cantaloupe Charentais* son procesos dependientes de etileno (Flores, 2001; Nuñez-Paleniús y col., 2006, 2008). Sin embargo, otros procesos que ocurren durante la maduración como la acumulación de azúcares, la acidez, el peso del fruto, el tamaño, el número de semillas y el contenido de carotenoides en la carne del fruto del melón se han descrito como procesos independientes de la síntesis de etileno (Nuñez-Paleniús y col., 2008).

Los genes ACC sintasa y ACC oxidasa también han sido descritos como reguladores en la expresión de etileno en varias especies y tejidos (Zarembinski y Theologis, 1994; Giovannoni, 2007). Ambos aumentan su actividad coincidiendo con el momento de máxima producción de etileno y maduración del fruto (Bleecker y Kende, 2000). Aun así, el proceso es mucho más complejo y se sabe que en la regulación intervienen otros genes entre ellos, factores de transcripción, ACO, ACS y la familia de receptores de etileno (Ezura y col., 2000).

Finalmente, dado que en *Arabidopsis* la familia de receptores (*ETR1*, *ERS1*, *ETR2*, *EIN4* y *ERS2*) ha sido bien identificada y descrita (Stepanova y Ecker, 2000) y se ha evidenciado

la participación de estos receptores en la maduración de varias frutas climatéricas, se decidió que *ETR1* era un buen candidato para el TILLING. El interés por descubrir el mecanismo molecular responsable de la regulación de la maduración es evidente tanto a nivel comercial como científico. Los mutantes en el gen *ETR1* podrían aportar nuevos conocimientos sobre la recepción del etileno y en consecuencia, sobre la maduración del fruto, aunque como ya se ha mencionado anteriormente, todos los procesos de maduración y el carácter climatérico no son exclusivamente dependientes de esta hormona (Giovannoni , 2001).

OBJETIVOS

Los objetivos propuestos para esta tesis doctoral son los siguientes:

- Desarrollar la primera plataforma de mutantes TILLING para la especie *Cucumis melo* a partir de la variedad de melón tipo Piel de sapo M62-113.
- Poner a punto el método TILLING y determinar los parámetros que optimicen la detección de mutantes y minimicen el número de falsos positivos, utilizando la detección de mutantes mediante digestión de heteroduplex con endonucleasas.
- Obtener las secuencias de los genes seleccionados *PDS*, *eIF4E*, *eIF(iso)4E* y *ETR1* a partir de M62-113. Realizar la búsqueda de mutantes para cada gen en la población de TILLING con el objetivo de encontrar mutantes en cada uno de ellos y determinar, si es posible, los efectos fenotípicos de estas mutaciones así como comprobar su segregación en generaciones posteriores.
- Establecer una comparación entre dos poblaciones de TILLING, la obtenida en este trabajo de tesis doctoral y la del INRA con el objetivo de comprobar los parámetros óptimos para la obtención de una población de TILLING eficiente con una tasa de mutación aceptable.
- Realizar un estudio de EcoTILLING en los genes *eIF4E* y *ETR1* entre M62-113 y 113 accesiones de melón, que representan la variabilidad de la especie, proporcionadas por el COMAV.
- Iniciar la construcción de un catálogo de mutantes fenotípicos en melón, catalogando las mutaciones encontradas en 400 familias M2, con las descripciones de los fenotipos mutantes encontrados.

2. MATERIALES Y MÉTODOS

PRIMERA PARTE: PROTOCOLO TILLING

2.1. CREACIÓN Y DESARROLLO DE UNA POBLACIÓN MUTAGENIZADA

Para la elaboración de la población de TILLING en melón se escogió la línea doble haploide de la variedad de melón tipo “Piel de Sapo” M62-113. Esta línea proviene de la empresa *Semillas Fitó*, la cual participó en el proyecto MELOGEN (GEN2003-20237 del MEC) dónde está enmarcado este trabajo. El agente químico EMS (Etil Metano Sulfonato) fue el mutágeno escogido para elaborar esta población de TILLING dado que provoca mutaciones puntuales, alquilando las guaninas y provocando que enlacen con las timinas y no con su base complementaria, citosina (Sega, 1984).

2.1.1 MUTACIÓN Y GERMINACIÓN DE LAS SEMILLAS (M0)

La mutación de las semillas se hizo en tres experimentos distintos, ya que el número de semillas era muy elevado (17.000 semillas aproximadamente) y tras la mutagénesis, las semillas debían repartirse en dos centros colaboradores donde posteriormente se realizaría el crecimiento de las plantas. En la primera tanda (año 2.005) se mutaron 8.000 semillas aproximadamente a una concentración de 1% EMS y se mandaron a la empresa *Semillas Fitó* para su crecimiento. En la segunda tanda (año 2.005) se mutaron otras 8.000 semillas aproximadamente, pero esta vez con tres concentraciones distintas de EMS (0,5%, 1% y 1,5%) y en esta ocasión, se mandaron al grupo colaborador de Valencia (COMAV-UPV (Universidad Politécnica de Valencia)). Finalmente, en el 2.006 se mutaron aproximadamente 1.000 semillas en una tercera tanda, ya que muchas de las plantas que se crecieron en Valencia durante el año 2005 murieron debido a una incidencia en campo del colapso del melón. En esta última tanda se usó una única dosis de 1.5 % de EMS. El protocolo de mutación de las semillas se organizó en tres días consecutivos. El primer día se hidrataron las semillas. Para ello se pusieron en recipientes de plástico en una proporción de 25 ml de agua cada 100 semillas. Las semillas se dejaron en agitación suave durante 16 horas. El segundo día se añadió en

campana extractora el agente mutágeno EMS. Se aplicaron concentraciones de 0.5%, 1% y 1.5% (Tabla 2.1) en lotes de 200 semillas por recipiente con 50 ml de agua. El tratamiento se aplicó durante 18 horas en la campana extractora con agitación suave.

Tabla 2.1: Datos correspondientes a las dosis de EMS utilizadas para mutar la población, los frutos obtenidos en la generación M1, el número y porcentaje de las familias con semillas viables y no viables en la generación M2. Finalmente el número de familias M2 incluidas en las placas pool que forman la población de TILLING.

Dosis EMS	Nº de frutos M1	Familias con semillas M2 no viables	Familias con semillas M2 con apariencia viable	Familias M2 en pools de DNA
0,5% COMAV	204	18 (8.82%)	186 (91.18%)	177
1% COMAV	209	25 (11.96%)	184 (88.04%)	151
1.5% COMAV	392	55 (14.03%)	337 (85.97%)	269
1% FITÓ	2.335	399 (17.08)	1.936 (82.92%)	1.771
TOTAL	3.140	497	2.643	2.368

El tercer día se recogieron las semillas con precaución en una malla de un poro inferior al tamaño de la semilla cerrándola con una brida. Se aplicaron tres lavados. Para el primero se introdujo la malla con las semillas dentro de un vaso de precipitados de 5 litros y se dejó en agitación durante 4 horas. Para el segundo lavado se repitió el mismo procedimiento pero en esta ocasión durante 1 hora. Finalmente, el último lavado se realizó dejando la malla en una pila bajo agua del grifo durante 1 hora.

Al material de laboratorio usado para los lavados de las semillas se le añadió una solución de 1M de NaOH para inactivar el EMS. El mismo día se prepararon las semillas para la germinación. Se repartieron 50 semillas en cada bandeja de plástico con una base de papel de filtro húmedo y se cubrieron con plástico transparente para mantener la humedad. De éste modo se dejaron toda la noche en cámara para la germinación en condiciones óptimas de luz (12-20 horas luz continua), temperatura (28°C-30°C) y grado de humedad adecuados (hidratación previa de las semillas y papel

húmedo en la base del recipiente). Al día siguiente, las semillas germinadas fueron trasladadas a sus respectivos destinos para ser plantadas.

2.1.2 CRECIMIENTO DE LAS GENERACIONES M1 Y M2

Un total de 17.000 semillas se repartieron en los dos destinos como se ha descrito anteriormente. El crecimiento en Cabrera de Mar en *Semillas Fitó* se hizo en invernadero y en condiciones óptimas para el melón (T° 25-30°C y 55-65% de humedad).

En Valencia se creció la primera tanda del 2005 en campo, pero tuvieron una grave infección por colapso y murieron gran parte de las plantas sobreviviendo una pequeña parte de las 8.000 semillas plantadas inicialmente. Por este motivo posteriormente, se plantaron 1.000 semillas más, tratadas al 1,5% de EMS, esta vez en invernadero para controlar más las condiciones ambientales y sobretodo evitar infecciones.

2.1.2.1 AUTOFECUNDACIÓN DE LA GENERACIÓN M1 EN INVERNADERO

Este trabajo fue realizado por el personal de soporte de cada centro donde fueron distribuidas las semillas (COMAV y *Semillas Fitó*). A medida que las plantas fueron creciendo se iban podando de manera que sólo quedara una rama principal y algunas secundarias dónde una vez aparecidas las flores se polinizaron manualmente. Se autofecundaron un total de 3.140 plantas M1 (Tabla 2.1).

Para la polinización, primero se eligió una flor femenina y se le extrajeron los pétalos con cuidado dejando el gineceo libre y sobretodo la parte superior del ovario (estigma) accesible. Con la ayuda de un clavo afilado se recogió un poco de polen de la antera de una flor masculina de la misma planta, con cuidado e intentando que éste permaneciera lo más seco posible. Se aplicó este polen encima del estigma intentando también que el ovario estuviera lo más seco posible ya que con la humedad el polen no fecunda eficientemente.

Posteriormente, se cubrió la flor con un sobre de papel para evitar polinizaciones cruzadas y se rotuló con la fecha para poder controlar los días de cuajado. También se cortaron las ramas laterales que emergían de la misma rama en la cual se encontraba la flor polinizada. De este modo la planta podía concentrar más recursos en el cuajado del fruto. Después de una semana aproximadamente se retiraba el sobre de papel para confirmar el cuajado. Por lo general se obtuvo un fruto por planta. Los melones se dejaron crecer durante 55 días (DAP) aproximadamente.

2.1.2.2 CRECIMIENTO DE LA GENERACIÓN M2

Un total de 2.368 frutos M1 produjeron semilla M2 viable (Tabla 2.1). A estos frutos se les extrajeron y limpiaron las semillas. Posteriormente se germinaron diez semillas de cada familia y de cada plántula se extrajo un disco de hoja. Los diez discos de hoja se mezclaron para extraer el ADN que posteriormente, se organizó en placas. El resto de semillas de cada familia M2 se guardó para futuras aplicaciones (Figura 2.1).

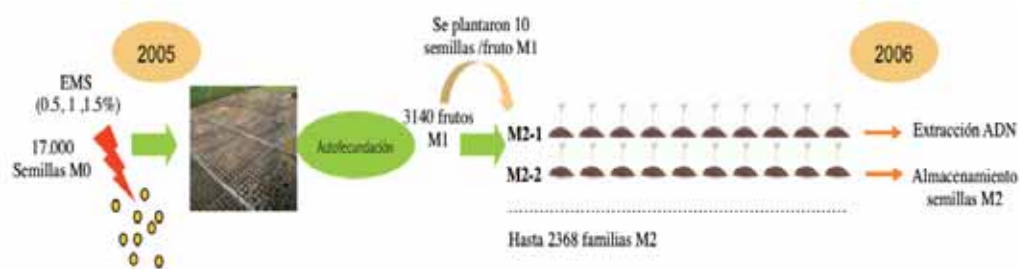


Figura 2.1: Esquema del procedimiento utilizado para obtener el ADN de la población de TILLING. En primer lugar se mutaron aproximadamente 17.000 semillas lo que permitió la autofecundación de 3.140 plantas M1 para la obtención de 2.368 familias M2. De cada familia M2 se germinaron diez individuos y se extrajo el ADN de cada una de las diez plantas para formar el pool de ADN. Este ADN es la base de la plataforma de TILLING.

2.1.3 EXTRACCIÓN DE ADN

2.1.3.1 EXTRACCIÓN DE ADN MEDIANTE KIT COMERCIAL

La extracción del ADN se hizo en el COMAV con el Kit comercial DNeasy 96 plant Mini Kit (250) (QIAGEN) de modo que se pudieron hacer un gran número de extracciones con un alto rendimiento y en un período de tiempo reducido. Una vez extraído el ADN, se mandó a nuestro laboratorio del IRTA en Cabrils donde se verificó la calidad y se cuantificó mediante geles de agarosa. Después se transfirieron los ADNs de tubos a placas de 96 (Figura 2.2), obteniendo finalmente un total de 25 placas stock.

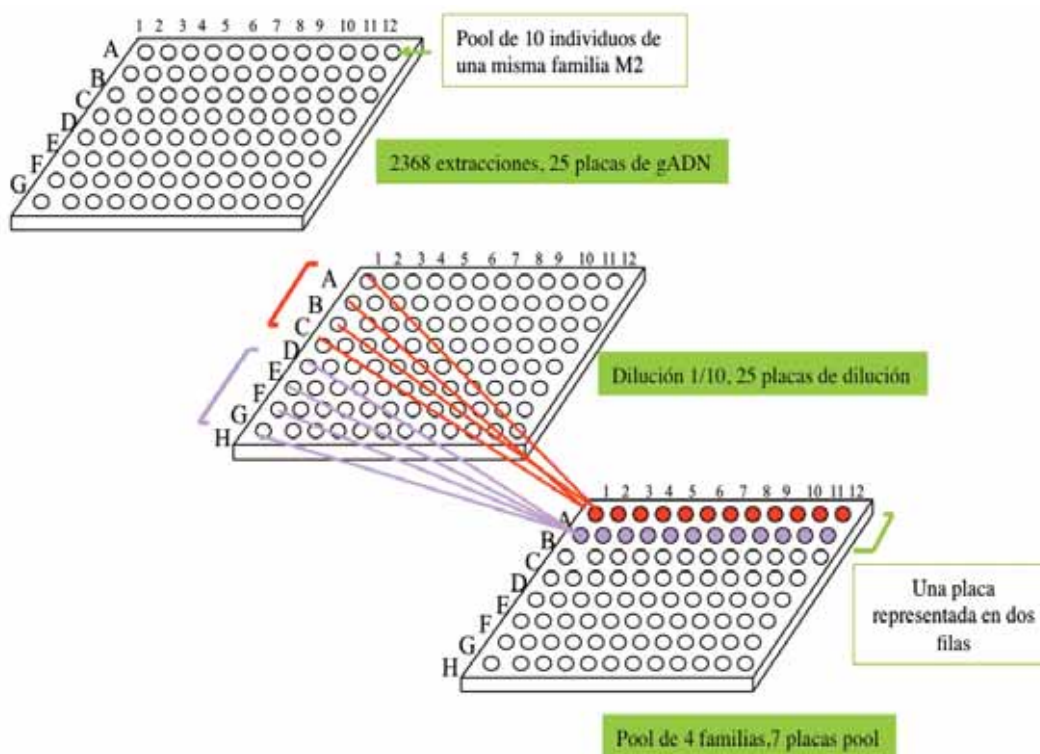


Figura 2.2: Esquema del diseño seguido para elaborar las placas pool. gADN: ADN genómico. Inicialmente se introdujeron todos los ADNs de cada familia M2 en placas *Stock* con una concentración media de 50 ng/ μ l. Cada pocillo contiene el ADN de una familia M2 compuesta por 10 individuos. Posteriormente se construyeron las placas de dilución, diluyendo a 5 ng/ μ l el ADN de cada familia. Finalmente se organizaron las placas pool mezclando las familias de cuatro en cuatro en el orden que muestra la figura.

Se realizaron copias de las placas stock pero esta vez con el ADN diluido 1/10. Se añadieron 20 μ l del ADN stock con 180 μ l de agua para obtener un volumen total de la dilución de cada familia de 200 μ l. Para la construcción de las placas pool se mezclaron equimolarmente los ADNs de cada cuatro familias. El diseño se muestra en la Figura 2.2.

2.1.3.2 MINI EXTRACCIÓN DE ADN

El protocolo usado está basado en el protocolo publicado por Doyle and Doyle (1.990) ligeramente modificado. El material de partida fue un pequeño disco de hoja en estadio de plántula. El disco de hoja se introdujo en un *ependorf* de 1,9 ml el cual se sumergió en nitrógeno líquido lo más rápidamente posible. Posteriormente se retiró el *ependorf* y

se trituró el material con la ayuda de una pequeña broca para devolver la muestra triturada al nitrógeno. Posteriormente, se le añadió 500 µl de tampón de extracción Doyle (Anexo II.f) precalentado a 65°C mezclándolo suavemente para después introducirlo al baño maría a 65°C durante 30 minutos.

Se añadieron 500 µl de cloroformo isoamílico en proporción 24:1; se mezcló y se centrifugó a 11.000 rpm a T^a ambiente durante 10 minutos. De esa centrifugación se recuperó la fase acuosa y se le añadieron 350 µl de 2-propanol (isopropanol) frío centrifugándolo 5 minutos a 10.000 rpm para poder precipitar los ácidos nucleicos. Al ADN precipitado se le añadieron 200 µl de tampón de lavado de Doyle (Anexo II.g) dejándolo reposar un mínimo de 20 minutos a -20°C, después se centrifugó 3 minutos a 10.000 rpm y a T^a ambiente. Finalmente se extrajo el sobrenadante para dejar el pellet secándose durante 30 minutos. Una vez seco se resuspendió el pellet con 50 µl de agua HPLC, y se dejaron las muestras a 4°C toda la noche para facilitar una buena resuspensión del ADN.

2.1.3.3 ELECTROFORESIS EN GEL DE AGAROSA

Para la comprobación de la calidad y cantidad del ADN de las extracciones, se separaron los fragmentos en geles de agarosa. Los geles contenían un 1% de agarosa (EEO de Pronadisa) en NEB 0.25x (Anexo II.i) y 1 µl/100 ml de bromuro de etidio. La electroforesis se realizó con tampón NEB, y se cargaron las muestras en agua junto con el colorante de carga (1/5) (Anexo II.h). Junto con las muestras se cargó también el marcador de peso molecular dependiendo del tamaño del fragmento a detectar (ADN del fago λ o *pUC mix*). Finalmente, se aplicó un voltaje de entre 85-100 V.

2.2 DETECCIÓN E IDENTIFICACIÓN DE LOS MUTANTES

Una vez elegido el gen diana y obtenida su secuencia específica de la variedad M62-113, con la que hemos desarrollado la población de TILLING, se procedió a la detección de los mutantes. El protocolo está basado en cuatro fases: 1) Amplificación del fragmento del gen mediante PCR, 2) Desnaturalización y renaturalización de los amplicones, 3) Digestión de los heterodúplex, 4) Visualización de los fragmentos de digestión en geles de acrilamida verticales (Figura 2.3).

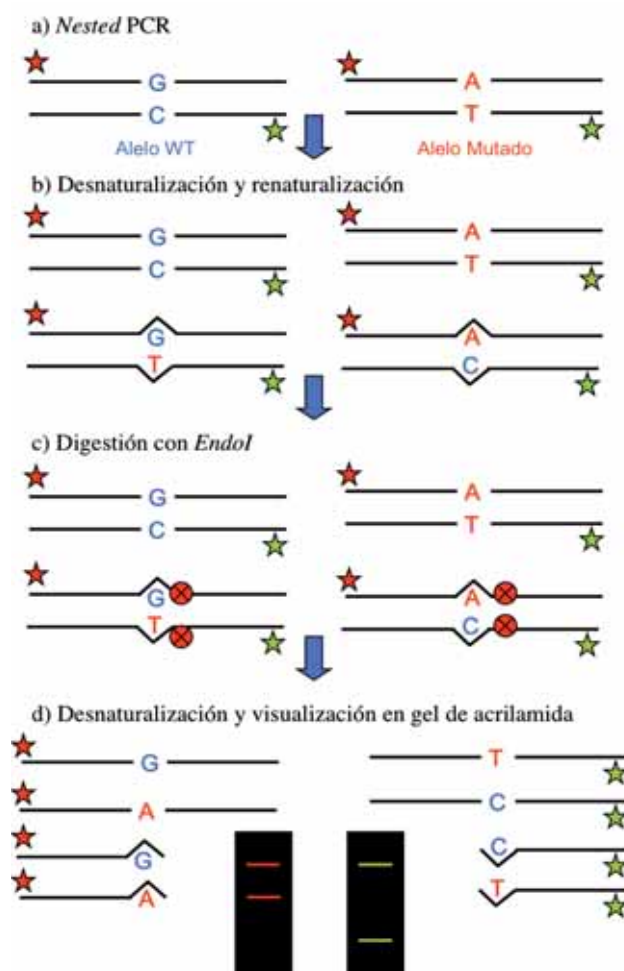


Figura 2.3: Esquema del protocolo de TILLING. a) *Nested PCR*, segunda amplificación (más específica) del fragmento a analizar con cebadores marcados con fluorescencia. b) Desnaturalización de los fragmentos y renaturalización provocando los heterodúplex entre los fragmentos no complementarios por contener una

mutación. c) Digestión con la endonucleasa específica de heterodúplex y por lo tanto de los fragmentos que contienen mutación. Los alelos silvestres no formarán heterodúplex y por lo tanto, no son digeridos. d) Desnaturalización y visualización de los fragmentos en geles de acrilamida. En el LI-COR 4300 podemos visualizar los dos canales (IRD 700 y IRD800) por separado y confirmar las mutaciones mediante el cálculo del tamaño del fragmento digerido. Con la suma de los fragmentos digeridos y visualizados en ambos canales debe obtenerse el fragmento no digerido original.

2.2.1 ELECCIÓN Y SECUENCIACIÓN DEL GEN DIANA

Para aplicar el método de TILLING primero hace falta definir el gen diana que queremos estudiar y del cual queremos encontrar mutantes. Una vez elegido el gen hace falta secuenciarlo en la especie o variedad con la que se haya obtenido la población de TILLING para un eficiente diseño de los cebadores.

Para la amplificación de los genes *PDS*, *eIF(iso)4E* y *ETR1* se buscaron ESTs en la base de datos de MELOGEN (actualmente se puede encontrar esta información en www.icugi.org) mientras que para el gen *eIF4E* los cebadores ya estaban diseñados anteriormente (Nieto y col., 2006). Para cada uno de ellos se encontró un unigen (www.icugi.org) y a partir de ese fragmento se diseñaron los cebadores para empezar a secuenciar el gen entero o regiones concretas del mismo. Para el diseño de cebadores se usó el programa primer3 (<http://frodo.wi.mit.edu>) y en determinadas ocasiones se diseñaron los cebadores manualmente.

Actualmente, y gracias a la disponibilidad de la secuencia del genoma del melón, no sería necesario secuenciar los ESTs puesto que ya dispondríamos de la secuencia específica de cada gen en la especie de estudio.

2.2.1.1 AMPLIFICACIÓN DE FRAGMENTOS MEDIANTE (PCR)

Una vez diseñados los cebadores, se optimizó la amplificación mediante PCR en gradiente de temperatura para determinar la temperatura de hibridación óptima. Esta PCR consiste en preparar la misma mezcla (de ADN y cebadores) y probarla a distintas temperaturas de hibridación para poder observar en cual es más específica.

Para preparar la PCR estándar se utilizaron de 10-50 ng de ADN, tampón Lab 10X (Lab 1X: MgCl₂ 1.5 mM, KCl 50 mM, Tris-HCl 10 mM pH 8.3, gelatina 0.001%), 0.4 mM de cada nucleótido, 0.4 μM de cada cebador, y en función de cada pareja de cebadores, de 1.5 mM a 3.5 mM de MgCl₂ y 2U de Taq polimerasa. El volumen final estándar utilizado era de 25 μl. La reacción se realizaba en termocicladores Gene Amp PCR System 9700,

2720 Thermal Cycler (Applied Biosystems) y MJPTC-200 Peltier Thermal Cycler (MJ Research) para a las reacciones de PCR en gradientes de temperatura.

El programa estandarizado consistía de un primer ciclo de desnaturalización de 2 minutos a 94°C, y 35 ciclos de desnaturalización de la cadena (20 segundos a 94°C), hibridación de los cebadores (temperatura específica para cada pareja), y elongación del fragmento (de 30 segundos a 1 minuto en función del tamaño del amplicón a 72 °C). Finalmente se añadió un ciclo de 5 minutos a 72°C para acabar la elongación y se mantuvo la reacción a 12°C.

2.2.1.2 SECUENCIACIÓN DE ADN

Los fragmentos resultantes fueron primeramente purificados con el *High Pure Product Purification Kit* (Roche) para posteriormente ser secuenciados. En cada reacción de secuenciación se añadieron 3 µl (10-25 ng) del fragmento de ADN purificado que deseaba secuenciarse, 2 µl de tampón TDS (200 mM Tris-HCl, PH9.0; 5 mM MgCl₂), 2 µl de Big Dye (*BigDye® Terminator v1.1 Cycle Sequencing Kit*) y finalmente 3 µl de agua HPLC para completar el volumen total de 10 µl. Para secuenciar el fragmento se hizo una PCR de secuenciación que constaba de 25 ciclos, de tres pasos; un primer paso para la desnaturalización a 96°C durante 10 segundos, un paso de hibridación de los cebadores a 50°C durante 5 segundos y un tercer paso de elongación a 60° durante 4 minutos.

Posteriormente se procedió a la precipitación del DNA en placa. A 10 µl del volumen de la reacción de secuenciación se le añadían 10 µl de agua y posteriormente, 2 µl de EDTA 125 mM, 2 µl de AcNa 3M pH 4,6, y 50 µl de EtOH absoluto. Después de mezclarlo se dejaba precipitar el DNA durante 15 minutos y se centrifugaba a 3.000 rpm durante 30 minutos, se invertía la placa y se eliminaba el sobrenadante. Luego se realizaba una centrifugación corta a 1.000 rpm para después añadirle 150 µl de EtOH al 70 %. Una vez mezclada la placa por inversión se centrifugaba a 3.000 rpm durante 10 minutos eliminando por absorción el sobrenadante. Finalmente se podía hacer una centrifugación corta con la placa invertida encima de papel absorbente para asegurar un secado más intenso del precipitado. Una vez precipitadas las secuencias se resuspendieron cada una de ellas con 12 µl de formamida HiDi y se cargaron en el secuenciador ABI3130XI.

2.2.1.3 ANÁLISIS DE LAS SECUENCIAS

Las secuencias fueron analizadas habitualmente con el programa *Sequencher 4.8*. De cada gen se utilizó el ortólogo de la especie más cercana que hubiera disponible en la base de datos como modelo. De este modo pudimos obtener la secuencia exacta de cada gen o del fragmento de interés en la variedad M62-113 y localizar la posición de intrones y exones.

2.2.2 DISEÑO DE CEBADORES PARA LA DETECCIÓN DE MUTANTES

Una vez obtenida la secuencia específica de los genes en la variedad M62-113, procedimos al diseño de cebadores que amplificarían la región de interés para posteriormente ser analizada mediante el método TILLING. Para la detección de los mutantes se requiere una amplificación muy específica, es por ello que se diseñaron dos parejas de cebadores para cada región. Una amplificaba un fragmento mayor de la región de interés mientras que la otra pareja de cebadores estaba diseñada a una distancia de entre 10 y 20 pb más adelante de la pareja anterior, haciendo así una amplificación de un fragmento previamente ya amplificado o doble amplificación (*Nested PCR*).

Para la elección del fragmento a amplificar se utilizó el programa CODDLE (*Codons Optimized to Discover Deleterious Lesions* (<http://www.proweb.org/input/>)). CODDLE permite identificar las zonas del gen que una vez mutadas tienen más probabilidad de variar el aminoácido en zonas conservadas de la proteína y por lo tanto, más probabilidad de afectar a la funcionalidad de la misma. El programa permite introducir variables como la especie que se muta, y el método y/o agente químico con el que se ha realizado la mutación para poder realizar el cálculo. Una vez identificada esta región se diseñaron los cebadores específicos con CODDLE, primer3 o mediante diseño manual (Anexo II.a).

2.2.3 AMPLIFICACIÓN DEL FRAGMENTO A ANALIZAR

Gracias al diseño de dos parejas de cebadores y a la *Nested PCR* se obtenía un producto mucho más específico. Para la primera PCR (apartado 2.2.1.1) se utilizaron los cebadores sin marcaje. La reacción de PCR se realizó a partir de la mezcla estandarizada mostrada en la Tabla 2.2.

Tabla 2.2: Composición de la reacción de PCR en el protocolo de TILLING (*Anexo II.b.**ADN procedente de cada pocillo de las placas pool)

Componente	Concentración inicial	Concentración final	V/reacción
Tampón de PCR*	10 mM	1 mM	2.5 µl
MgCl ₂	25 mM	2,5 mM	2.5 µl
dNTPs	5 mM	0,2 mM	1 µl
Cebador F	10 mM	0,4 mM	1 µl
Cebador R	10 mM	0,4 mM	1 µl
Taq Polimerasa	5 u/µl	5 u	1 µl
ADN**	5 ng/µl (aprox)	10 ng	2 µl
Volumen total			25 µl

2.2.3.1 NESTED PCR

Para la segunda PCR (*Nested PCR*) (Figura 2.3.a) se utilizó el protocolo anterior aunque en este caso se añadían los cebadores marcados para LI-COR, de modo que los cebadores directos están marcados con el fluorocromo IRD700 (rojo) y los cebadores reversos con el fluorocromo IRD800 (verde). El ADN añadido era el producto de la primera PCR. El programa de PCR utilizado era el mismo que en la primera amplificación aunque cambiando la temperatura de hibridación para cada pareja de cebadores. En estas amplificaciones debía tenerse precaución y se debían envolver los tubos con papel de aluminio en todo momento ya que los fluorocromos IRD se degradan al exponerse a la luz.

Finalmente, la visualización del producto de PCR se realizó en geles de agarosa al 1% y tinción con bromuro de etidio. Se usó marcador de peso molecular adecuado a cada fragmento (*pUC Mix* para fragmentos de 100-1000 pb y λ digerido con *HindIII+HaeIII* para fragmentos de más de 1000 pb) para comprobar que el amplicón obtenido era el correcto.

2.2.4 FORMACIÓN DEL HETERODÚPLEX

Este paso es necesario para la detección de la mutación (Figura 2.3.b). Cada pocillo de las placas pool consta de ADN de 4 familias M2 distintas (cada una de ellas con 10 individuos, por lo tanto en cada pocillo hay 40 individuos representados). El ADN está mutado al azar y en cada pocillo puede haber presencia de alelos silvestres y alelos

mutados. La formación del heterodúplex sólo se dará entre fragmentos con algún cambio de base o SNP, es decir, entre alelos distintos (mutado y no mutado).

Para la formación del heterodúplex se sometieron las PCRs con marcaje (*Nested PCR*) a un ciclo de desnaturalización y luego una renaturalización lenta. Para ello se introdujo la *Nested PCR* en el termociclador MJPTC200 con un programa de 94°C durante 10 minutos y luego un gradiente de 94°C a 8°C disminuyendo -0.1°C/s y finalizando a 4°C.

2.2.5 DIGESTIÓN DEL HETERODÚPLEX

2.2.5.1 ELECCIÓN DE LA CONCENTRACIÓN DE LA ENZIMA *EndoI*

Para elegir correctamente la dilución de la enzima *EndoI* se usó la información proporcionada por el laboratorio del INRA, que fue el centro que nos proporcionó la enzima concentrada. En el Anexo II.c puede verse el gráfico elaborado por este laboratorio dónde se puede determinar la dilución de enzima necesaria en función de la concentración de ADN obtenido en el producto de la PCR.

La concentración media de ADN obtenida en nuestro laboratorio en las condiciones citadas anteriormente era del rango de 40-50 ng, por este motivo y habiendo realizado pruebas test para corroborar la información del gráfico se utilizó la dilución 1/5000 (Anexo II.c) de la enzima *EndoI*.

2.2.5.2 DIGESTIÓN DEL PRODUCTO DE LA PCR CON *EndoI*

Para cada reacción de digestión (Figura 2.3.c) (y teniendo en cuenta que la cantidad de producto de PCR siempre oscilaba entre 40 y 50 ng) se añadió 3 µl del producto de PCR, 3 µl de tampón para la digestión (Anexo II.e), 3 µl de *EndoI* (1/5.000) y 21 µl de agua HPLC para llegar a los 30 µl de volumen total. Esta reacción se incubaba a 45°C durante 20 minutos. Posteriormente, para detener la reacción de digestión, se añadieron 5 µl de EDTA 0.15 M manteniendo la reacción en hielo y protegida de la luz.

2.2.5.3 PURIFICACIÓN DEL PRODUCTO DE LA DIGESTIÓN

Para la purificación primero debía prepararse la placa de purificación llenando la placa de carga con *Sephadex* en polvo (*Sephadex*TM G-50 Superfine, Amersham Biosciences), y expandir con la espátula que lleva adjunta la placa de carga *Sephadex* G50 (MACL09645, Millipore). Una vez bien llenos los pocillos y la placa limpia se invirtió sobre una placa filtro Millipore (MultiScreen-HV) dejando que el *Sephadex* cayera en cada pocillo. Añadimos 325 µl de agua HPLC a cada pocillo de la placa filtro para que el

Sephadex polimerizara y formara la resina que purificaría el producto de la digestión. Una vez añadida el agua hizo falta esperar un mínimo de una hora para que la resina absorbiera suficientemente el agua.

Después de una hora se centrifugó la placa filtro con una placa debajo (Microplate Greiner de 96 pocillos) a 2.000 rpm durante 2 minutos para secar la resina. Una vez expulsada el agua la placa estaba lista para cargar los 35 μ l de la digestión de heterodúplex. Para aislar la muestra se colocó la placa de recolección (Placa ABI), habiendo cargado previamente 5 μ l de azul con formamida en cada pocillo, debajo de la placa filtro (prestando atención a que la orientación de las dos placas sea la misma). Posteriormente se añadieron los 35 μ l de la digestión en cada pocillo y se centrifugó otra vez 2 minutos a 2.000 rpm, eluyendo así la muestra. Para el secado y concentrado de la muestra se depositó la placa en el *Speed Vac* (Jouan RC 1009) a 65°C durante 45 minutos. De este modo se redujo el volumen de agua y se obtuvo la muestra concentrada y a punto para la carga en el gel de acrilamida vertical.

2.2.6 SEPARACIÓN DE LAS MUESTRAS MEDIANTE GELES DE ACRILAMIDA VERTICALES Y DECONVOLUCIÓN

2.2.6.1 PREPARACIÓN DEL GEL DE ACRILAMIDA

Para la separación de las muestras hemos utilizado el secuenciador LI-COR 4300. Para la preparación del gel, primero se montaron adecuadamente los cristales del LI-COR (cristales tratados específicamente para la correcta captación y el análisis del láser de la misma máquina). Entre los dos cristales se colocaron los espaciadores de 25 mm de LI-COR, y se encajaron los cristales con las piezas para este montaje. Una vez montados se preparó la mezcla con 25 ml de Acrilamida (*AccuGel™ 19:1 Sequencing Grade; National Diagnostics*), 20 μ l de TEMED y 100 μ l de APS (persulfato de amonio). La mezcla se agitó ligeramente y se absorbió en una jeringa de 50 ml para inyectarla inmediatamente en el montaje de los cristales. Finalmente se puso el peine en la parte superior del gel para marcar el espacio dónde se cargarían posteriormente las muestras. Una vez polimerizado el gel (entre 30 y 45 minutos) se montó dentro del LI-COR 4300. Todas las piezas se encajaron tal y como muestra la Figura 2.4. Se usó TBE 10X diluido a partir del stock comercial (KB^{PLUS}TBE Buffer Pack, LI-COR) como tampón para correr las muestras.

Las condiciones de precalentamiento del gel usadas en el LI-COR 4300 fueron las estandarizadas para el TILLING que corresponden a 1500 V, 40 mA, 40 W y 50°C

durante 15 minutos. Una vez hecho el precalentamiento del gel se procedió a cargar las muestras.

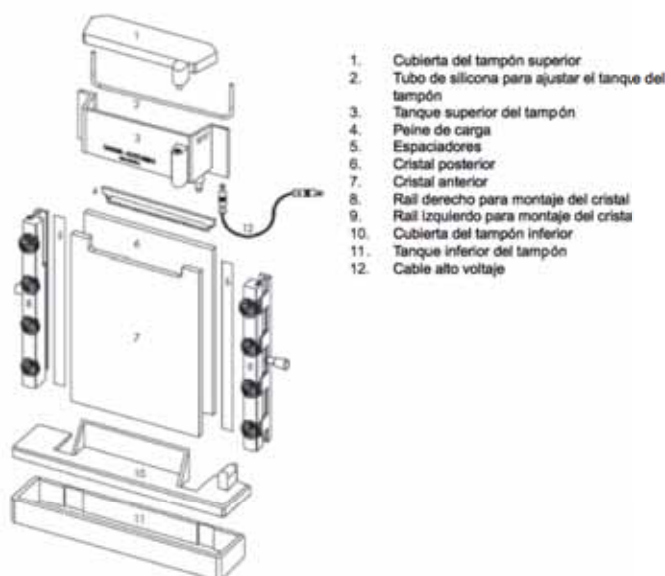


Figura 2.4: Representación del montaje necesario para correr el gel en el LI-COR 4300. El gel se introduce en la máquina LI-COR 4300 y se controla mediante un computador que controla las condiciones de electroforesis (imagen extraída de www.licor.com)

2.2.6.2 CARGA DE LAS MUESTRAS EN EL GEL DE ACRILAMIDA

Las muestras concentradas mediante el *Speed vac* se desnaturalizaron en un termociclador a 94°C durante 3 minutos e inmediatamente se pusieron en hielo. El marcador de 700 pb (LI-COR Size Standard 50-700 bp IRDye® 700 y 800;) para el LI-COR también debía desnaturalizarse. Las muestra se cargaron en una placa de carga (96Well Membrane Tray, #TAY96, de The Gel Company) para así posteriormente absorber las muestras con los peines desechables (CAJ96 de The Gel Company). Después de dos minutos (en los que la muestra se absorbe bien en el peine) se colocó el peine en la parte superior del gel y se dejaron correr las muestras durante 4 minutos con el peine encajado. Finalmente, se detuvo el voltaje para quitar el peine desechable dejando correr las muestras el tiempo necesario en función del tamaño del amplicón (para 1000 pb, 4-6 horas) (Figura 2.3.d).

2.2.6.3 ANÁLISIS DEL GEL DE ACRILAMIDA: DECONVOLUCIÓN

Los mutantes se identificaron finalmente en los geles de acrilamida. Para detectar un mutante hace falta hacer un análisis detallado de la imagen en cada carril y recorrerlo manualmente para poder detectar los pequeños puntos que aparecen como resultado de la digestión de un heterodúplex con la enzima *EndoI*.

Para el análisis del gel se analizó la imagen (con la ayuda del programa *PhotoShop*) ampliándola hasta abarcar todos los carriles para encontrar los cortes de los heterodúplex. La imagen se analizó conjuntamente en los dos canales (700 y 800) y posteriormente se verificó en cada canal el punto de identificación del heterodúplex comprobándose que la suma del tamaño del corte en los dos canales, coincidía con el tamaño total del fragmento analizado (Figura 2.5).

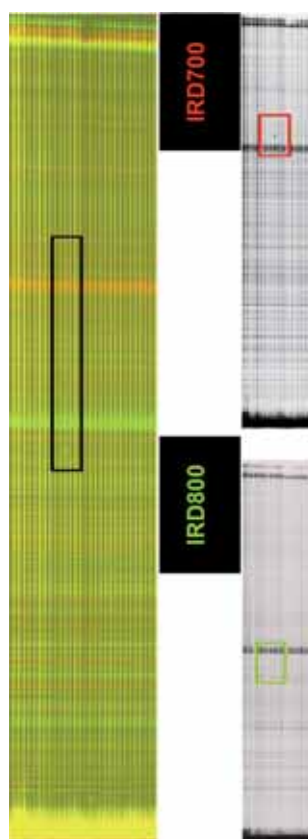


Figura 2.5: Visualización de las muestras en el gel de TILLING. La digestión de un heterodúplex se puede visualizar mediante dos puntos en el mismo canal. Uno (rojo) pertenece al fragmento digerido desde el cebador directo (marcado con IRD 700) hasta el heterodúplex, y el otro fragmento corresponde desde el

cebador inverso (marcado con IRD 800) hasta el heterodúplex. De este modo la suma del tamaño de los dos fragmentos debe dar el total del tamaño del amplicón, que corresponde al homodúplex (banda amarilla de la parte superior del gel). Para visualizar con más detalle y asegurar la especificidad del corte se puede visualizar también el gel en cada canal independiente, en este caso sólo veremos el fragmento marcado para ese canal (imágenes en blanco y negro a la derecha).

Una vez identificados los carriles que contenían una digestión se procedió primero a la identificación del pocillo de la placa mezcla y posteriormente a identificar la familia M2 concreta que contenía la mutación dentro de la mezcla (recordemos que las mezclas están formados por cuatro familias). A este proceso se le llama deconvolución.

Para la deconvolución se elaboró una plantilla (Figura 2.6) dónde señalando la posición del pool mutante, se podían identificar con seguridad las cuatro familias M2 que lo formaban, gracias a la estructura y a los colores de la misma.

Una vez identificadas las cuatro familias se repetía el protocolo (desde la primera PCR, *Nested PCR*, formación de heterodúplex, digestión, purificación, y gel de acrilamida) para poder identificar cuál de las cuatro familias M2 era la que contenía la mutación.

		1	2	3	4	5	6	7	8	9	10	11	12	Pool1	Pool2	Pool3	Pool4	Pool5	Pool6	Pool7+
ABCD	A													D1	D5	D9	D13	D17	D21	D25
EFGH	B		2											D2	D6	D10	D14	D18	D22	c12P1
ABCD	C						1							D3	D7	D11	D15	D19	D23	c12P2
EFGH	D													D4	D8	D12	D16	D20	D24	c12P3
ABCD	E													D5	D9	D13	D17	D21	D25	c12P4
EFGH	F													D6	D10	D14	D18	D22	D26	c12P5
ABCD	G													D7	D11	D15	D19	D23	D27	c12P6
EFGH	H													D8	D12	D16	D20	D24	D28	c12P6

Figura 2.6: Plantilla diseñada para optimizar la deconvolución. Primero se marca la posición en la placa donde se ha detectado el pool mutante. También se coloreará con el color correspondiente a la placa pool dónde hemos encontrado el mutante (en el ejemplo la posición del mutante detectado es C4 de la placa pool 6 (color fucsia)). Posteriormente se procede a identificar la placa dilución de la que proviene el mutante: desplazándonos a la derecha (flecha 1 gris) desde nuestro ejemplo (C4, fucsia) tenemos la fila determinada y nos detendremos en la columna del mismo color . De este modo podremos identificar la placa dilución donde se encuentra el mutante (en el ejemplo la placa dilución es la 22 (D22)). Una vez identificada la placa dilución es necesario identificar la posición de las cuatro familias que componen el pool en la placa dilución. Para ello, utilizamos la flecha 2 y nos desplazamos a la izquierda desde la localización de nuestro mutante. Inmediatamente sabemos que las familias que componen el pool mutante están en las posiciones (A-B-C-D) (gracias a la flecha 2) de la columna 4 de la placa de dilución 22.

2.2.7 SECUENCIACIÓN Y ANÁLISIS DE LA MUTACIÓN

Para identificar el cambio de base producido por el EMS hizo falta secuenciar las familias M2 mutantes. Para ello se amplificó mediante *Nested PCR*s ADN de las familias mutantes de cada gen analizado. La amplificación se comprobó mediante gel de agarosa y el fragmento se purificó mediante Sephadex. Para la secuenciación se utilizó el protocolo descrito en el apartado anterior (2.2.1.2).

Una vez obtenidas las secuencias se analizaron con el programa *Sequencher 4.8* y se buscó en el cromatograma la mutación en forma de doble pico (ya que cada familia contiene ADNs de 10 individuos y en consecuencia se trata de una mezcla de alelos mutantes y silvestres). Una vez detectada la mutación se analizó con el mismo programa si la mutación estaba localizada en zona intrónica o exónica y en este último caso si comportaba un cambio de aminoácido. Para ello, se tradujo la secuencia obtenida del mutante y se comparó con la secuencia de la proteína original. Así se pudo verificar si había un cambio de aminoácido y cómo de drástico podía ser el cambio mediante análisis con el programa SIFT.

2.2.8 CONFIRMACIÓN DE LA SEGREGACIÓN DE LA MUTACIÓN EN LA FAMILIA MUTANTE

Esta parte del protocolo sólo pudo realizarse con los mutantes de la Fitoeno desaturasa. Sin embargo, trabajos posteriores realizados por otros colaboradores del laboratorio han permitido que tengamos datos de segregación de otros mutantes del gen *eIF4E* (éstos se presentaran posteriormente en el apartado 4.2.2).

Una vez identificada la familia M2 mutante se germinaron 25 semillas de la misma familia (proporcionadas por COMAV-UPV, Valencia, donde se hallan almacenadas). Éstas se germinaron para observar segregación de la mutación y algún fenotipo. En estadio de plántula se procedió a hacer una extracción de ADN de las plantas mutantes para la *PDS* según el protocolo descrito en el apartado 2.1.3.2 para poder secuenciar y analizar (según apartado 2.2.1.2) la región donde se encontraba la mutación, y de este modo comprobar que había segregación del alelo mutante e identificar un individuo con el alelo mutante en homocigosis.

2.3 ECOTILLING

2.3.1 PREPARACIÓN DEL ADN

Para el EcoTILLING se preparó el ADN de 113 variedades de melón que se recibieron del banco de germoplasma del COMAV (Anexo II.j). De cada uno de ellos se hizo la cuantificación mediante Nanodrop (*Nanodrop® Spectrophotometer ND-1.000*). Una vez determinada la concentración de ADN de cada variedad, se diluyó a 50 ng/μl y se añadió el volumen adecuado de ADN de M62-113 (ADN de referencia) para obtener una mezcla equimolar en la que ambos ADNs estuvieran a 50 ng/μl.

Se diseñó una placa donde cada pocillo contenía la mezcla de las diferentes variedades de melón con el ADN de referencia. Para determinar las diferencias alélicas (haplotipos) entre las variedades se realizó el protocolo de TILLING con los genes *eIF4E* y *ETR1*. En los geles de acrilamida y las secuenciaciones posteriores se pudieron identificar los distintos haplotipos para cada gen.

2.3.2 DETERMINACIÓN DE LOS HAPLOTIPOS

Para la determinación de los haplotipos se seleccionaron las variedades que presentaban heterodúplex en los geles y se secuenciaron con el mismo protocolo indicado anteriormente. Una vez obtenidas las secuencias se analizaron con el programa *Sequencher 4.8* y se identificaron los haplotipos divergentes a la variedad referencia. Finalmente, se compararon los datos fenotípicos procedentes del COMAV para cada variedad con los haplotipos para intentar determinar si existía alguna correlación entre la función del gen y su fenotipo.

SEGUNDA PARTE: PROTOCOLOS GENERALES

Los protocolos descritos a continuación han sido utilizados frecuentemente durante el período de esta tesis aunque no están directamente relacionados con el protocolo de TILLING.

2.4 EXTRACCIÓN DE ARN

Se usaron muestras de pulpa de fruto de melón T111 15 días y 46 días postantesis. El protocolo utilizado corresponde al TRI-REAGENT (SIGMA).

Se tomaron 2 g de cada muestra de pulpa y se pulverizaron con nitrógeno líquido en un mortero previamente esterilizado. Una vez pulverizado, el material se transfirió a un tubo Falcon de 50 ml y se le añadió 20 ml de TRI-REAGENT. Posteriormente se repartió la mezcla en partes iguales en tubos Falcon de 50 ml, y se les añadió otra vez a cada uno 0,2 ml por ml de volumen de TRI-REAGENT. Se agitó y se centrifugó la mezcla a 3.000 rpm durante 45 minutos a 4°C. Posteriormente se recuperó la fase acuosa y se repartió en tubos ependorf de 1,5 ml añadiéndole a cada uno 0,5 ml de isopropanol centrifugándolos a 8.000 rpm durante 30 minutos a 4°C. De este modo se obtuvo un pellet, que se dejó secar entre 10-15 minutos a temperatura ambiente para posteriormente resuspenderlo con 20 µl de agua. La comprobación de la cantidad y calidad del RNA se hizo mediante gel de agarosa. Previamente se hicieron varios lavados de NaOH (1M) al material de electroforesis para evitar la acción de RNAsas.

2.5 GENOTIPADO CON MARCADORES MOLECULARES

2.5.1 CAPS (*Cleaved Amplified Polymorphic Sequence*)

Una vez identificado un polimorfismo en una secuencia puede analizarse si en él existe alguna diana de enzima de restricción. De este modo, si se realiza la amplificación del fragmento y posteriormente, la digestión, el patrón obtenido será distinto entre el genotipo que tiene el polimorfismo y el que no. Esta diferencia de patrón se podrá visualizar en un gel de agarosa (Konieczny y Ausubel, 1993). La concentración de agarosa de estos geles dependerá del tamaño esperado de los fragmentos de digestión. A menor tamaño de fragmento, más densidad debe tener el gel, llegando a usarse normalmente geles de agarosa al 2%. Una vez amplificadas las regiones del fragmento

analizado, se procede a preparar las digestiones con las enzimas de restricción. Éstas se realizaron habitualmente en un volumen total de 20 µl que contienen 2 µl de tampón 10x (correspondiente a cada enzima de restricción), 2.5 U de enzima, 10 µl del producto amplificado. La mezcla se incubó a la temperatura óptima de cada enzima (37°C-65°C) durante 2 h (también variable en función de la enzima). El producto de digestión se visualiza tras electroforesis en agarosa.

2.5.2 MICROSATÉLITES O SSRs (*Simple Sequence Repeats*)

El protocolo usado en el laboratorio para la amplificación de marcadores SSRs estaba puesto a punto con marcaje radioactivo (Katzir y col., 1996). Para la reacción de amplificación se usó el tampón de SSR 1x (Tris-HCl 75 mM pH 8.8, (NH₄)SO₂ 20 M, Tween-20 0.01 %), MgCl₂ 3.5 mM, 0.6 µM de cada cebador (dependiendo del SSR), 166 µM de nucleótidos dATP, dTTP, dGTP, 2 µM dCTP, 0.1 µl de α³³ P-dCTP (3.000 Ci/mmol), 2 U *Taq* polimerasa y 60 mg de ADN molde en un volumen total de 15 µl. Las condiciones de amplificación fueron un primer ciclo de desnaturalización a 94°C durante 1 min, posteriormente 35 ciclos de amplificación de 30 s a 94 °C, 30 s de hibridación a 59°C y 30 s de elongación a 72°C y por último, una extensión final de 5 min a 72°C.

Posteriormente a la muestra se le añadió 10 µl de tampón de carga (95% formamida, EDTA 10 mM, azul de bromofenol 0.01 %, xilencianol 0.01%). Las muestras fueron desnaturalizadas durante 10 minutos a 100°C para poder proceder a la separación mediante electroforesis vertical en condiciones desnaturalizantes en geles de acrilamida durante 2-3 horas.

Después de añadirle a la muestra el tampón de carga y aplicarle el proceso de desnaturalización, se cargaron 4 µl de los productos de amplificación en el gel de acrilamida. El contenido de acrilamida utilizado para los SSRs es habitualmente de un 6% de acrilamida (19:1 acrilamida/bisacrilamida), tampón TBE 1x, urea 7.5 M, 0.05% APS (persulfato de amonio) y 0.1% TEMED (N,N,N',N'-Tetramethylenediamina). Una vez polimerizado el gel, se calentó hasta alcanzar los 50°C para posteriormente cargar las muestras y aplicar un voltaje de 60 W durante 2-3 horas.

Después el gel fue transferido a un papel Whatman 3 MM, secado al vacío y expuesto en un film fotográfico AGFA Curix RP2 hasta el día siguiente.

2.5.3 DETECCIÓN DE SNPs MEDIANTE LA TÉCNICA DE SNAPSHOT

Esta técnica se basa en el protocolo del Kit comercial ABI Prism SNaPshot ddNTP Primer Extension Kit (Applied Biosystems). Normalmente fue usada en los casos en que no se detectaba ninguna enzima de restricción que pudiese diferenciar ambos alelos para diseñar un marcador tipo CAPS. Para iniciar el protocolo debía obtenerse el producto de PCR con los cebadores específicos y comprobar la cantidad así como la calidad de la amplificación. Para este protocolo era necesario diseñar un cebador interno de unas 20 pb que terminara la posición 3' justo antes del nucleótido polimórfico.

Se tomaron 50 ng del producto amplificado y se le añadieron 2 µl de la enzima Exo-SAP-IT (USB, Cleveland, Ohio) para incubar la mezcla durante 15 minutos a 37 °C. Esta enzima contiene una exonucleasa I y una fosfatasa alcalina de gamba que son hidrolíticas y se utilizan para eliminar los restos de dNTPs y cebadores que quedan después de realizar la PCR, de modo que no puedan interferir en los pasos siguientes del protocolo. Posteriormente se inactivó la enzima a 80 °C durante 15 minutos.

De esta reacción se utilizaron 25 ng como molde para la extensión, juntamente con 2.5 µl de SnaPshot Ready Reaction Mix y 2 pmol del cebador interno previamente diseñado, en un volumen total de 5 µl. La reacción de extensión consistió en 25 ciclos de una primera desnaturalización a 96°C durante 10 s, la hibridación del cebador interno a 50°C durante 30 s y la extensión a 60°C durante 30 s. Finalmente se mantuvo la reacción a 4 °C. Posteriormente se trató la reacción con 1 U de fosfatasa intestinal bovina (Roche) a 37°C durante 1 h, para degradar el cebador interno que pudiera quedar en exceso. Se inactivó la enzima llevando la reacción a 75°C durante 15 min. Luego se añadió LIZ500 como marcador de peso molecular. El cebador extendido se visualizó en un secuenciador ABI Prism 3100 (Applied Biosystems) en las condiciones recomendadas en el protocolo de SNaPshot.

2.6 CLONAJE EN PLÁSMIDO

Para el clonaje de ADN en plásmido se utilizaron las cepas de *E. coli* JM 109 competentes por choque térmico que venían adjuntas al Kit del vector de clonaje pGEM-T "Easy Vector System" (Promega). Para iniciar el clonaje se amplificó el fragmento en cuestión mediante PCR. Los cebadores utilizados fueron en cada caso los específicos de la región a amplificar. Posteriormente se tuvo que purificar el fragmento de PCR mediante el Kit comercial Roche (*High Pure PCR Product Purification Kit*). Una

vez hecha la comprobación de la calidad y cantidad del producto mediante gel de agarosa, se realizó la ligación tal y como indica el fabricante del Kit pGEM-T Easy Vector System II. Posteriormente se realizó la transformación en las células competentes JM 109 que también incluía el Kit comercial.

2.6.1 TRANSFORMACIÓN DE CÉLULAS COMPETENTES MEDIANTE CHOQUE TÉRMICO

Entre 1-3 μ l de la reacción de ligación fue añadida a 50 μ l de células competentes que estuvieron previamente incubadas en hielo durante 30 min. El choque térmico se realizó a 42°C durante aproximadamente 50 s y rápidamente se incubó en hielo durante 2 min. Luego se transfirieron las células a un tubo ependorf con 950 μ l de LB (1% bacto-triptona, 0.5 % extracto de levadura, 0.5 % NaCl pH 7.0) para incubarlas durante 1 h a 37°C y con una agitación de 200 rpm.

Las células transformadas fueron seleccionadas en placas LB agar (LB, 1.5 % bacto-agar) con antibiótico de resistencia ampicilina (50 μ g/mL), X-Gal (5-bromo-4-chloro-3-indolil- β -galactosidasa) (20 ng/mL) e IPTG (isopropiltio- β - D-galactosidasa) (40 μ g/mL). Las placas se dejaron incubar toda una noche a 37°C para poder seleccionar las células transformadas.

2.6.2 EXTRACCIÓN DEL ADN PLASMÍDICO (MINIPREP)

Para la purificación del ADN plasmídico se usó el Kit Qiagen® Miniprep (Qiagen), basado en la lisis alcalina de las células bacterianas (Birnbom y Doly, 1979) con alguna modificación. Se utilizaron 1,5 mL de cultivo bacteriano que se había crecido durante toda la noche anterior a 37°C. Posteriormente, se centrifugó a 8.000 rpm durante 3 min de modo que el precipitado fue resuspendido con la solución I junto con RNAsa (100 μ g/mL). Posteriormente se añadieron 400 μ l de la solución II (que produce la lisis celular) y se incubó en hielo durante 5 min con una breve agitación para no dañar el ADN plasmídico. Al añadir 300 μ l de solución III se neutralizó la lisis (puesto que contiene una alta concentración de sales y estas permiten la renaturalización del ADN plasmídico y su estabilización en la solución). Después una centrifugación a 13.000 rpm durante 7 min a 4 °C permitió la precipitación de los restos celulares. El ADN plasmídico contenido en el sobrenadante se transfirió a un tubo nuevo evitando el precipitado. Se le añadió un volumen de fenol-cloroformo y se centrifugó durante 5 min a 13.000 rpm. Se recuperó la fase acuosa añadiéndole 500 μ l de isopropanol y

mezclándolo bien por inversión. Se centrifugó 10 min a 13.000 rpm y se limpió el pelet obtenido. Finalmente se lavó el pelet con etanol al 70% centrifugándolo también durante 5 min a 13.000 rpm. Luego se dejó secar el pelet y se resuspendió en 50 μ l de agua desionizada estéril obteniendo un rendimiento de 1-4 μ g de ADN por 1,5 mL de cultivo.

3. RESULTADOS

3.1 CONSTRUCCIÓN DE UNA PLATAFORMA DE MUTANTES TILLING EN MELÓN

Durante los últimos años la técnica del TILLING se ha usado ampliamente para el apoyo de la mejora genética en plantas (Wang, 2006), aunque en melón todavía no se había desarrollado ninguna población de mutantes. Empresas privadas como *Semillas Fitó* se habían mostrado interesadas en ayudar a desarrollar herramientas que ayudaran a optimizar la selección y mejora de esta especie. El desarrollo de una plataforma de mutantes representaría una colección de individuos con nuevos alelos que podrían, en algunos casos, proporcionar nuevos caracteres de interés para la especie. La que aquí se describe es la primera población de TILLING elaborada en *Cucumis melo* (aunque durante los cuatro años de este trabajo ya se ha obtenido otra (Dahmani-Mardas y col., 2010)), que ha requerido un largo proceso de puesta a punto y optimización del protocolo. Este trabajo se encuentra enmarcado dentro del proyecto MELOGEN (GEN2013-20237, 2004-2007) que tenía como objetivo desarrollar herramientas genómicas para el melón y que contó con la colaboración de la empresa *Semillas Fitó* y el COMAV (Instituto de Conservación y Mejora de la Agrodiversidad Valenciana) para la puesta a punto de la plataforma de TILLING.

3.1.1 LA POBLACIÓN DE TILLING DE MELÓN

3.1.1.1 COMPROBACIÓN DE LA HOMOCIGOSIS DE LA LÍNEA M62-113

Para desarrollar la población de TILLING en melón se decidió usar la línea pura doble haploide (LDH) de melón tipo “Piel de sapo” variedad *Inodorus* M62-113. La empresa *Semillas Fitó* fue la encargada de subministrar las semillas de esta línea para generar la población de TILLING.

Para corroborar que la línea era realmente homocigota, se analizó con una batería de microsatélites (SSRs) tanto la línea original M62-113 cómo las líneas parentales usadas frecuentemente en el laboratorio, la línea “Piel de Sapo” T111 y la accesión coreana PI 161375. Analizando 2 SSRs por grupo de ligamento (Tabla 3.1 y Figura 3.1) los resultados mostraron que el genotipo de la línea usada era igual que el genotipo de

T111, y distinto del PI y a su vez, pudimos confirmar que era homocigota para los 24 SSRs ensayados.

Tabla 3.1: Descripción de los SSRs utilizados para la comprobación de la homocigosis de la línea M62-113.

GL	Nombre marcador	SSR	Cebador directo/reverso	Referencia	Tamaño alelos (pb)	
					PS/M62-113	PI
I	ECM85	TJ85	AGGACAGCGGAGCTTTTCTT	Fernández-Silva y col., 2008	107	110*
			TGAAATCGAAGTCCACTCTGAA			
I	CMCTN4	NR4	AAAACAAAAGCTCTCCACGA	Gonzalo y col., 2005	126	130
			CTTCCTTTATTATGCCTACG			
II	CMAGN39	NR39	GGGCCATTTCTTTTACAT	No publicado	140	160
			TCTCTTAACCTTCTCTCTCC			
II	CMAGN68	NR68	GGAAGGAAATTAGCATGCAC	Gonzalo y col., 2005	180	195
			GCCACTCTGTCTTTCTTCC			
III	TJ30	TJ30	TTAGGGAAGGCAATCAATCG	Gonzalo y col., 2005	178	150*
			AGGAGGAGGGAATGCTTTGT			
III	ECM125	TJ125	GGAAAACGCAAATCAGTGAG	Fernández-Silva y col., 2008	135	110*
			CTGAACGTGGACGACATTTT			
IV	ECM108	TJ108	CGTTGGGGACCATAGGAGTT	Fernández - Silva y col., 2008	113	120*
			TTCTCAACCCTCTGCCCTTA			
IV	CMTCN6	NR6	GCATTGCTCGATCAGTTTTAC	Gonzalo y col., 2005	140	150
			ACTCCGTCAAGATCCCAAAA			
V	CMAGN52	NR52	CCACCAACATAACACACAAC	Gonzalo y col., 2005	120	140
			CTCTCACACTGTTGGGAAGA			
V	CMCTN2	NR2	CTGAAAGCAGTTTGTGTCGA	Gonzalo y col., 2005	165	175
			AAAGAAGGAAGAGGCTGAGA			
VI	CMTCN18	NR18	ACCAATCCATCACTCTCACT	Gonzalo y col., 2005	140	210
			GAAAGAATGGGGAAAAGAG			
VI	CMTCN41	NR41	CCCCAAGATTCGTATTAATC	Gonzalo y col., 2005	120	145
			TGGTAGTAGAGATGATATAC			

VII	CMAGN75	NR75	TGGGTTTCTCTACTACTG	Gonzalo y col., 2005	150	175
			TGCTTTTACTCTCATTCAAC			
VII	ECM172	TJ172	CCCCTTCTCCATTCTACC	Fernández-Silva y col., 2008	111	130*
			TGAGACGTGGCAGAGAAGAA			
VIII	TJ2	TJ2	GAGGAATCCGAGACCACAAC	Gonzalo y col., 2005	152	140*
			GCCAAGTGTGTGTTGAAAA			
VIII	ECM55	TJ55	ACGCGACGATTCCCTCT	Fernández-Silva y col., 2008	111	120*
			CAAACAAACCCGGATCAAAG			
IX	CMTC47	CMTC47	GCATAAAAGAATTTGCAGAC	Danin-Poleg y col., 2000	154	145*
			AGAATTGAGAAGAGATAGAG			
IX	CMATN22	NR22	CGGCAATCATCTTATCTTTC	Gonzalo y col., 2005	166	170
			AAGATTGAAGTGGGAAAATG			
X	CMCTN71	NR71	TCAATTTTTGCCAAACAAGC	Gonzalo y col., 2005	165	205
			CAAGGACACAGATTTAATAC			
X	CMTCN8	NR8	CCTCCGCCACATATTACAAT	Gonzalo y col., 2005	140	155
			TTCATCTTGACACGTAAGAG			
XI	ECM63	TJ63	CCATTCATCTCCTCTCAATTC	Fernández-Silva y col., 2008	100	130*
			GGCGGCAAGTTTAAAGAACC			
XI	CMATN89	NR89	CACTACCTTAAACAGAATTG	Gonzalo y col., 2005	138	175
			GGACAATTTAGGGAGGATC			
XII	TJ29	TJ29	AGCCTAAGCCACCGATTTTT	Gonzalo y col., 2005	137	120*
			TTCCAAGTGGGGTTATGAG			
XII	CMAGN33	NR33	CTGTCTGCTATTCTCCACTTGG	Gonzalo y col., 2005	110	125
			TGTATGCCACGTAGCGAAAC			

*Tamaño aproximado determinado mediante visualización de los geles.

compatible con una buena tasa de supervivencia y fertilidad de las plantas (Winkler y col., 2005). En nuestro laboratorio se realizó un experimento previo al desarrollo de la población de TILLING para poder establecer el rango de dosis más adecuado. Éste consistió en mutar tandas de 100 semillas con distintas concentraciones de EMS (0.1, 0.5, 1.0, 1.5 y 2.0 %) y observar la tasa de germinación al cabo de dos semanas (Figura 3.2). El primer efecto destacable pudo observarse en las plantas mutadas a una concentración de 0.5 % de EMS ya que un 22% de las semillas no se desarrollaron con normalidad. Este valor aumentaba hasta un 50% en las plantas tratadas con dosis a partir del 1% de EMS. Por lo tanto, este resultado indicaba que con rangos superiores a 0.5% de EMS la germinación en la M0 disminuía considerablemente.

En cuanto a la germinación de la generación M2 se observó una pequeña disminución en la dosis de 0.5%, en la cual germinaron el 93.51% de las semillas, mientras que en las dosis al 1% y 1.5% la disminución de la germinación fue algo mayor obteniendo germinaciones del 82.73 y 75.94%, respectivamente.

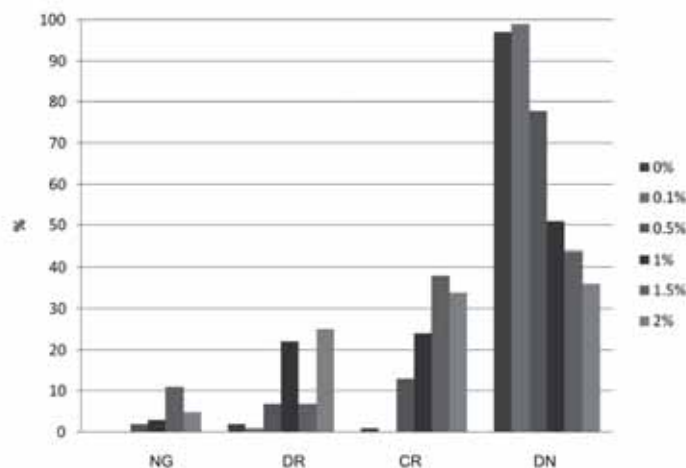


Figura 3.2: Porcentajes de germinación en tandas de 100 semillas M0 tratadas con distintos porcentajes de EMS (de 0% a 2%), dos semanas después del tratamiento según la observación del crecimiento de la generación M1. (NG= semillas no germinadas, DR= sólo desarrollo raíz, CR= cotiledones anormales y desarrollo raíz, DN= desarrollo normal). Estudio realizado por Antonio Monforte, Iban Eduardo y Jordi Garcia Mas durante el año 2004 (IRTA).

Finalmente se decidió usar las dosis de 0,5%, 1% y 1,50% (Tabla 3.2) para mutar las semillas ya que éstas producían unos niveles aceptables de supervivencia en la M1

(germinación superior al 40% (Figura 3.2)), fertilidad en la M1 y vigor en la germinación y desarrollo de la generación M2. El tratamiento se hizo en 17.000 semillas, aplicando distintas dosis de EMS. Las semillas tratadas se repartieron entre los dos centros colaboradores (*Semillas Fitó*, en Cabrera de Mar (Barcelona) y COMAV-UPV (Valencia)) para su crecimiento. En cada lugar se crecieron las plantas obteniendo la generación M1, aunque en el caso del COMAV hubo una drástica pérdida de plantas debido a una elevada incidencia de colapso del melón en el año 2005. Las plantas M1 se autofecundaron y se obtuvieron 3140 frutos, las semillas de las cuales, se plantaron para obtener la generación M2. De cada uno de los frutos se plantaron 10 semillas, obteniendo así la generación M2. Finalmente, para cada familia se cogió un disco de hoja de cada una de las diez plántulas M2, que se mezclaron para extraer el ADN. El número de familias M2 en las que se obtuvo ADN fue 2368. Estos ADNs fueron enviados a nuestro laboratorio en Cabrils donde se transfirieron a placas para posteriormente organizar las mezclas que formarían la población de TILLING.

Tabla 3.2: Dosis de EMS usadas para mutar las semillas, número de plantas con frutos obtenidos en la generación M1 y número de familias M2 con ADN disponible.

Dosis EMS	Número de frutos recogidos en las plantas M1	Número de familias M2 incluidas en la población de TILLING
0,50%	204	177
1%	2544	1922
1.5%	392	269
TOTAL	3140	2368

3.1.1.3 PREPARACIÓN DE LAS MEZCLAS DE ADN DE LA POBLACIÓN DE TILLING

La población de TILLING está representada en 7 placas pool (o mezcla) donde se han mezclado los ADNs de las 2368 familias en mezclas de cuatro familias. La población consta de tres tipos de placas: unas con los ADNs originales de las extracciones para cada familia M2 (placas "Stock"), otras con el ADN diluido 1/10 (placas "Dilution") a una concentración de (5 ng/ μ l) y finalmente las placas de trabajo que incluyen la mezcla de cuatro familias. Las mezclas se han elaborado con un criterio de orden muy estricto puesto que posteriormente a la detección de los mutantes debe poderse aislar e identificar fácilmente la familia que contiene la mutación.

El orden del ADN en las placas corresponde al orden de las tres tandas de mutación de las semillas. La nomenclatura usada para las familias M2 la establecieron los dos centros colaboradores en función del orden de recogida de la planta en cada centro, así todas las familias que se crecieron en *Semillas Fitó* (Cabrera de Mar) comienzan por "F", mientras que las que se crecieron en el COMAV-UPV (Valencia) comienzan por "C". Las semillas de todas las familias están correctamente guardadas y etiquetadas en el COMAV-UPV. El listado y la posición para cada familia M2 en las placas de la población, se encuentran detallados en el Anexo III.a.

3.1.1.4 PUESTA A PUNTO DEL MÉTODO TILLING

Dado que el método TILLING no había sido usado para desarrollar poblaciones de mutantes en melón, era necesaria la optimización de la técnica en nuestro laboratorio.

Para poder desarrollar óptimamente la población de mutantes se decidió realizar una estancia en el grupo del Dr. Abdel Bendahmane del URGV-INRA (Evry, Francia). En una primera fase en el año 2005 se aprendió la técnica para poder desarrollarla en nuestro laboratorio. La segunda fase en el 2008 sirvió para analizar los genes en las distintas poblaciones de mutantes que se habían desarrollado en el laboratorio del INRA. De ese modo se pretendía compararlas y poder determinar la eficiencia de la población IRTA. Por este motivo, en los apartados siguientes se irán describiendo mutantes caracterizados a partir de la población del IRTA y del INRA.

3.1.1.4.1 POBLACIÓN DE TILLING DEL INRA

La población de TILLING del INRA tiene algunas diferencias respecto a la población del IRTA ya que consta de varias subpoblaciones obtenidas por distintas empresas privadas de semillas (Takii, ASL y Gautier). A diferencia de la nuestra, está realizada con una variedad tipo *Charentais*. El número de familias que la componen es de 2483 y éstas están mutadas usando dosis comprendidas entre 1 y 3% de EMS, por lo tanto la concentración de EMS se aumentó considerablemente respecto a la dosis usada en la población IRTA. Además, las mezclas están compuestas por 8 familias M2 mientras que las nuestras contienen 4 familias M2 (Tabla 3.3). También es importante destacar que para mutar la población del INRA se usaron tres protocolos distintos en los cuales no se prehidrataron las semillas antes de aplicar el EMS, aunque los protocolos son confidenciales y pertenecen a cada una de las empresas privadas que los desarrolló. Es importante tener en cuenta este punto puesto que puede influir de un modo determinante en la eficiencia final de mutación de la población.

Tabla 3.3: Tabla comparativa de las dos poblaciones analizadas en este trabajo mediante TILLING.

	Nº familias M2	% EMS	Aplicación del EMS	Variedad	Familias/mezcla
Población IRTA	2368	0.5, 1, 1.5%	con prehidratación de las semillas	Piel de Sapo	4
Población INRA	2483	1, 1.5, 2%	sin prehidratación de las semillas	Charentais	8

3.1.1.4.2 TAQ POLIMERASA EMPLEADA PARA TILLING

Los protocolos de TILLING recomiendan el uso de *Taq* Polimerasa con actividad correctora o de alta fidelidad. Está descrito que la *Taq* tiene una tasa de error de 1.1×10^{-4} errores/pb (Barnes, 1994) por ello se iniciaron los experimentos con *Taq* polimerasas comerciales que contenían actividad correctora como la *Pfu* (Roche) obtenida de *Pyrococcus furiosus*, con una tasa de error 1.6×10^{-6} errores/pb (Lundberg y col., 1991) o la *Expand Taq polymerase* (Roche) con una tasa de error de 4.8×10^{-6} , aunque el coste de las reacciones era excesivo. Se probaron las mismas ampliaciones con *Taq* sin actividad correctora y se determinó (mediante TILLING y secuenciación) si se habían añadido errores durante la amplificación, pudiendo probar que no era así. Finalmente, se decidió que todas las ampliaciones se harían con *Taq* sin actividad correctora, verificando posteriormente la mutación mediante secuenciación. Este punto es importante puesto que reduce el coste de la técnica considerablemente.

3.1.1.4.3 OPTIMIZACIÓN DE LA TÉCNICA TILLING

Una herramienta indispensable para poner a punto el TILLING es disponer de un control. Se decidió usar un SNP previamente identificado en nuestro laboratorio. En 2006 Nieto y col. identificaron el gen *nsv* (correspondiente al gen *elF4E* de melón) que confiere resistencia al virus MNSV (*Melon Necrotic Spot Virus*). El gen *nsv* codifica para un factor de iniciación de la traducción de proteínas y una mutación en este gen confiere resistencia a ciertos virus (Robaglia y Caranta, 2006; Nieto y col., 2005). La mutación que da resistencia a MNSV es un cambio de aminoácido de una leucina en el genotipo resistente PI 161375 (PI) a una histidina en el genotipo susceptible (T111) en la posición 228 de la proteína. Esta posición se localiza en el exón 5, más concretamente en la posición 683 del gen (Nieto y col., 2006). Decidimos usar este SNP como control positivo para poner a punto el método. Para ello, se secuenció el fragmento del gen

eIF4E donde está descrita la mutación en la línea M62-113 y se comparó con la secuencia de las líneas T111 y PI. La secuencia de M62-113 y T111 en esta región es idéntica. Asimismo se localizaron otros polimorfismos entre M62-113/T111 y PI (Figura 3.3 a).

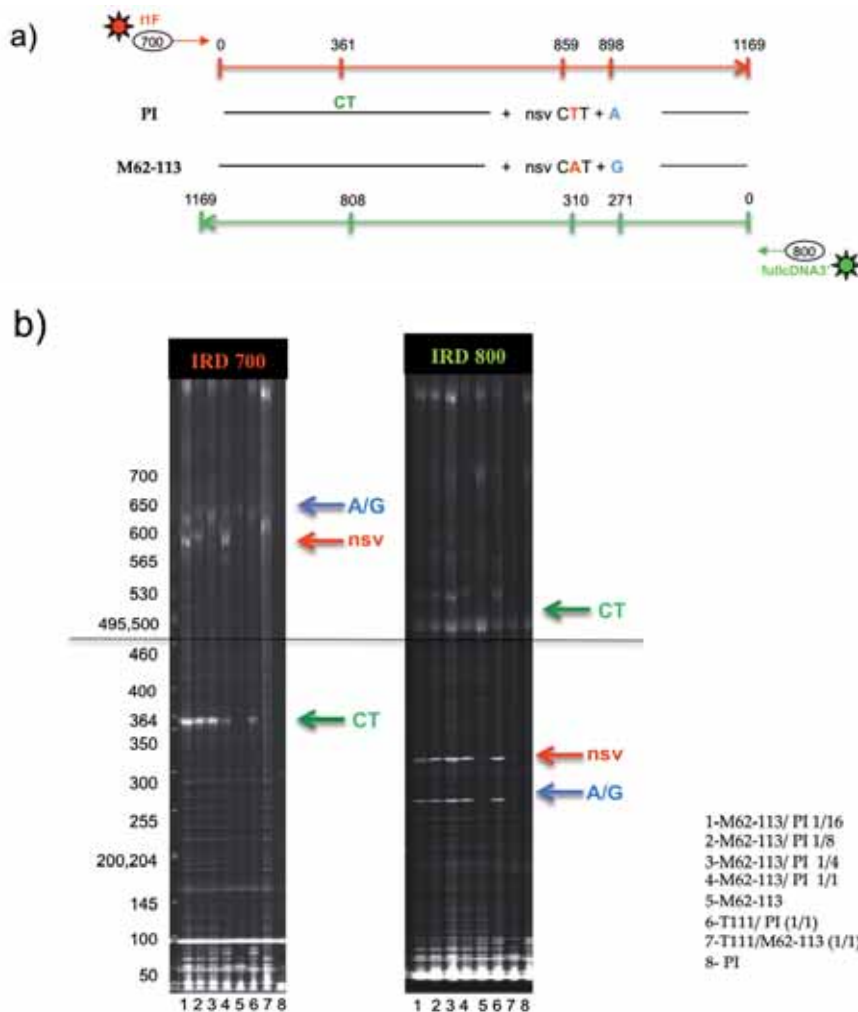


Figura 3.3: a) Esquema del fragmento secuenciado del gen *eIF4E* con las posiciones de los polimorfismos encontrados entre las variedades M62-113 y PI. b) Gel de LI-COR una vez aplicado el método de TILLING a distintas muestras de ADNs. Cada panel representa los canales IRD700 e IRD800, respectivamente. * La línea en gris determina el tamaño a partir del cual el gel perdía resolución (y por lo tanto los tamaños de los heterodúplex visualizados en el gel no coinciden con el tamaño real) como consecuencia del aumento excesivo de la temperatura del gel de acrilamida. Se comentará en detalle en el texto.

A parte del SNP responsable de *nsv* a 859/310 pb de distancia de los cebadores *IIF/fullcDNA3'* encontramos una inserción CT en el genotipo PI en la posición 361/808. El tercer polimorfismo encontrado era un SNP que en el genotipo M62-113 contenía una G mientras que en el genotipo PI contenía una A en la posición 898/271. De modo que al aplicar el protocolo de TILLING en la amplificación de este fragmento usando como ADN molde la mezcla de ambas variedades obteníamos un patrón de digestión característico en los geles de acrilamida (Figura 3.3 b).

Se usaron cebadores ya diseñados anteriormente para amplificar la zona del SNP (Nieto y col., 2007) (*Intron 1F y full cDNA3'*) y posteriormente se realizó el protocolo variando diversos parámetros para poder determinar las condiciones óptimas y poder estandarizarlas. Estos parámetros fueron i) la cantidad de ADN que debíamos añadir en la digestión, ii) el porcentaje de dilución de la endonucleasa, iii) el sistema de purificación de la reacción de digestión, iv) la concentración de la muestra (el tiempo de centrifugación y la temperatura aplicada en el *SpeedVac*) y v) el modo de carga en el LI-COR (con peine o manual). Finalmente, después de varias pruebas, se estandarizaron las condiciones en las que el SNP control se veía con mayor resolución. Se demostró que la purificación de la reacción de digestión era necesaria puesto que si no se realizaba aparecían unas bandas en el gel de LI-COR que interferían en la visualización de las digestiones. Las condiciones óptimas de concentración de la muestra en el *SpeedVac* fueron de 45 minutos a 65°C, puesto que a menor tiempo y menor temperatura la muestra quedaba demasiado diluida para obtener una visualización correcta en el LI-COR. Finalmente, se comprobó que el método de carga usando los peines desechables (CAJ96) proporcionaba una visualización con mejor resolución para la identificación de los carriles que contienen mutantes así como la capacidad de carga (96 muestras por gel mientras que la carga manual sólo permite la carga de 64 muestras).

En la Figura 3.3.b puede verse el resultado de una de las pruebas para determinar los parámetros óptimos del protocolo y comprobar la sensibilidad del método. En el canal IRD700 se visualiza la digestión desde el cebador directo (*IIF*) hasta el punto de la digestión, mientras que en el canal IRD800 se visualiza el fragmento desde el cebador reverso (*fullcDNA3'*) hasta el mismo punto de la digestión (Figura 3.3a). En consecuencia, los tamaños de los fragmentos digeridos que se visualizan en ambos canales deben sumar lo mismo que la totalidad del tamaño del fragmento sin digerir. Del carril 1 al 8 pueden verse las digestiones resultantes según los distintos ADNs empleados. Las mezclas que contengan ADNs con secuencias distintas mostraran la

digestión producida por *EndoI* (carriles del 1-4 y 6) mientras que las muestras que contengan ADNs iguales en secuencia (7) o no sean mezclas de ADNs (5 y 8) no mostraran digestión. Puede observarse que en cualquiera de las diluciones hechas en las mezclas de los mismos ADNs (carriles 1-4) la digestión podía detectarse con la misma intensidad. Por lo tanto se comprobó que en las distintas diluciones de ADN 1/16, 1/8, 1/4 y 1/1 (carriles 1, 2, 3 y 4) los heterodúplex se visualizaban perfectamente y prácticamente con la misma intensidad. Las muestras son mezclas de distintos ADNs y de distintos heterodúplex. Las endonucleasas no cortan totalmente todos los heterodúplex simultáneamente. Por este motivo en los geles se pueden visualizar los fragmentos de los distintos tamaños dependiendo del heterodúplex que haya digerido la endonucleasa. También por ello, en los geles se pueden visualizar el tamaño total del fragmento analizado (perteneciente a aquellas muestras en las que no se ha digerido ningún heterodúplex) y los tamaños de cada heterodúplex (pertenecientes a aquellas muestras en las que se ha digerido cada uno de ellos).

Otra optimización posterior fue cambiar el modelo de LI-COR. El nuevo modelo 4300, nos permitió mejoras importantes: reducir el tiempo para correr las muestras, aumentar la resolución y analizar IRD-700 y 800 conjuntamente, lo que permite la identificación precisa de las digestiones y a la vez, descartar los artefactos.

Mientras que en el modelo antiguo (Li-COR IR²) el tiempo de electroforesis duraba aproximadamente siete horas, con el LI-COR 4300 conseguimos reducirlo a 3-4 horas dependiendo del fragmento a analizar. Además, esto permitía que los geles tuvieran más resolución ya que con 7 horas de migración la acrilamida aumentaba excesivamente de temperatura y, en consecuencia, la parte superior del gel perdía mucha resolución (Figura 3.3b, tamaños superiores a 460 bp).

3.1.2 CONCLUSIONES

Tras el desarrollo y optimización de la plataforma de TILLING en melón cabe destacar que es importante un estudio previo de mutagénesis para obtener una eficiencia de mutación adecuada. Para ello, el estudio del desarrollo de la planta tras la germinación así como el de la capacidad de producir semilla M2 viable es muy importante. El protocolo de mutación de las semillas debe ser optimizado previamente no sólo para fijar la dosis de mutágeno más eficiente sino también para aplicarlo de un modo eficaz. Según nuestros datos el porcentaje de mutágeno óptimo en melón está probablemente dentro del rango del 1,5-2,5%.

Una vez desarrollada la población es necesaria la puesta a punto del método de detección siendo indispensable un control positivo que permita optimizar todos los pasos del protocolo, así como disponer de una tecnología que permita la detección rápida y eficaz. En el caso de este trabajo se decidió usar como control el gen *elf4E* puesto que ya se sabía que contenía un polimorfismo entre los genotipos del melón T111 y el melón coreano PI. Pudimos determinar que no parece necesario el uso de *Taq* con actividad correctora y que el método es preciso con la muestra de ADN mutante diluida hasta 16 veces. Al mismo tiempo, demostramos que el paso de la purificación de la digestión con *EndoI* es indispensable para una visualización óptima de los fragmentos en los geles y que el LI-COR 4300 rinde una muy alta resolución y reduce el tiempo de migración de las muestras.

3.2 BÚSQUEDA DE MUTACIONES EN GENES DE IMPORTANCIA AGRONÓMICA Y ESTUDIO DE SU EFECTO FENOTÍPICO

El método TILLING está diseñado para usarlo generalmente como herramienta de genética inversa: primero se elige un gen diana del cual queremos encontrar mutantes y una vez encontrados y estudiados sus fenotipos se puede formular una hipótesis sobre el efecto o la funcionalidad del gen analizado. Los genes diana en este trabajo han sido seleccionados básicamente por su posible interés agronómico, excepto en el caso del gen de la *PDS* (Fitoeno desaturasa) que fue elegido como control. Los efectos de algunas mutaciones en este gen han sido descritos en otras especies y se caracterizan por un fuerte efecto prematuro y que puede detectarse visualmente, enanismo o albinismo de la planta, es por ello que este gen era un buen candidato como control fenotípico en este trabajo (Qin y col., 2007).

Los genes de la familia de los factores de iniciación de la traducción (*eIF4E* y *eIF(iso)4E*) son buenos candidatos para TILLING por su implicación en resistencias a virus, más concretamente, a potivirus (Robaglia, 2006; Ruffel y col., 2004). Finalmente el *ETR1* (*Ethylene Receptor 1*) puede ser un gen interesante por su implicación en la recepción de etileno y en consecuencia, en la maduración del fruto (Klee, 1999; Wilkinson y col., 1995).

La estrategia seguida para poder secuenciar los genes seleccionados fue encontrar ESTs o unigenes de estos genes en bases de datos de ESTs de melón. Una vez encontrada la secuencia de interés se procedió a diseñar cebadores para poder amplificar y secuenciar preferiblemente la región exónica del gen en la línea M62-113. Para cada gen disponíamos de diferente información, ya que en algunos casos (como en el *eIF4E*) el gen ya estaba caracterizado y existía un estudio previo de EcoTILLING (Nieto y col., 2007). Para los otros genes la única información de la que se disponía fueron secuencias parciales de ESTs (Tabla 3.4). En algunos casos se secuenció sólo una región del gen, escogida según la información disponible, o tras el uso del programa CODDLE (*Codons Optimized to Discover Deleterious LEsions*).

Tabla 3.4 Datos utilizados para obtener las secuencias de los genes candidatos. En la primera columna los unigenes seleccionados a partir de la base de datos ICuGI, en la segunda el número de accesión, si está disponible, y finalmente, el código de la anotación del gen correspondiente en MELONOMICS.

Gen	Unigen ICuGI	Número de accesión	Código MELONOMICS
<i>eIF(iso)4E</i>	MU44901	EU232172 región: 9935..11914 (Meyer y col., 2008)	MELO3C023037
<i>eIF4E</i>	MU54639	DQ393830 (mRNA completo) (Nieto y col., 2006)	MELO3C002698
<i>ETR1</i>	MU48499	-	MELO3C003906
<i>PDS</i>	MU47114 (antiguo MU9070)	-	MELO3C017772

En términos generales el procedimiento empleado fue que, una vez secuenciado todo el gen o un fragmento del mismo, se desarrollaron cebadores específicos para la zona a analizar mediante TILLING y se analizó cada gen en toda la población de mutantes (7 placas de mezcla). Una vez detectados los mutantes se hicieron las deconvoluciones correspondientes (identificación de la familia responsable de la mutación dentro de la mezcla). Posteriormente, se secuenciaron y analizaron las mutaciones determinando si éstas estaban en una región intrónica o exónica. Las localizadas en una región exónica se estudiaron con más detalle identificando cómo afectaba a la proteína y si la mutación provocaba un cambio de aminoácido. En aquellos casos en que el cambio era significativo (según valor de programa informático SIFT) se solicitó la semilla de la familia mutada y se observó el fenotipo. Se extrajo el ADN de los individuos para comprobar la segregación de la mutación en la familia, aunque este último paso sólo se realizó en el gen *PDS*.

3.2.1 *PDS* (FITOENO DESATURASA)

Para empezar a usar la población de TILLING se decidió seleccionar un gen control con el que *a priori* fuera más sencillo analizar el fenotipo de los mutantes. El gen escogido fue el de la *PDS* (Fitoeno desaturasa) que está involucrado en la ruta de síntesis de carotenos. Qin y colaboradores demostraron que la disrupción de este gen en *Arabidopsis* provoca fenotipos albinos y enanos como consecuencia del efecto de interrupción en la biosíntesis de clorofila, carotenoides y giberilinas (Qin y col., 2007). Así pues, si detectábamos una familia mutante en la generación M2 en la que la mutación implicara un cambio de aminoácido, esperábamos observar algunos de los

fenotipos descritos anteriormente en otras especies. Esto demostraría que la población contiene mutantes, que el método de detección es correcto y que además, pueden obtenerse mutantes con fenotipo asociado.

Simultáneamente a este trabajo se desarrolló una base de datos de ESTs de melón en el marco del proyecto ICuGI (www.icugi.org). De este modo tuvimos a nuestro alcance una herramienta eficiente para buscar ESTs de un gen determinado y poder proceder al diseño de cebadores que nos sirvieran para obtener la secuencia completa del gen en la línea M62-113. Para la *PDS* encontramos un único EST (MU9070), actualmente denominado MU47114.

3.2.1.1 DISEÑO DE CEBADORES PARA EL GEN *PDS*

Partiendo del EST MU9070 que contenía una alta homología (93%) con el gen de la *PDS* de *Momordica charantia* (AAR86105.1), diseñamos doce cebadores combinándolos entre sí (Figura 3.4) para poder obtener la secuencia genómica en la línea M62-113 que incluyera las regiones intrónicas. Finalmente amplificamos y secuenciamos una región genómica de 4537 pb.

Al no disponer de la secuencia completa del gen de la *PDS* en melón buscamos ortólogos en especies cercanas. Pudimos identificar que dicha secuencia correspondía a los últimos 6 exones de la *PDS* de *Momordica charantia* (AAR86105.1) (región marcada en gris en la Figura 3.5). Sin embargo, no se siguió con la secuenciación del gen completo dado que en ese momento no se disponía de más ESTs y además, se consideró que el fragmento secuenciado bastaba para iniciar la búsqueda de mutantes dado que contenía aproximadamente la mitad de región exónica (exones del 9 al 14). Actualmente, se dispone de la secuencia completa del genoma del melón, herramienta que ha permitido identificar el gen completo de la *PDS*. En la Figura 3.5 se puede observar la estructura del gen, que contiene 14 exones.

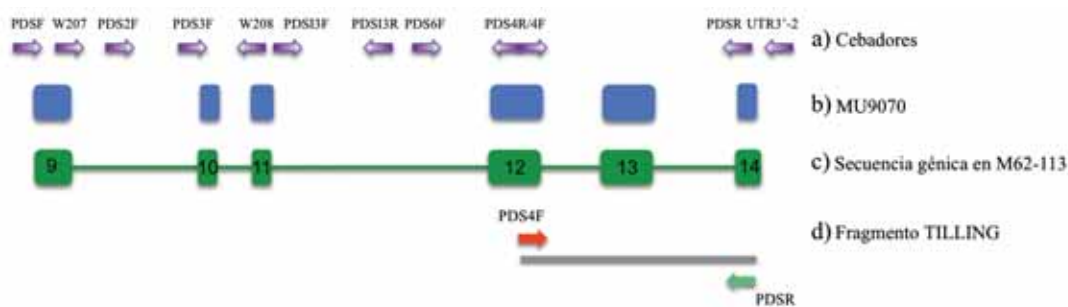


Figura 3.4: Esquema representativo de la región de *PDS* de melón entre los exones 9 y 14. a) En morado, los cebadores utilizados para la secuenciación del gen *PDS* (Anexo II) b) Esquema de la secuencia de referencia de la cual se partió para el diseño de cebadores y que corresponde al EST MU9070 extraída de la base de datos del ICuGI. c) Secuencia obtenida después de la amplificación y secuenciación con los cebadores indicados. d) Amplicón escogido para el análisis mediante TILLING.

Se ha podido realizar un alineamiento múltiple con el gen de la *PDS* identificado en melón (MELO3C017772) y otras especies con tal de observar las regiones más conservadas y posteriormente analizar la posición de las mutaciones detectadas (Figura 3.5). Además, y gracias a herramientas que actualmente están disponibles en melón como el Filoma (phylomedb.org), se ha obtenido el árbol filogenético del gen pudiendo corroborar que la secuencia de la que disponíamos corresponde al gen de la *PDS* (Figura 3.6, flecha roja) ya que el árbol filogenético lo relaciona con su ortólogo de *Arabidopsis* (Figura 3.6, flecha azul), que está correctamente anotado. Al mismo tiempo, esta herramienta nos permite afirmar que hay una sola copia de este gen en el genoma del melón (Figura 3.6). En el mismo árbol filogenético aparecen también dos genes de la Zeta-caroteno desaturasa (ZDS) (Phy003MFP3_CUCME y Phy003A1AY_CUCME), cercana filogenéticamente a *PDS*.

Una vez obtenida esta región de la secuencia del gen escogimos el fragmento a analizar mediante TILLING. Para elegirlo debía tenerse en cuenta que el tamaño óptimo de resolución de los fragmentos en el LI-COR 4300 es de 1000 pb aproximadamente. Por este motivo, los amplicones analizados en este trabajo no sobrepasan los 1200 pb. Para analizar genes de tamaño mayor de 1 Kb hace falta diseñar varios amplicones a partir de distintas partes del gen.

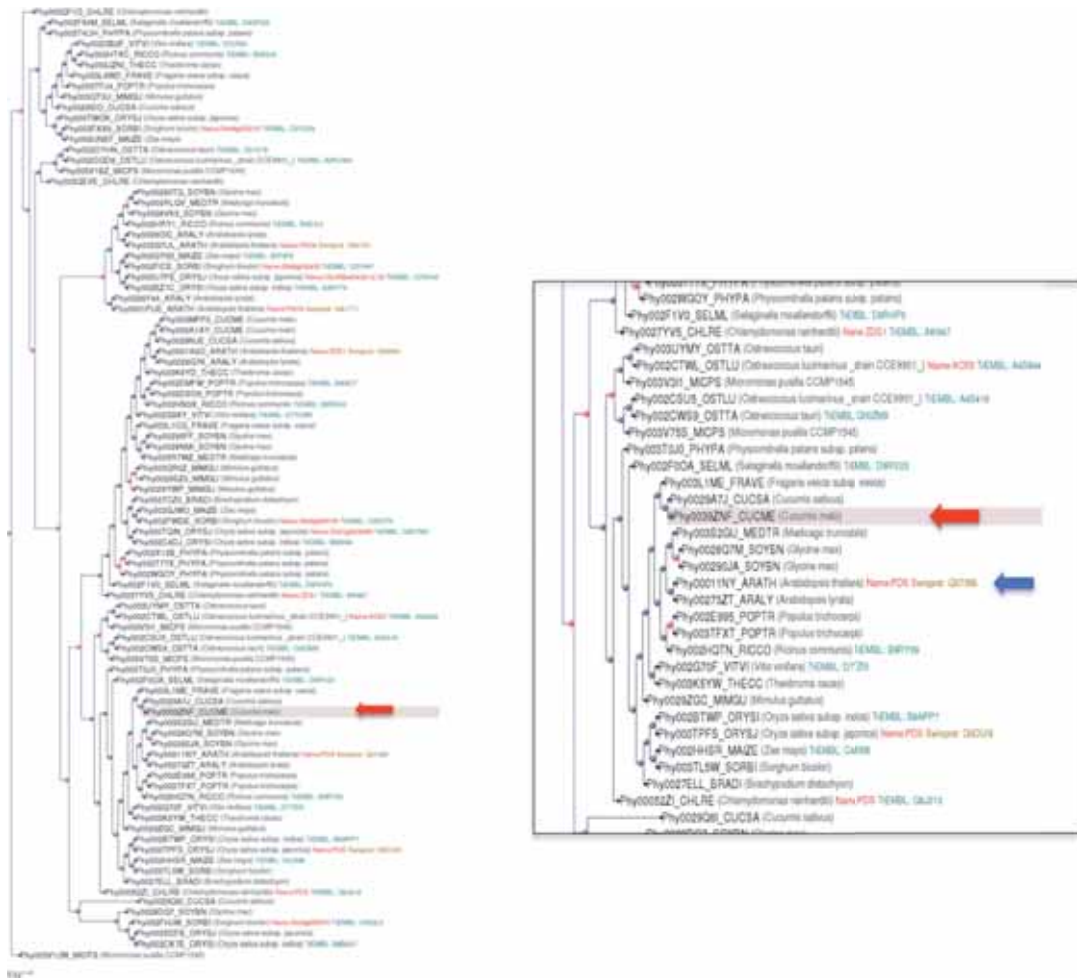


Figura 3.6: Árbol filogenético del gen de la PDS de melón (Phy0039ZNF o MELO3C017772) y de otras 23 especies vegetales. A la derecha el detalle de la posición de la PDS de melón (flecha roja) así como de su ortólogo en *Arabidopsis* (flecha azul) en el árbol filogenético (fuente: Phylomedb.org).

Al mismo tiempo, el programa CODDLE (Spieth y Lawson, 2005) nos indica la región con más probabilidad de sufrir un cambio de aminoácido tras la mutagénesis y que este cambio sea significativo para la funcionalidad de la proteína. Teniendo en cuenta estos datos se eligió el fragmento detallado en la Figura 3.4 d) con un tamaño total de 1189 pb y que incluye prácticamente los tres últimos exones.

En los primeros geles realizados con la PDS se incluyó, a parte del control positivo que se había utilizado hasta ese momento (*eIF4E-nsv*) (apartado 3.1.1.4.3), una muestra que contenía la mezcla de los amplicones PDS a partir de las variedades T111 y PI. Al

aplicar el protocolo de TILLING en esta muestra observamos que aparecían varios heterodúplex (Figura 3.7). Decidimos identificar los SNPs entre T111 y PI, encontrando nueve polimorfismos. Decidimos usar este fragmento como nuevo control positivo en los geles de TILLING, puesto que contenía más SNPs y la resolución era mejor que la obtenida para el fragmento del gen *elF4E* comentado anteriormente.

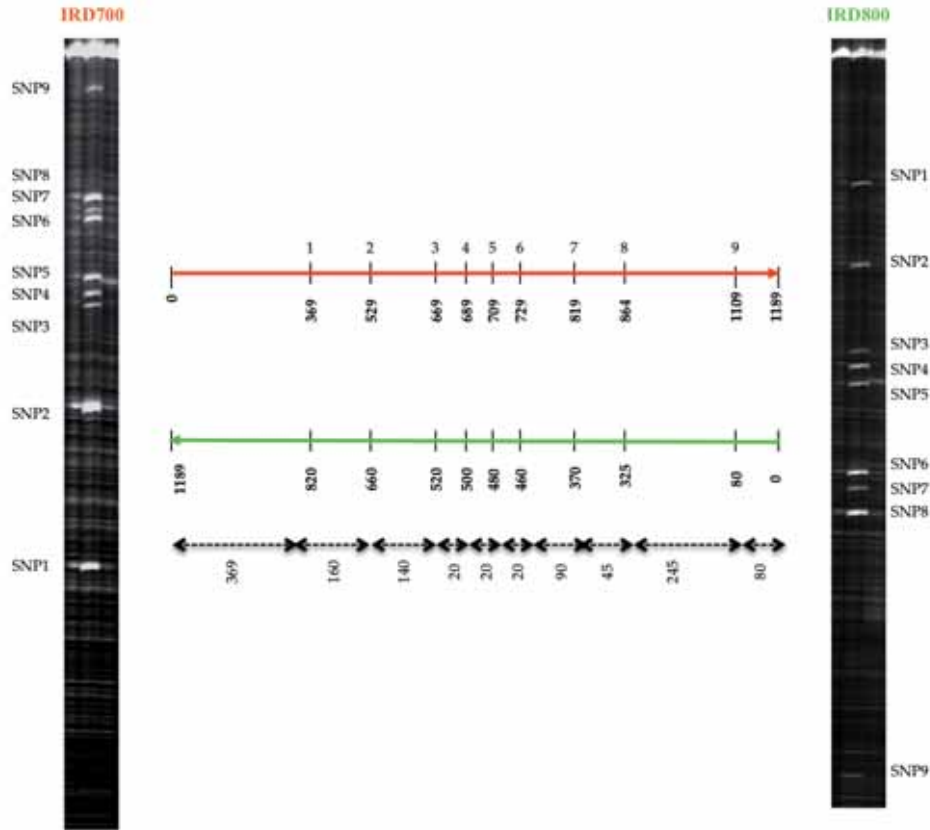


Figura 3.7 Esquema del fragmento de la *PDS* analizado mediante TILLING. A ambos lados de la figura están las digestiones en los dos canales del LI-COR (derecha, IR800, color verde; izquierda, IR700, color rojo) mientras que en la parte central puede verse una representación del fragmento analizado y la posición (partiendo de cada uno de los dos cebadores utilizados (directo: PDS4F y reverso PDSR)) de los SNPs detectados mediante secuenciación entre T111 y PI. En la última barra se detallan las distancias entre SNPs.

3.2.1.2 DETECCIÓN, IDENTIFICACIÓN Y CARACTERIZACIÓN DE MUTANTES EN *PDS*

Tras rastrear 2368 familias M2 de la población IRTA y 2483 de la población INRA se detectaron dos (F1151 y C384) y cuatro (1182, 2495, 1226 y 1460) mutantes respectivamente (Figura 3.9, Tabla 3.5) para el gen *PDS*. Cada familia mutante se

identificó mediante deconvolución y posteriormente se detectó y comprobó el SNP mediante secuenciación y análisis de los cromatogramas. En todos los mutantes se identificó la mutación en el cromatograma como un doble pico del nucleótido en la posición de la mutación, debido a que el fragmento analizado se había amplificado a partir de una mezcla de ADN de diez individuos de cada familia M2. Mediante el estudio del cromatograma (Figura 3.8) se determinó la posición de la mutación y se identificó el cambio de nucleótido.

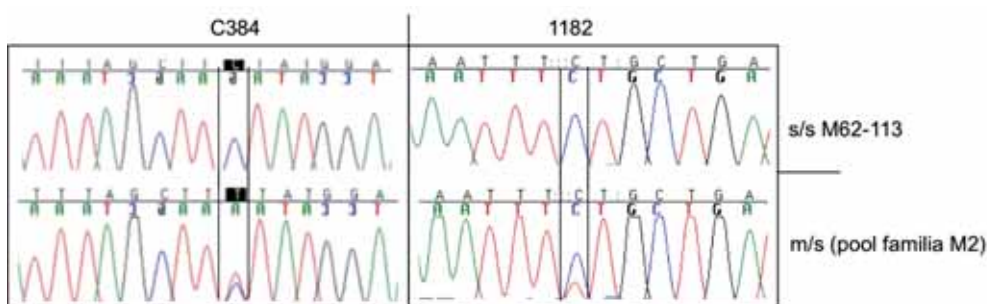


Figura 3.8: Cromatogramas de las dos familias con mutaciones en la zona exónica de la *PDS* (C384 y 1182). Los cromatogramas de la parte superior corresponden a los genotipos homocigotos silvestres (s/s) de M62-113, mientras que los cromatogramas de la parte inferior corresponden a los genotipos heterocigotos mutantes (m/s). En estos últimos se puede observar el doble pico que representa la heterocigosis del alelo (puesto que el ADN *Stock* de cada familia está compuesto por una mezcla de 10 individuos).

En la Figura 3.9 podemos observar la representación del gen de la *PDS* indicando el fragmento analizado mediante TILLING y las posiciones dónde se han detectado las mutaciones en las familias de cada población (IRTA, triángulos en rojo y INRA triángulos en verde). En la Tabla 3.5 se describen con más detalle las características de cada mutación. También, el porcentaje de EMS al que estaba mutagenizada la semilla y el cambio de aminoácido que provoca la mutación. Dos de los mutantes encontrados (1182 y C384) presentan la mutación en región exónica, concretamente en los exones 12 y 13 del gen, respectivamente. Los cuatro mutantes restantes (1226, 1460, 2495, F1151) presentan las mutaciones en el intrón 12. Todos los cambios fueron de citosina a timina, hecho que confirma el efecto del EMS (transiciones G-C/A-T) (Henikoff y Comai, 2003b).



Figura 3.9: Representación de la parte secuenciada de la *PDS* (Fitoeno desaturasa) dónde las cajas representan los exones y la parte lineal los intrones. En la parte superior, se detalla la región del gen de 1188 pb analizado mediante TILLING y en la parte inferior la posición de los mutantes. Los triángulos rojos representan los mutantes encontrados en la población del IRTA mientras que los verdes representan los encontrados en la población del INRA (esta codificación de color se mostrará a lo largo del trabajo).

Tabla 3.5: Mutantes encontrados en el fragmento de 1188 pb del gen *PDS* analizado. En fondo rojo aquellos mutantes encontrados en la población IRTA y en fondo verde aquellos encontrados en la población INRA. *Posiciones respecto al cebador directo PDS4F.

Nº	Familia	Posición	Mutación*	Cambio aa	%EMS	SIFT
1	1182	Exón 12	C395T	Ser485Phe	2%	0.01
2	2495	Intrón 12	C/T		2%	
3	1226	Intrón 12	C/T		2%	
4	1460	Intrón 12	C/T		2%	
5	F1151	Intrón 12	C/T		1%	
6	C384	Exón 13	C560T	Ser540Phe	1%	0.00

Las familias C384 y 1182 contienen una mutación en región exónica que provoca el mismo cambio de aminoácido de serina a fenilalanina (ya que ambas cambian del codón TCT a TTT), la primera en la posición 485 y la segunda en la posición 540 de la secuencia proteica del gen de la *PDS* (MELO3C017772) (Figura 3.5, posiciones marcadas en azul y verde).

La predicción de cómo este cambio de aminoácido puede afectar a la proteína puede comprobarse mediante el análisis con el programa SIFT (*Sorting Intolerant From Tolerant*, <http://blocks.fhcrc.org/sift/SIFT.html>) (Pauline y Henikoff, 2003). Introduciendo en el SIFT la secuencia proteica o la del gen e indicando la sustitución aminoacídica, hace una aproximación matemática mediante comparación con secuencias ortólogas y da una puntuación de cómo este cambio puede afectar a la funcionalidad de la proteína. Según SIFT las substituciones con una puntuación menor de 0.05 afectarán a la funcionalidad

de la proteína mientras que una puntuación mayor indica que probablemente el cambio no afecta a la función de la proteína. En todos los mutantes de todos los genes analizados en este trabajo se ha hecho este análisis. Para la *PDS*, en ambos casos SIFT predice que la mutación puede causar un efecto en la función de la proteína considerándola no tolerada. Teniendo en cuenta que la serina es un aminoácido polar no cargado y la fenilalanina es un aminoácido no polar (Figura 3.10), además del resultado de la predicción de SIFT, decidimos seguir con la observación fenotípica de estos mutantes.

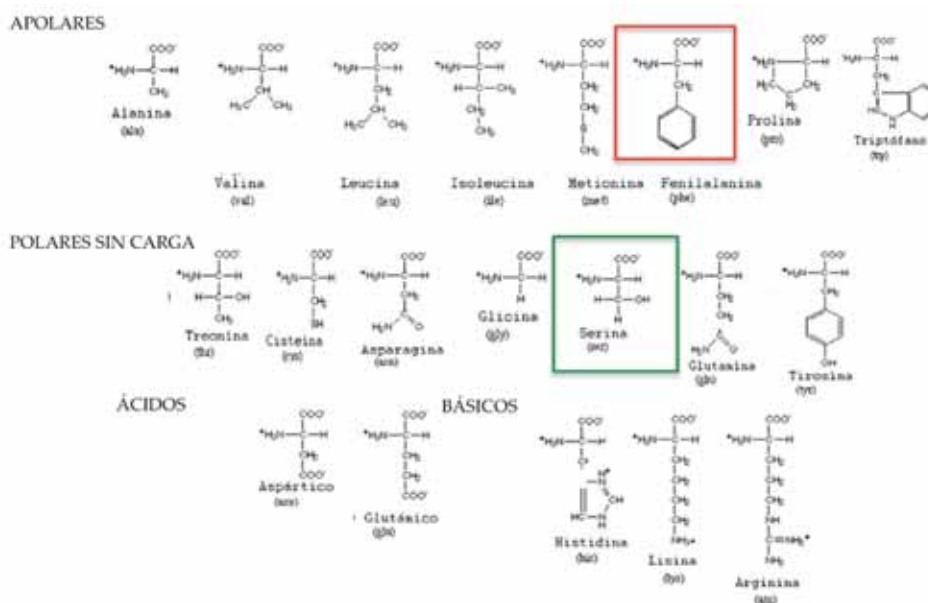


Figura 3.10: Representación de los 20 aminoácidos. Dependiendo del grupo radical éstos pueden clasificarse en apolares, polares no cargados y dentro de los polares cargados entre ácidos y básicos. En el caso de las familias mutantes C384 y 1182 el cambio de serina a fenilalanina puede modificar las propiedades de la estructura puesto que se produce una sustitución de un aminoácido sin carga a uno apolar pero con un grupo radical mucho más grande.

3.2.1.3 ESTUDIO GENOTÍPICO Y FENOTÍPICO EN LAS FAMILIAS C384 Y F1115

Los primeros mutantes detectados en el IRTA mediante TILLING fueron los del gen de la *PDS*. Se continuó el análisis de estas familias porque disponíamos de las semillas de las familias M2 para poder analizar el fenotipo. En cambio, no se continuó el análisis de la familia 1182 puesto que este mutante se detectó posteriormente en el INRA y no hubo tiempo material de solicitar las semillas M2 y fenotiparlas. Para la familia C384 era

factible visualizar algún fenotipo puesto que la mutación estaba en región exónica y producía un cambio de aminoácido (Tabla 3.5). Por el contrario, para la familia F1151, también de la población del IRTA, no se esperaba ningún efecto fenotípico puesto que la mutación se encuentra en región intrónica. Sin embargo, también se realizó el estudio de la segregación de la mutación para esta familia. El porcentaje de germinación de las semillas M2 fue elevado, en la familia C384 fue del 80% mientras que en la familia F1151 fue del 88% (Tabla 3.6).

Tabla 3.6: Germinación de la semillas M2 en las familias C384 y F1151.

Familia	Semillas plantadas	Germinadas	Plantas adultas	Plantas con fruto
C384	20	16	14	8
F1151	25	22	22	20

Para la familia F1151 se trasplantaron al invernadero las 22 plantas germinadas y no se observó ningún fenotipo. En esta familia el porcentaje de éxito de cuajado de los frutos fue elevado (90,9%) (Tabla 3.6). Los frutos de ambas familias fueron medidos, fotografiados y analizados fenotípicamente pero no se observó ninguna variación importante respecto al control (Figura 3.11). Las semillas M3 de estos frutos se recogieron y almacenaron.



Figura 3.11: Imagen de tres frutos de plantas M2 recogidos de las familias C384 (izquierda), F1151 (derecha), y control M62-113 (centro).

De la familia F1151 secuenciamos 15 individuos de los cuales encontramos 4 homocigotos silvestres, 7 heterocigotos y 4 homocigotos mutantes, de modo que se cumplía la segregación esperada 1:2:1 (Tabla 3.7).

En la familia C384 pudo verse un fenotipo característico que consiste en un abultamiento o malformación de la hoja así como un crecimiento ralentizado de la planta. Este fenotipo se detectó en 5 plántulas aunque no en todas ellas con la misma intensidad (Flechas azules en Figura 3.12).

En un estadio más adelantado de plántula (dos semanas tras la germinación) pudo observarse este fenotipo de un modo más acentuado en 6 de las 16 plántulas germinadas. Los cotiledones estaban muy arrugados y se secaban. Las primeras hojas de estas plántulas mostraban un retraso de crecimiento en comparación a las que no presentaban este fenotipo y en algunas de ellas no había ni siquiera presencia de las primeras hojas. De estas plantas murió una de la cual no pudimos recoger muestra. Al cabo de 4 semanas aproximadamente, las plantas se pasaron al invernadero para poder seguir el estudio fenotípico en la planta adulta. En planta adulta, no se observaron características fenotípicas diferenciales de las plantas control, pero una vez autopolinizadas, y teniendo en cuenta que crecíamos un fruto por planta, sólo se obtuvieron 8 frutos de 14 plantas (un 57%) (Tabla 3.6).

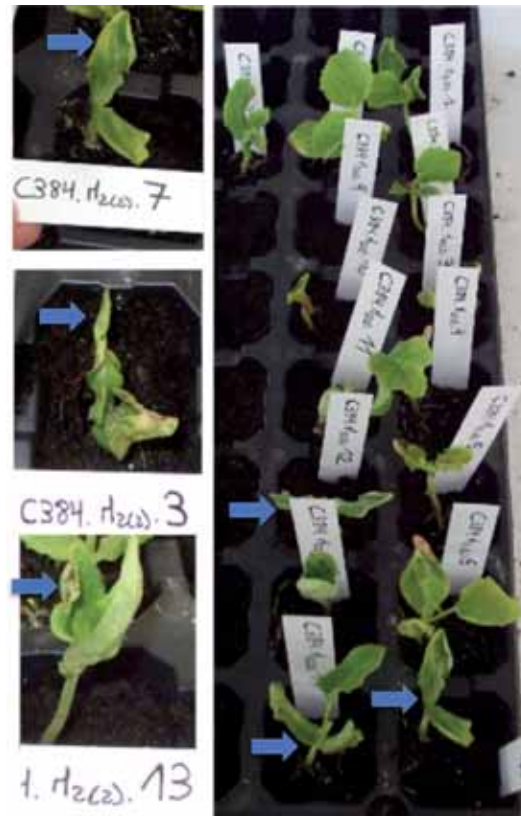


Figura 3.12: Plántulas M2 de la familia mutante C384. Se observó un fenotipo característico consistente en el abultamiento de los cotiledones (flechas azules).

Al secuenciar los 15 individuos germinados de la familia C384 obtuvimos 6 homocigotos silvestres y 9 heterocigotos (Tabla 3.7, experimento 1). No encontramos ningún homocigoto mutante, aunque hay que tener en cuenta que no pudimos recoger material de la planta que murió en estadio de cotiledón. No se pudo establecer una relación directa con el fenotipo detectado en esta familia y el genotipo de los individuos (Figura 3.12). Es por ello que se decidió repetir el experimento (experimento 2) para intentar recoger material de esas plántulas que se morían prematuramente.

Para ello, se germinaron 20 semillas M3 que se habían obtenido de una planta M2 heterocigota para la mutación. Se obtuvieron 15 plantas, lo que representa un porcentaje de germinación del 75%. El fenotipo observado anteriormente (hojas secas, abultadas) se repitió pero además esta vez se obtuvo una planta albina.

Tabla 3.7: Resultados de la segregación de la mutación en las familias mutantes de la *PDS*. Detalle de los individuos encontrados en cada clase genotípica y los datos del test de *Chi cuadrado* (χ^2) con su probabilidad. Para la familia C384 se realizaron dos experimentos. En el primero, realizado en la generación M2, no obtuvimos ningún homocigoto mutante, aunque se murió una plántula antes de poder recoger material. El segundo experimento se realizó con la generación M3, en la cual si obtuvimos un mutante homocigoto que coincidió con el único fenotipo albino.

Familia	wt/wt (M62-113)	wt/m	m/m	χ^2	Probabilidad
F1151	4	7	4	0,0167	>0,9
C384 (experimento 1, M2)	6	9	-	2,9	0,5-0,1
C384 (experimento 2, M3)	14		1	8	0,01

Se recogió material de la hoja y se genotiparon 15 individuos. Esta vez el número de homocigotos silvestres (M62-113) fue de 14, ninguna planta resultó heterocigota y pudimos verificar mediante secuenciación que la única planta homocigota mutante era la planta con fenotipo albino, tal y como puede verse en la Figura 3.13.

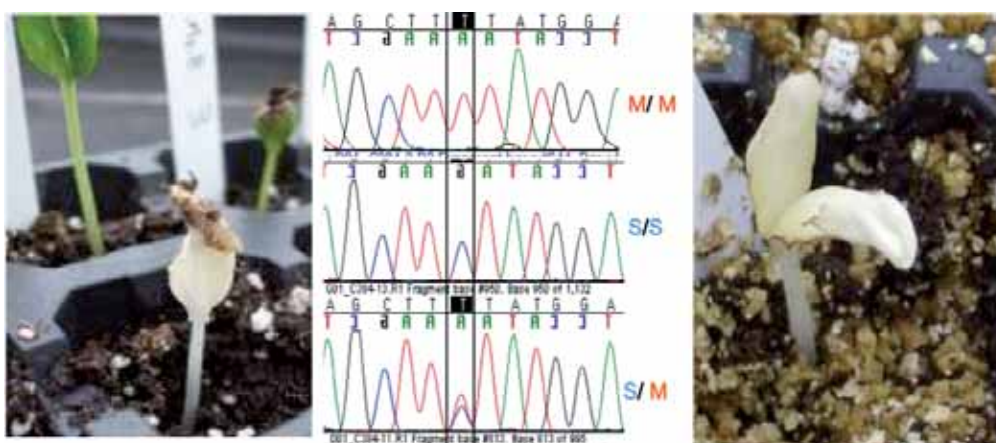


Figura 3.13: En el centro, los cromatogramas obtenidos al secuenciar el fragmento del gen de la *PDS* en los distintos individuos de la familia mutante C384. Puede observarse (de arriba abajo) el genotipo homocigoto mutante que corresponde a la planta M3 albina (imágenes en los laterales), el genotipo homocigoto silvestre, y finalmente el genotipo heterocigoto. La planta albina fue el único individuo homocigoto para la mutación encontrado en la familia C384. (M: Mutante, S: Silvestre).

3.2.2 *eIF4E* (FACTOR DE INICIACIÓN DE LA TRADUCCIÓN)

Las resistencias recesivas en plantas han sido ampliamente investigadas viéndose que los factores de iniciación de la traducción (*eIF4E*) están directamente relacionados con algunas resistencias a virus y más concretamente a los potivirus (Robaglia y Caranta, 2006; Ruffel y col., 2004). En nuestro laboratorio disponíamos de la secuencia del gen *eIF4E* (unigen ICuGI MU54639, número de accesoión ABD57970) y los cebadores para el análisis estaban ya diseñados para EcoTILLING (Nieto y col, 2007). Por ello, se usaron los mismos cebadores que permitían el análisis del gen en dos amplicones, un primer amplicón que abarcaba 500 pb y permitía amplificar el primer exón y un segundo amplicón de 1169 pb que permitía analizar del segundo al quinto exón (Figura 3.14).

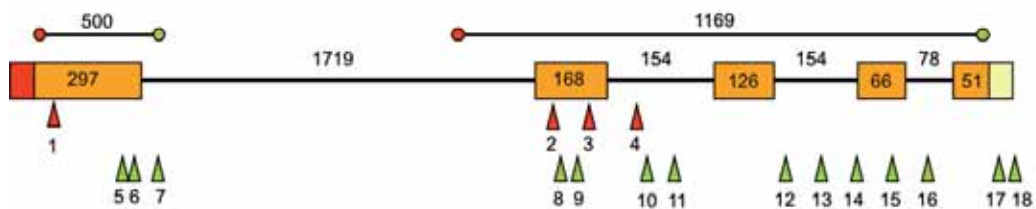


Figura 3.14: Representación del gen *eIF4E* en melón (*nsv*). En la parte superior se muestran los dos amplicones usados de 500 y 1169 pb. Los triángulos detallan las posiciones en las que se han encontrado mutaciones. En rojo, los mutantes encontrados en la población del IRTA y en verde los mutantes encontrados en la población del INRA. En rojo y amarillo las regiones 5' y 3' UTR, respectivamente.

Es importante mencionar que gracias a la reciente obtención de la secuencia del genoma del melón (García-Mas y col., 2012) se ha podido confirmar que sólo existe una copia del gen *eIF4E* en esta especie (Figura 3.15), a diferencia de *Arabidopsis thaliana* que contiene 3 genes, o 2 en tomate. Las tres especies, sin embargo, sólo tienen una copia de la isoforma *eIF(iso)4E* (Rodríguez-Hernández y col., 2012).

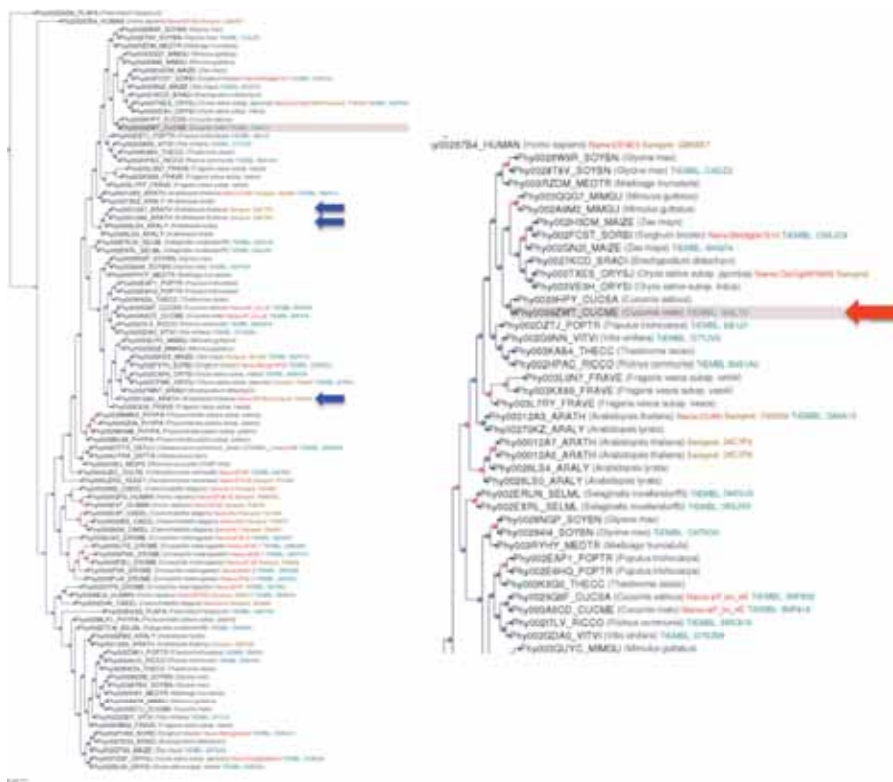


Figura 3.15: Árbol filogenético del gen *eIF4E*. A la derecha se observa la sección ampliada donde se localiza este gen de melón (flecha roja). Se puede observar que sólo hay una copia (fuente: Phylomedb.org) sin embargo, *Arabidopsis thaliana* contiene tres copias (flechas azules).

3.2.2.1 DETECCIÓN, IDENTIFICACIÓN Y CARACTERIZACIÓN DE LOS MUTANTES EN EL GEN *eIF4E*

Para el primer amplicón (Tabla 3.8) se detectó un mutante en la población del IRTA (F2036, mutante nº 1) y tres en la población del INRA (53 (5), 2532 (6) y 1420 (7)). Los mutantes 1, 5 y 6 tienen la mutación en el primer exón mientras que el mutante 7 tiene la mutación en la región del primer intrón aunque muy cercana al final del exón. Al secuenciar los SNPs se ha comprobado que todos los cambios excepto dos (mutante 4, F959, G/C y mutante 10, familia 1287, C/A), son transiciones de C/G a T/A, tal y como era esperado al usar EMS. De las tres mutaciones descritas en el exón 1, dos provocan cambio de aminoácido sin embargo, el SNP del mutante 6 es silencioso. Las mutaciones del mutante 1 (R25K) y del mutante 5 (E97K) provocan sustituciones importantes para la funcionalidad de la proteína según SIFT.

Tabla 3.8: Mutantes encontrados en el primer amplicón de 500 pb del gen *eIF4E*. En rojo, aquéllos encontrados en la población IRTA y en verde en la población INRA. *Posición respecto el ATG en el ARNm.

Nº	Familia	Posición	CDS*	Cambio aa	%EMS	SIFT
1	F2036	Exón 1	G 74 A	Arg25Lys	1%	0.00
5	53	Exón 1	G 289 A	Glu97Lys	1%	0.02
6	2532	Exón 1	C 294 T	Phe98Phe	2%	
7	1420	Intrón 1	G/A		2%	

Para el segundo amplicón (Tabla 3.9) se identificaron un total de 14 mutantes de los cuales 11 se han detectado en la población del INRA y 3 en la población del IRTA. De estos, 5 tienen la mutación en región exónica, 7 en región intrónica y 2 en el extremo 3'UTR. Cuatro de los mutantes que tienen la mutación en la región exónica, provocan un cambio de aminoácido (F57, 1229, 444, 2060) y la familia F1880 contiene una mutación silenciosa.

La familia 1667 contiene el SNP en el primer nucleótido donde empieza el intrón 3, por lo tanto, podría verse afectada la función de *splicing* y en consecuencia la funcionalidad de la proteína. Todas las mutaciones identificadas en este amplicón son transiciones C/G-T/A excepto en F959, en el que el cambio es una transversión de guanina a citosina y 1368, que contiene un cambio de timina a citosina.

Tabla 3.9: Mutantes identificados en el segundo amplicón de 1169 pb del gen *eIF4E*. En rojo, aquéllos encontrados en la población IRTA y en verde en la población INRA.

Nº	Familia	Posición	CDS	Cambio aa	%EMS	SIFT
2	F57	Exón 2	G 385 A	Glu129Lys	1%	0.00
3	F1880	Exón 2	G 465 A	Thr155Thr	1%	
4	F959	Intrón 2	G/C		1%	
8	1229	Exón 2	C 401 T	Ala134Val	2%	0.10
9	444	Exón 2	G 407 A	Gly136Glu	2%	0.00
10	1287	Intrón 2	C/A		2%	
11	1115	Intrón 2	C/T		2%	
12	1667	Intrón 3	G/A primer nucleótido del intrón		2%	
13	2646	Intrón 3	C/T		2%	
14	1233	Intrón 3	C/T		2%	
15	2060	Exón 4	G 616 A	Glu206Lys	2%	0.09

16	2035	Intrón 4	C/T		2%	
17	1368	3'UTR	T/C		2%	
18	2163	3'UTR	G/A		2%	

En resumen, en el gen *eIF4E* se han encontrado un total de 18 mutantes teniendo en cuenta los resultados de los dos amplicones. Ocho de ellos (44,4%) en región exónica, 8 en región intrónica (44,4%) y 2 en el extremo 3'UTR (11,11%). De los mutantes en región exónica, 3 (37,5%) pertenecen a la población del IRTA mientras que 5 (62,5%) a la población del INRA. En la Figura 3.16 se presenta el alineamiento entre el gen *eIF4E* y su isoforma, *eIF(iso)4E*, de modo que puede observarse la homología entre ambos y además la localización de las mutaciones identificadas en el primero de ellos en posiciones exónicas. Algunas de estas mutaciones están localizadas en regiones muy conservadas para ambos genes.



Figura 3.16: Alineamiento entre las proteínas de *eIF4E* y el *eIF(iso)4E* de melón indicando las mutaciones detectadas en *eIF4E* (las verdes correspondientes a mutaciones detectadas en familias de la población INRA y las rojas en la población IRTA).

Es importante destacar que de las cuatro familias mutantes detectadas en la población del IRTA, 3 tienen la mutación en región exónica (75%). Sin embargo, en la población del INRA, de un total de 14 familias mutantes sólo 5 se localizan en región exónica (35,7%).

Según los valores de SIFT, las mutaciones que afectaran a la funcionalidad de la proteína serán las de las familias F2036, 53, F57 y 444, mientras que las mutaciones de

las familias 1229 y 2060 serán toleradas según SIFT (valor > 0,05) y la funcionalidad de la proteína no se verá afectada. Respecto a la distribución de las mutaciones debemos destacar que se han encontrado mutantes en todos los exones excepto en el exón 3. Por otro lado, la distribución de las mutaciones en las partes intrónicas analizadas parece ser homogénea.

3.2.2.2 PREDICCIONES DE LA ESTRUCTURA EN 3D DE LOS MUTANTES *eIF4E*

Para complementar los estudios de predicción con SIFT de cómo pueden afectar las sustituciones aminoacídicas a la funcionalidad de la proteína del gen *eIF4E*, se realizó un estudio detallado de la estructura en 3D de *eIF4E*. El gen *eIF4E* tiene un gran interés puesto que la obtención de una variedad resistente a determinados virus sería un hallazgo importante tanto a nivel científico como comercial. Por este motivo realizamos un estudio más exhaustivo de predicción de cómo podían afectar las mutaciones detectadas a la estructura de *eIF4E*. El modelaje en 3D se realizó con el programa bioinformático PyMOL. En las figuras que se presentan a continuación se pueden visualizar las localizaciones de las mutaciones identificadas que podrían afectar a la funcionalidad de *eIF4E*.

El mutante número 2 (familia F57, mutación Glu129Lys) (según Tabla 3.9) (Figura 3.17) parece contener la mutación más importante en cuanto al posible efecto sobre las interacciones entre la proteína y su ligando. Durante el modelaje se estableció la posible zona de interacción entre *eIF4E* y el 5'CAP, ya que PyMOL ofrecía la posibilidad de definir virtualmente (a través de comparaciones con las bases de datos de cristalización de uniones de proteínas con su sustrato o ligando) el centro activo y el ligando (Figura 3.18). En este caso el ligando es la 7-metil guanosina trifosfato ($C_{11}H_{18}N_5O_{11}P_2^+$) o 5'CAP, molécula que se añade al ARNm y lo protege de su degradación. Esta molécula es el lugar de reconocimiento para el inicio de la traducción y por lo tanto la que interacciona directamente con la proteína *eIF4E* en el momento del inicio de la traducción (Robaglia y Caranta, 2006). Esta mutación es especialmente interesante porque parece que obligaría a establecer enlaces distintos con el ligando, lo que podría tener efecto en la funcionalidad de *eIF4E*.

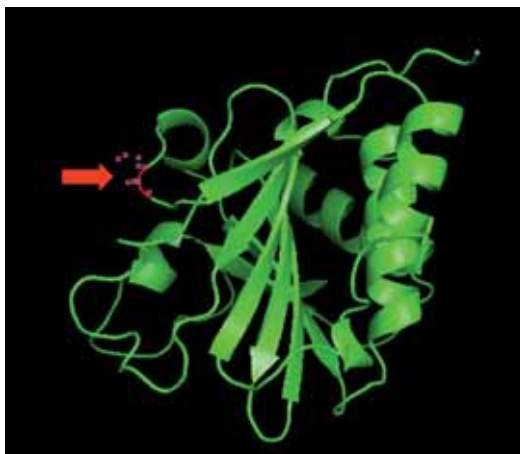


Figura 3.17: Modelaje de la proteína *eIF4E* en 3D mediante el *Output* del programa PyMOL del mutante 2 (familia F57). En rojo y marcada mediante la flecha, la posición de la mutación Glu129Lys.

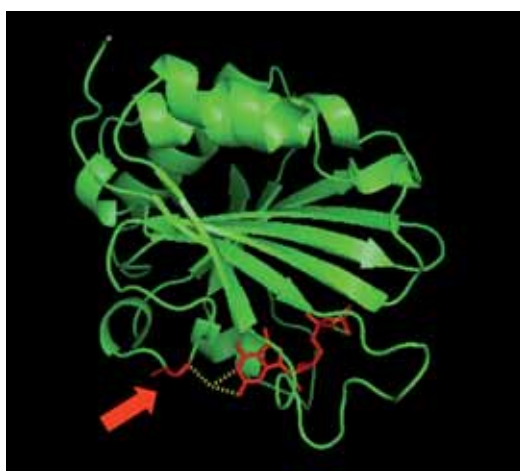


Figura 3.18: Modelaje de la proteína *eIF4E* en 3D mediante el *Output* del programa PyMOL con posible ligando (molécula 5'CAP del ARNm). En amarillo, detalle de las posibles interacciones entre el ligando (en rojo) y el aminoácido Glu129 (en rojo y flecha).

El mutante número 5 (familia 53, mutación Glu97Lys) (según Tabla 3.8) (Figura 3.19) contiene una mutación que puede provocar un efecto comprometido para la funcionalidad de la proteína. El ácido glutámico de la posición 97 en el genotipo silvestre crea un enlace de hidrógeno con la serina 93 (Figura 3.20). En cambio, en el mutante el ácido glutámico queda sustituido por una lisina y ésta pasaría a

interaccionar con la treonina localizada en la posición 91. Por lo tanto, la mutación produciría un cambio en las interacciones entre aminoácidos y podría afectar la estructura o la estabilidad de *eIF4E*.



Figura 3.19: Modelaje de la proteína *eIF4E* en 3D mediante el *Output* del programa PyMOL del mutante 5, familia 53. En rojo y marcada mediante la flecha, la posición de la mutación Glu97Lys.

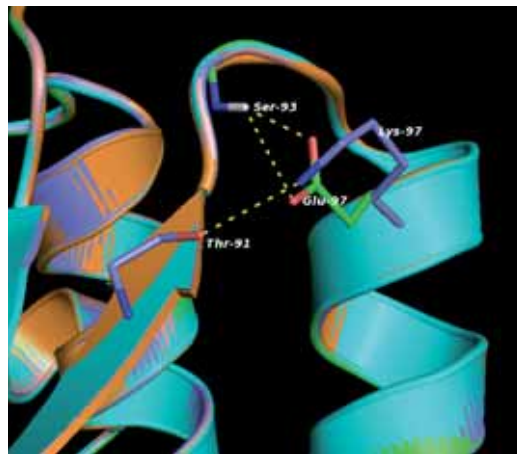


Figura 3.20: *Output* del programa PyMOL del mutante 5, familia 53. Detalle de cómo cambiaría el enlace en el mutante. En el genotipo silvestre interacciona el ácido glutámico-97 con la serina-93 y en cambio, en el mutante interaccionaría la lisina-97 con la treonina-91.

El mutante número 8 (familia 1229, mutación Ala134Val) (según Tabla 3.9) tiene la mutación localizada en la superficie de *eIF4E* (Figura 3.21) por lo que su efecto debería ser menor puesto que el centro activo de *eIF4E* se localiza en una posición interna.

Además, la valina y la alanina no poseen carga, y por lo tanto la sustitución no tendría probablemente ningún efecto a nivel estructural.

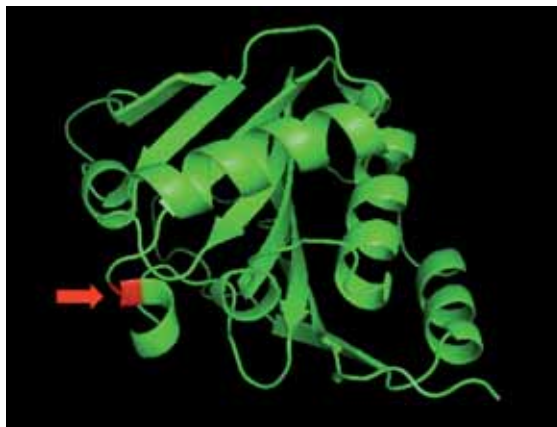


Figura 3.21 Modelaje de la proteína *eIF4E* en 3D mediante el Output del programa PyMOL del mutante 8, familia 1229. En rojo y marcada mediante la flecha, la posición de la mutación Ala134Val.

El mutante número 9 (familia 444, mutación Gly136Glu) (según Tabla 3.9) introduce un cambio de glicina a ácido glutámico, un residuo de cadena larga lateral con carga. La posición de este residuo no es superficial sino interna (Figura 3.22).



Figura 3.22: Modelaje de la proteína *eIF4E* en 3D mediante el Output del programa PyMOL del mutante 9, familia 444. En fucsia y marcada mediante la flecha, la mutación Gly136Glu.

Se han incluido en la Figura 3.23 todas las mutaciones comentadas anteriormente para el gen *eIF4E*. Las mutaciones 1 y 15 no se analizaron.

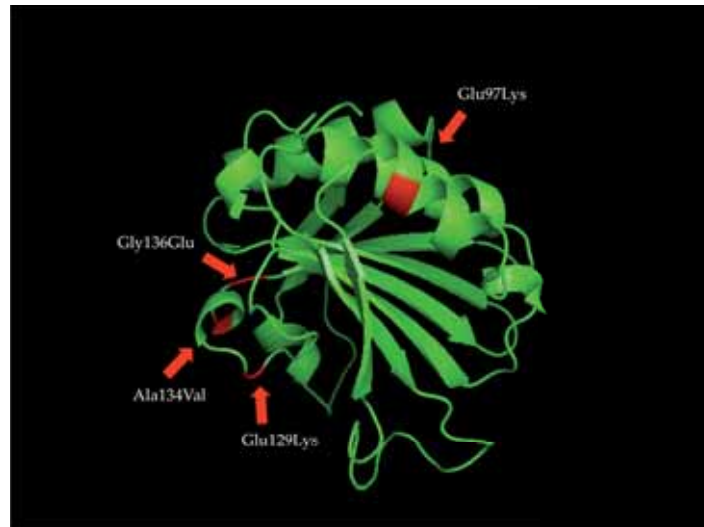


Figura 3.23: Modelaje de la proteína en 3D del gen *eIF4E*. En rojo las mutaciones detalladas anteriormente por su posible implicación en el cambio de funcionalidad de *eIF4E*.

3.2.3 *eIF(ISO)4E* (ISOFORMA DEL FACTOR DE INICIACIÓN DE LA TRADUCCIÓN 4E)

En el laboratorio ya se había iniciado la secuenciación del gen *eIF(iso)4E* (unigen MU44901 de la base de datos de ICuGI) y por lo tanto disponíamos de una parte importante de la secuencia. Aun y así, se diseñaron cuatro nuevos cebadores (Figura 3.24 a) para poder terminar de secuenciar los fragmentos que faltaban para obtener la secuencia completa. Una vez obtenida, resecuenciamos en la variedad M62-113. En ese momento no se había publicado todavía la secuencia completa del gen, sin embargo, actualmente sí está disponible (GeneBank: EU232172 (secuencia del BAC que contiene *eIF(iso)4E* en la región: 9935-11914)) (Meyer y col., 2008). Decidimos iniciar la detección de los mutantes usando un amplicón en la región del gen detallada en la Figura 3.24, de 1014 pb y que contenía tres exones.

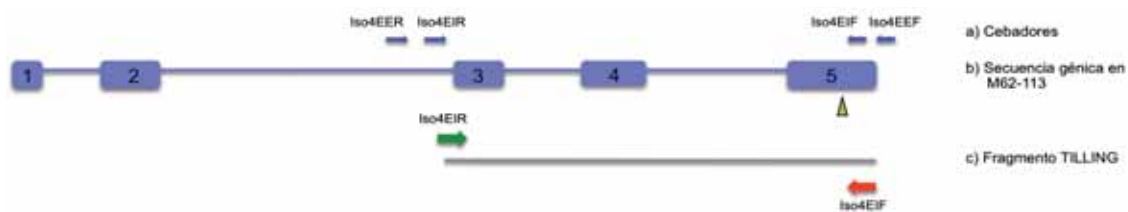


Figura 3.24: (b) Esquema del gen *eIF(iso)4E* compuesto de cinco exones (cajas azules). También puede observarse (c) el fragmento analizado mediante TILLING de 1014 pb así como (a) los cebadores utilizados tanto para secuenciar la región como para detectar los mutantes (Anexo II.j). Los triángulos detallan las posiciones en las que se han encontrado mutaciones. En verde los mutantes encontrados en la población del INRA.

Actualmente este gen ha sido identificado en la secuencia del genoma como MELO3C023037 con una longitud total de 1991 pb. Tal y como se ha podido observar en la Figura 3.16, los genes *eIF4E* y su isoforma contienen regiones muy conservadas. También gracias a la obtención de la secuencia del genoma del melón (García-Mas y col., 2012) se ha podido confirmar que sólo existe una copia del gen *eIF(iso)4E* (Figura 3.25).



Figura 3.25: Árbol filogenético del gen *eIF(iso)4E*. A la derecha se observa la sección ampliada donde se localiza este gen de melón (flecha roja). Se puede observar que sólo hay una copia anotada en el genoma (fuente: Phylomedb.org).

3.2.3.1 DETECCIÓN, IDENTIFICACIÓN Y CARACTERIZACIÓN DE LOS MUTANTES

En el gen *eIF(iso)4E* se ha identificado un único mutante en la población INRA y ninguno en la población IRTA (Tabla 3.10). El mutante detectado pertenece a la familia 1587 y tiene un SNP en el primer exón del gen. Esta mutación provoca un cambio de aminoácido de arginina a lisina, sin embargo, según los valores obtenidos en SIFT, podría ser tolerado y no afectar a la función de la proteína.

Tabla 3.10: Características del mutante encontrado en el gen *eIF(iso)4E* en la población de INRA con el nombre de la familia, posición, tipo de mutación y % de EMS utilizado. La mutación según SIFT no es significativa y no afectaría a la función de la proteína.*Posición respecto al cebador iso4EIF.

Nº	Familia	Posición	CDS *	Cambio aa	%EMS	SIFT
1	M1587	Exón1	G128A	Arg43Lys	2%	1.00

3.2.4 ETR1 (RECEPTOR DE ETILENO 1)

El *ETR1* (Receptor de Etileno 1) pertenece a una familia multigénica (Sato-Nara y col, 1999) y está implicado en la señal de transducción del etileno (Bleecker y Kende, 2000). Están descritas mutaciones en este gen en *Arabidopsis* que provocan insensibilidad al etileno (Wilkinson y col., 1995) y también se ha demostrado que hay una mayor expresión de este gen en semillas y placenta (Sato-Nara y col., 1999). En la base de datos (<http://www.icugi.org>) disponíamos del unigen MU48499 descrito como parte del gen *ETR1*, por ello se inició la secuenciación a partir de este EST. Finalmente, secuenciamos una gran parte del gen utilizando 13 cebadores en distintas combinaciones, detallados en la Figura 3.26.

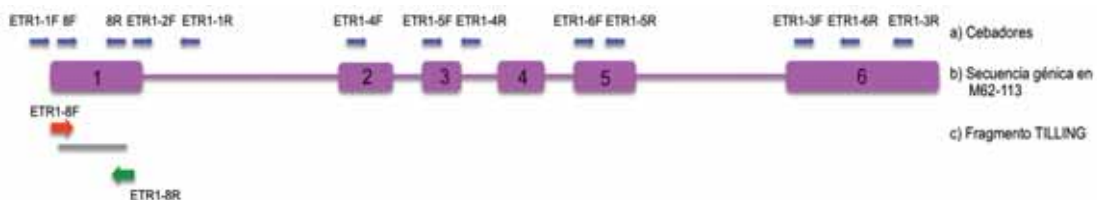


Figura 3.26: Esquema del gen *ETR1* (b), de los cebadores diseñados para su secuenciación (a) (Anexo II.j) y del fragmento del gen de 615 pb analizado mediante TILLING (c).

Actualmente se ha podido obtener la secuencia completa del gen *ETR1* en melón. El gen identificado como MELO3C003906 tiene una longitud total de 4988 pb y está compuesto

por seis exones tal y como se había identificado mediante la secuenciación en la variedad M62-113. Como en los genes presentados anteriormente, también se ha podido obtener el árbol filogenético del gen *ETR1* comprobando que forma parte de una familia multigénica y como tal, encontramos dos genes parálogos en melón (Figura 3.27).

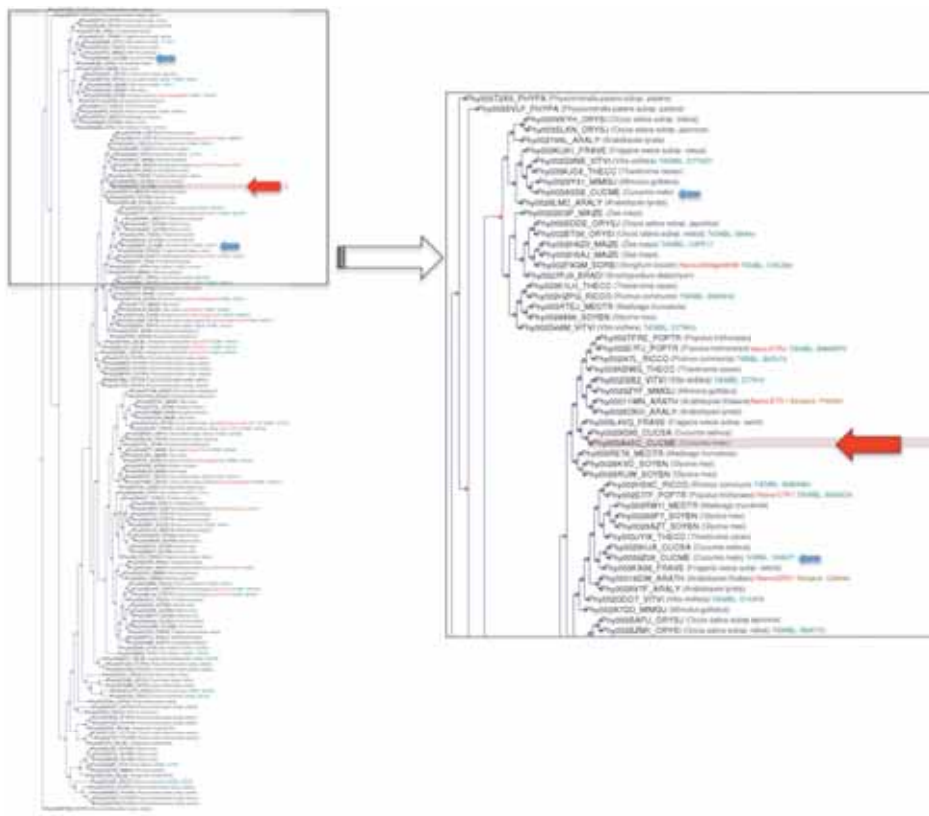


Figura 3.27: Árbol filogenético del gen *ETR1*. A la derecha se observa la sección ampliada donde se localiza este gen de melón (flecha roja). Se pueden observar a la izquierda (señalados con flechas azules) los parálogos al gen *ETR1* (flecha roja) (fuente: Phylomedb.org).

3.2.4.1 DETECCIÓN, IDENTIFICACIÓN Y CARACTERIZACIÓN DE LOS MUTANTES

Tras el análisis mediante TILLING del gen *ETR1* no se detectaron mutantes en la población del IRTA. Por otro lado, en la población del INRA no se pudo analizar este gen puesto que los cebadores diseñados a partir de la variedad usada en la población IRTA, no funcionaron correctamente y no amplificaron en la variedad *Charentais*. Se diseñaron nuevos cebadores aunque siempre a partir de la secuencia de referencia

obtenida de la variedad M62-113. Este segundo diseño de cebadores tampoco amplificó la variedad *Charentais*. Finalmente, no hubo más tiempo para analizar cuál era el motivo por el cuál los cebadores no hibridaban correctamente y por ello se descartó este gen en el análisis de TILLING en la población INRA. Por el contrario, el gen *ETR1* si resultó interesante para el análisis de EcoTILLING que se presentará posteriormente

3.2.5 EFICIENCIA DE LA MUTAGÉNESIS EN LA POBLACIÓN DE TILLING

En la bibliografía se encuentran distintos modos de expresar la tasa de eficiencia de mutación de una población de TILLING. El cálculo básicamente debe tener en cuenta el número de amplicones analizados, el número de familias M2 de las que está compuesta la población y el número de mutantes encontrados. En la Tabla 3.11 se pueden ver las tasas de mutación descritas en diversas poblaciones de TILLING para distintas especies. Hay que tener en cuenta que la tasa obtenida no se ha calculado de igual manera en cada población, por lo que cualquier comparación es aproximativa.

En este trabajo, el cálculo se ha realizado teniendo en cuenta el número de pares de bases analizados en 3 de los genes que se han estudiado (*PDS*, *eIF4E* y *eIF(iso)4E*). Al no poder analizar el gen *ETR1* en la población del INRA se decidió descartar este gen y no tenerlo en cuenta para el cálculo de la tasa de mutación. Así, el cálculo se ha hecho del siguiente modo:

- 1- N° de bases analizadas \times N° de familias que se han analizado = N° bases analizadas en total
- 2- N° bases analizadas en total / N° de mutantes encontrados en cada población = pb/ Mutante

Tabla 3.11: Eficiencia de mutación de distintas poblaciones de TILLING en distintas especies. Las dos últimas, sombreadas, corresponden a las desarrolladas en melón.

Especie	Eficiencia de mutación	Referencia
Zebrafish	1/235Kb-1/527 Kb	Wienholds y col., 2003 Draper y col., 2004;
<i>Drosophila melanogaster</i>	1/156 Kb	Winkler y col., 2005
<i>Arabidopsis thaliana</i>	1/300 Kb	Greene y col., 2003
<i>Zea mays</i>	1/500 Kb	Till y col., 2004

<i>Oryza sativa</i>	1/500 Kb	Wu y col., 2005
<i>Hordeum vulgare</i>	1/1000 Kb	Caldwell y col., 2004
<i>Triticum aestivum</i> (6x)	1/24 Kb	Slade y col., 2005
<i>Triticum turgidum</i> subsp. <i>durum</i> (4x)	1/40 Kb	Slade y col., 2005
<i>Medicago truncatula</i>	1/485 Kb	Le Signor y col., 2009
<i>Cucumis melo</i> (Charentais) (INRA)	1/573 Kb	Dahmani-Mardas y col., 2010
<i>Cucumis melo</i> (Piel de Sapo) (IRTA)	1/1.5 Mb o 0.33/500 Kb	Gonzalez-To y col., 2011

En la población IRTA se han analizado 3.871 pb (1.188 en *PDS*, 1.669 en *eIF4E* y 1.014 en *eIF(iso)4E*) en un total de 2.368 familias. El número total de pb analizados es de 9,17 Mb. Teniendo en cuenta que para los tres genes se han encontrado 6 mutantes, podemos expresar la tasa de mutación como que cada 1.528 Kb podríamos encontrar un mutante. O lo que es lo mismo, 0,33 mutantes/ 500 Kb para poder comparar mejor con las tasas obtenidas en otras poblaciones de otras especies (Tabla 3.11).

Cabe decir que para que una tasa de mutación sea más fiable hace falta analizar un mayor número de genes a los analizados en este trabajo. El número de genes analizados durante el desarrollo de otras poblaciones es generalmente 10 (Wu y col., 2005; Caldwell y col., 2004; Greene y col., 2003; Dahmani-Mardas y col., 2010).

Está descrito que para que una población de TILLING sea útil su tasa de eficiencia debe de ser >1 mutante/500 Kb (Till, 2007). Sin embargo, debemos tener en cuenta que esta tasa de mutación se ha calculado en arroz y en una población de 768 individuos (Till, 2007).

La tasa obtenida en la población del INRA, tras analizar 11 genes ha sido de 1/573 Kb (Dahmani-Mardas y col., 2010). Pero para poder establecer una mejor comparación entre las poblaciones IRTA e INRA, se ha tenido en cuenta para esta última sólo el análisis con los genes que se han analizado en ambas poblaciones (*eIF4E*, *eIF(iso)4E*, *PDS*). Una vez realizado el cálculo de la tasa de mutación de la población INRA con los criterios descritos anteriormente se ha obtenido una tasa de mutación de 1 mutación cada 500 Kb, a partir de 19 mutantes en 2.483 familias.

Tabla 3.12: Tasa de mutación en las poblaciones de TILLING del IRTA y el INRA.

	Mutantes	Genes analizados	pb analizados	Familias	Tasa
Población IRTA	6	3	3872	2368	1/1,5Mb
Población INRA	19	3	3872	2483	1/500Kb

Finalmente, teniendo en cuenta la eficiencia requerida según Till, 2007 se ha calculado el rendimiento de la población:

$$\frac{\text{Eficiencia real}}{\text{Eficiencia teórica}} \times 100 = \% \text{ de rendimiento Eficiencia esperada}$$

Por lo tanto en la población IRTA se ha obtenido un 33% $((1/1500 \text{ Kb}/1/500 \text{ Kb}) \times 100)$ de rendimiento de mutación mientras que en la población del INRA, un 100% $((1/500 \text{ Kb}/1/500 \text{ Kb}) \times 100)$. La eficiencia de mutación en la población del INRA ha sido claramente superior (tres veces más, aproximadamente) a la encontrada en la población del IRTA.

3.3 EcoTILLING

En este trabajo se han analizado mediante EcoTILLING los genes *eIF4E* y *ETR1*. Los genes *PDS* y *eIF(iso)4E* fueron analizados en el COMAV, en el marco del proyecto MELOGEN. Las 113 variedades de melón analizadas corresponden a parte de la colección del banco de germoplasma del COMAV y representan una gran variabilidad a nivel fenotípico ya que son originarias de 39 países distintos y pertenecientes a distintos grupos. Más concretamente encontramos 1 variedad perteneciente al grupo dudaim, 26 melo, 3 conomon, 12 cantalupensis, 44 inodorus, 7 flexuosos, 2 momordica, 1 adzhur-chate, 10 agrestis, 4 chandalak, 1 battik y 1 ameri (Anexo II.k). Se ha observado que existe variabilidad natural entre las 113 variedades analizadas y la línea de referencia M62-113 para estos dos genes.

3.3.1 ANÁLISIS MEDIANTE EcoTILLING EN EL GEN *eIF4E*

Para el análisis de EcoTILLING se usó la misma estrategia usada para el TILLING de modo que se analizaron los dos amplicones del gen *eIF4E* en las 113 accesiones. El primer amplicón incluye el primer exón mientras que en el segundo presenta los cuatro exones restantes que contiene el gen *eIF4E* (Figura 3.14). Al igual que para el método de TILLING, mediante la secuenciación de las variedades (se escogieron dos variedades por cada SNP visualizado en el gel) pudimos determinar la posición de los SNPs y, al mismo tiempo, comprobar si éstos provocaban cambios en la secuencia aminoacídica.

En el primer amplicón se detectaron 7 SNPs (1.1 a 1.7) (Figura 3.28 y Tabla 3.13), de los que sólo el SNP 1.5 provoca un cambio de aminoácido: Asp71Gly. El programa SIFT asigna un valor de 0.00 a este cambio y por lo tanto, determina que puede afectar a la funcionalidad de la proteína. Sólo la accesión PI 505602 contiene este SNP, en heterocigosis.

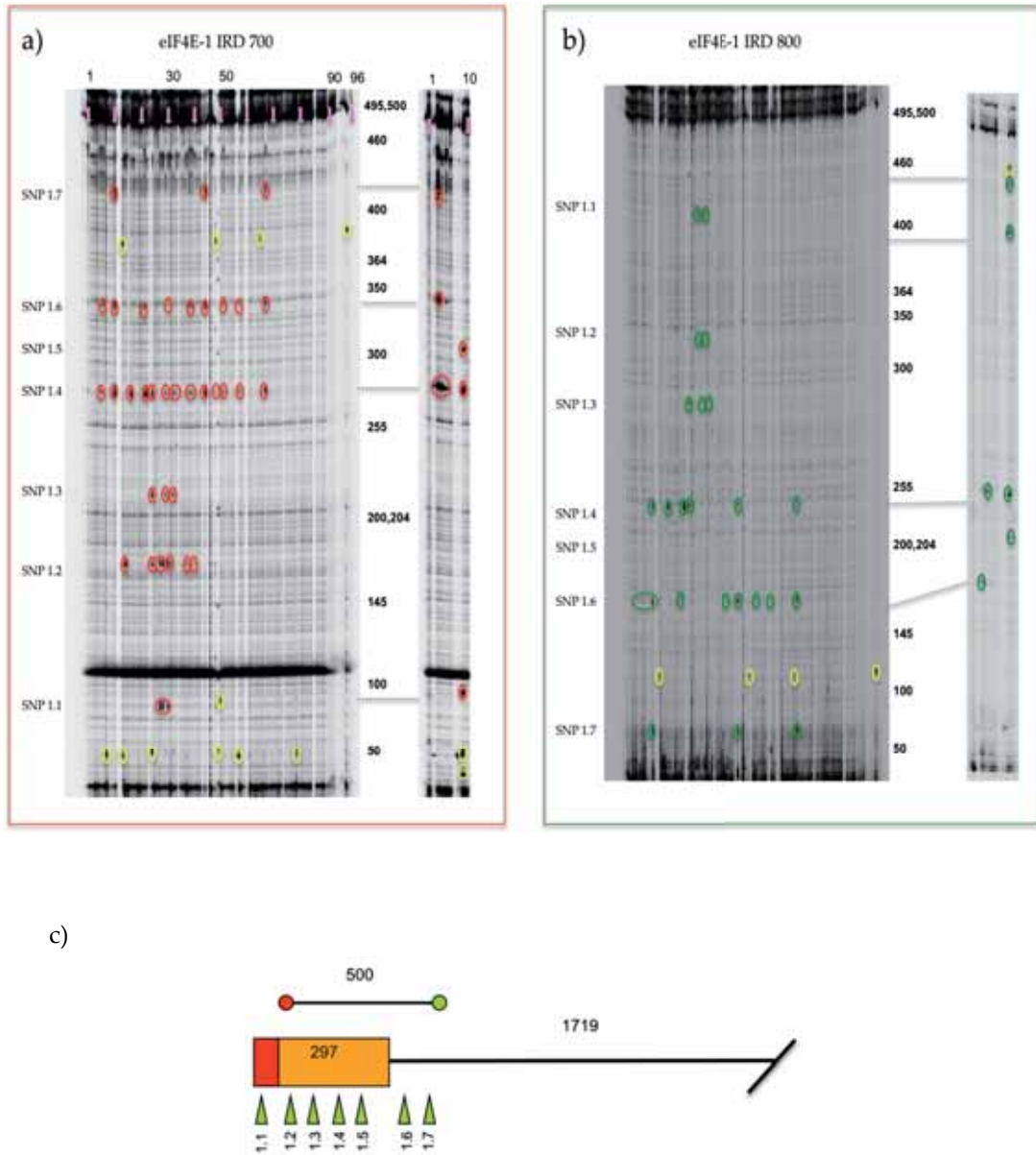


Figura 3.28: Gel de EcoTILLING para el primer amplicón de *eIF4E*. a) canal IRD700, b) canal IRD800. Cada carril corresponde a la mezcla de una de las 113 variedades con la referencia M62-113. En rojo y verde se han detallado los heterodúplex detectados (SNP1.1-1.7). En amarillo se detallan algunos heterodúplex falsos o artefactos descartados mediante secuenciación. Algunos heterodúplex no se visualizaron correctamente en ambos canales sino sólo en uno de ellos, c) localización de los SNPs detectados mediante EcoTILLING en el primer amplicón.

El resto de SNPs detectados en el primer amplicón de *eIF4E* son silenciosos o no están en región exónica. El SNP 1.6 (Figura 3.29) es un polimorfismo complejo, formado por una inserción de dos timinas y un microsatélite. Por lo tanto, el heterodúplex formado en este caso puede ser debido a la inserción de las dos timinas o al microsatélite que está localizado justo a continuación. En la Figura 3.29 puede verse como una de las variedades que contienen este polimorfismo (CUM474) contiene la inserción de dos timinas seguidas del microsatélite (CT)₇, mientras que la variedad de referencia M62-113 no contiene la inserción y el número de repeticiones del microsatélite es (CT)₆.

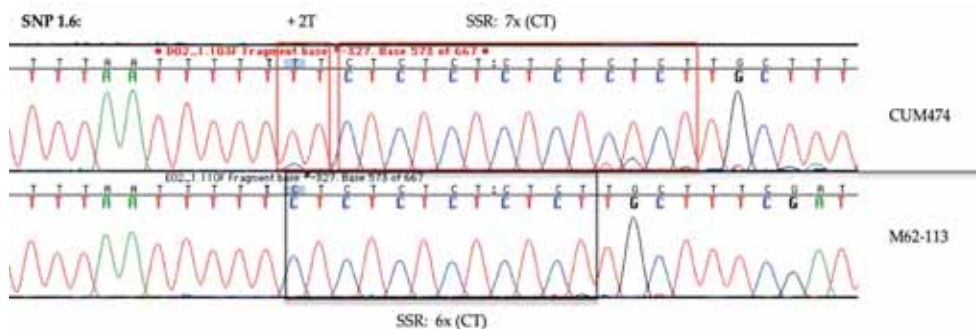


Figura 3.29: Cromatograma de las variedades CUM474 y M62-113 en la región que contiene el SNP 1.6 . El SNP 1.6 está compuesto por una inserción de dos timinas y una repetición adicional CT en CUM474.

Tabla 3.13: Posición, cambio de nucleótido y de aminoácido de cada SNP detectado, variedades que lo contienen y el tamaño del fragmento de digestión según IRD700 y IRD800 para el primer amplión del gen *elF4E*. En sombreado gris, el SNP 1.5 que provoca un cambio no sinónimo en la posición 71 de la secuencia proteica. *PS: Piel de Sapo, Ac: Accesoión.

SNP	Posición	SNP (PS-Ac)*	aa	SIFT	Nombre Variedades	Tamaño fragmento canal IRD 700	Tamaño fragmento canal IRD 800
SNP1.1	5'UTR	C(-42)G	-	-	PI 505602, PI 282448 PI 266947, PI 124112	50 pb	450 pb
SNP1.2	Exón1	T81A	Arg27Arg	-	CUM206, CUM198 PI 436534, PI 124112 PI 266947, PI 282448 PI 618854, ASI-C-10	172 pb	328 pb
SNP1.3	Exón1	C123G	Val41Val	-	PI 185111, PI 436534 CUM 287	214 pb	286 pb
SNP1.4	Exón1	G186T	Leu55Leu	-	CUM294, CUM259 PI 140632, CUM363 CUM225, CUM372 PI 436534, CUM-287 PI 185111, PI 164320 SUD-C-2, AN-C-109 C-550, C-62 CUM416, CUM468 CUM474, CUM500 ns130032, PI 282448	277 pb	223 pb
SNP1.5	Exón1	A212G	Asp71Gly	0.00	PI 505602	303 pb	197 pb
SNP1.6	Intrón 1	TT (CT) ₇ 243 --(CT) ₆	Thr81Thr	-	CUM294, CUM259 PI 140632, CUM225 PI 282448, CUM344 CUM53, SUD-C-2 C-550, C-62 CUM416, CUM474	334 pb	166 pb
SNP1.7	Intrón 1	C328T	-	-	PI 140632, SUD-C-2 AN-C-8, CUM474	419 pb	81 pb

Para el segundo amplicón se detectaron 5 SNPs (Figura 3.30, Tabla 3.14) pero sólo el SNP 2.5 está localizado en región exónica. Sin embargo, el cambio de aminoácido que provoca es sinónimo (Lys230Lys) y por lo tanto no afectaría a la funcionalidad de la proteína. En la Figura 3.30 puede verse el gel con los heterodúplex identificados para este amplicón. Desde el punto de vista fenotípico los haplotipos detectados en este segundo amplicón no tienen ningún interés ya que, o están localizados en región intrónica, o provocan cambios sinónimos.

Observando las distintas combinaciones entre los SNPs identificados en las 113 accesiones se han asignado un total de 19 haplotipos. De las 113 accesiones, 68 no presentan ningún polimorfismo respecto la variedad de referencia y por lo tanto tienen el mismo genotipo en los fragmentos analizados del gen *eIF4E* asignándoles así el haplotipo cero (H0). En cambio, las 45 accesiones restantes presentan SNPs que las diferencian de la variedad M62-113.

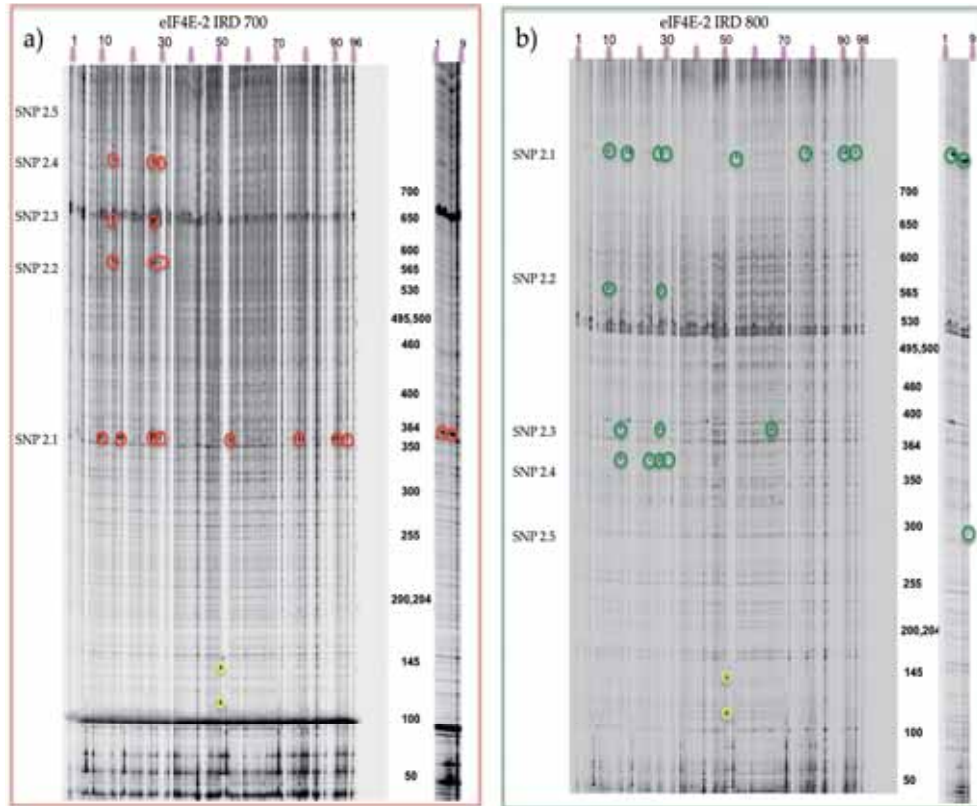
En la Tabla 3.15 se presentan los polimorfismos que presenta cada accesión, las cuales han sido ordenadas según la cantidad de SNPs que contienen, empezando por las que contienen sólo un SNP y continuando por las que contienen combinaciones de dos, tres, cuatro, seis y siete SNPs.

Observando la Tabla 3.15 se puede afirmar que los SNPs que aparecen con más frecuencia son el SNP 1.4 para el primer amplicón y el SNP 2.1 para el segundo, que aparecen en 20/113 y 16/113 accesiones, respectivamente. A la inversa, los menos frecuentes y por lo tanto los más específicos de accesión son los SNPs 1.1, 1.3, 1.5 y el SNP 1.7 para el primer amplicón, apareciendo en tan sólo 4/113, 3/113, 1/113 y 4/113 accesiones, respectivamente. Para el segundo amplicón el SNP 2.5 es el menos frecuente apareciendo únicamente en la accesión PI 505602, que también es la única que contiene el SNP 1.5.

En total, se han detectado 12 polimorfismos respecto a la variedad M62-113 en 45 accesiones de las 113 analizadas, en una región de 1669 pb (500+1169 pb).

El haplotipo más frecuente es el H5 (SNP 2.1), encontrado en 10 accesiones, seguido del H2 (SNP 1.4) encontrado en 7 accesiones. Existen 11 haplotipos que se encuentran sólo en una accesión (H4, H7, H8, H10, H11, H13, H14, H15, H16, H17 y H18). El más relevante es el H13, ya que es el único que contiene los SNPs 1.5 y 2.5 (Tabla 3.15), siendo el 1.5 (Figura 3.31) el único que produce un cambio de aminoácido y tiene un valor 0.00 en SIFT. La accesión que lo contiene es la PI 505602, perteneciente al grupo *agrestis* y proveniente de Zambia. Sin embargo, a

este cambio de aminoácido no se le puede atribuir un fenotipo concreto relacionado con resistencia puesto que no hay información bibliográfica ni experimental con la que se pueda relacionar este genotipo con una resistencia frente algún virus.



c)

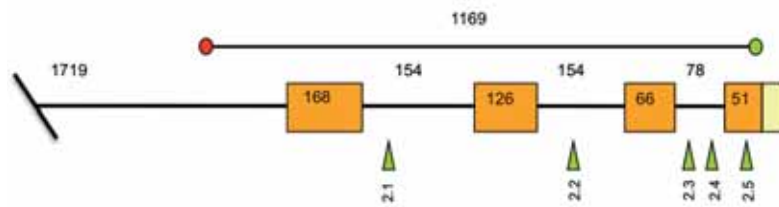


Figura 3.30: Gel de EcoTILLING para el segundo amplicón de *eIF4E*. a) canal IRD700 b) canal IRD800. Cada carril corresponde a la mezcla de una de las 113 variedades con la referencia M62-113. En rojo y verde se han detallado los heterodúplex detectados (SNP 2.1-2.5). En amarillo se detallan algunos

heterodúplex falsos o artefactos descartados mediante secuenciación, c) localización de los SNPs detectados mediante EcoTILLING en el segundo amplicón.

Tabla 3.14: Posición, cambio de nucleótido y de aminoácido de cada SNP detectado, variedades que lo contienen y el tamaño del fragmento de digestión según IRD700 y IRD800 para el segundo amplicón del gen *elF4E*. *PS: Piel de Sapo, Ac: Accesoión.

SNP	Posición	SNP(PS-Ac)*	aa	SIFT	Nombre Variedades	Tamaño fragmento canal IRD 700	Tamaño fragmento canal IRD 800
SNP2.1	Intrón2	2285/+CT	-	-	CUM438, PI 140632 PI 490388, CUM363 CUM338, PI 282448 PI 266947, PI 124112 ASI-C-9, PI 169329 PI 169368, C-241 J.A 8215, CUM500 Asic-6, y285	361 pb	808 pb
SNP2.2	Intrón3	A2515G	-	-	CUM438, CUM198 PI 282448, CUM287 PI 266947, PI 124112 PI 490388, PI 276660	591 pb	578 pb
SNP2.3	Intrón 4	C2717T	-	-	PI 266947, PI 124112 CUM198, PI 167057	793 pb	376 pb
SNP2.4	Intrón4	C2748T	-	-	PI 282448, PI 266947 PI 124112, CUM198 CUM438	824 pb	345 pb
SNP2.5	Exón5	G2788A	Lys230Lys	-	PI 505602	864 pb	305 pb

Tabla 3.15: Haplotipos encontrados mediante el análisis por EcoTILLING de los dos amplicones de eIF4E. En sombreado gris se ha detallado el SNP 1.5 que provoca un cambio de aminoácido en la accesión PI505602.

Nombre variedad	SNP1.1	1.2	1.3	1.4	1.5	1.6	1.7	2.1	2.2	2.3	2.4	2.5	Haplotipo: SNPs
CUM206		x											H1:1.2
PI 618854		x											
ASI-C-10		x											
CUM363				x									H2: 1.4
CUM225				x									
CUM372				x									
PI 164320				x									
AN-C-109				x									
CUM468				x									
Ns130032				x									
CUM225						x							H3: 1.6
CUM344						x							
CUM53						x							
AN-C8							x						H4:1.7
PI 140632								x					H5: 2.1
CUM363								x					
CUM338								x					
ASI-C-9								x					
PI 169329								x					
PI 169368								x					
C-241								x					
J.A 8215								x					
Asic-6								x					
Y285								x					
PI 490388									x				H6:2.2
PI 276660									x				
PI 167057										x			H7:2.3
PI 185111			x	x									H8: 1.3+1.4
CUM294				x		x							H9: 1.4+1.6
CUM259				x		x							
C-550				x		x							

C-62				x		x								
CUM416				x		x								
CUM500				x				x						H10: 1.4+2.1
PI 436534		x	x	x										H11: 1.2+1.3+1.4
SUD-C-2				x		x	x							H12: 1.4+1.6+1.7
CUM474				x		x	x							
PI 505602	x					x								H13:1.1+1.5+2.5
CUM287			x	x					x					H14: 1.3+1.4+2.2
CUM438								x	x		x			H15: 2.1+2.2+2.4
CUM198		x							x	x	x			H16: 1.2+2.2+2.3+2.4
PI 140632				x		x	x	x						H17: 1.4+1.6+1.7+2.1
PI 266947	x	x						x	x	x	x			H19: 1.1+1.2+2.1+2.2+2.3 +2.4
PI 124112	x	x						x	x	x	x			
PI 282448	x	x		x		x		x	x		x			H18:1.1+1.2+1.4+1. 6+2.1+2.2+2.4

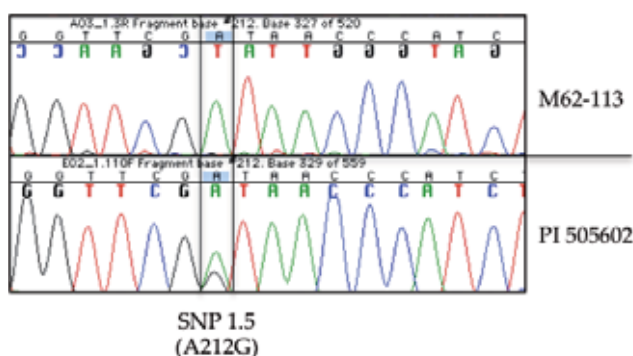
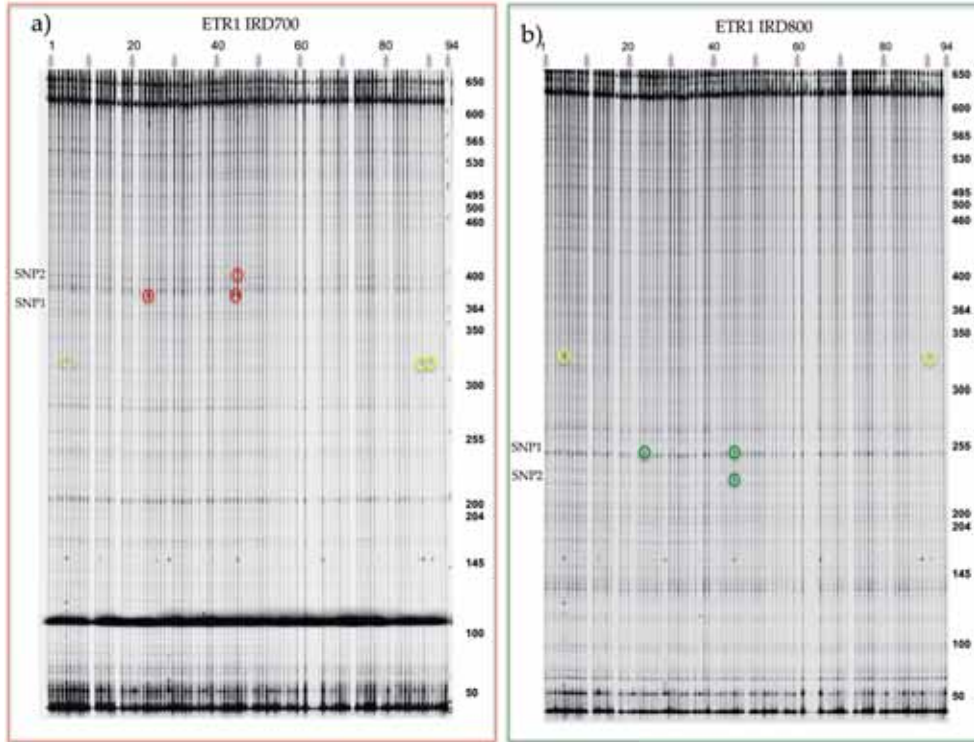


Figura 3.31: Cromatogramas de la variedad de referencia M62-113 y la accesión PI 505602, que contiene el SNP 1.5 responsable de un cambio de aminoácido (ácido aspártico a glicina).

3.3.2 ANÁLISIS MEDIANTE EcoTILLING EN EL GEN *ETR1*

El amplicón rastreado mediante EcoTILLING corresponde al primer exón de *ETR1*, de 615 bp (Figura 3.26). Se han detectado dos polimorfismos entre la variedad M62-113 y las 113 accesiones analizadas. En la Figura 3.32 se puede observar el gel con la posición de los dos SNPs detectados. Para cada gen analizado mediante

EcoTILLING tuvieron que realizarse dos geles dado que el número de variedades a analizar era superior al número de muestras que pueden cargarse en un gel de TILLING. Este es el motivo por el cual en la Figura 3.32 no están representadas todas las variedades y por lo tanto no están detallados todos los heterodúplex.



c)

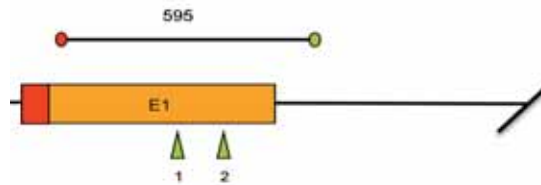


Figura 3.32: Gel de EcoTILLING para el amplicón del gen *ETR1*. a) canal IRD700, b) canal IRD800. Cada carril corresponde a la mezcla de una de las 113 variedades (aunque en cada gel sólo se cargaban 96/113) con la referencia M62-113. En rojo y verde se han detallado los heterodúplex detectados (SNP 1 y 2). En amarillo se detallan algunos heterodúplex falsos o artefactos descartados mediante secuenciación, c) localización de los SNPs detectados mediante EcoTILLING en el primer exón del gen *ETR1*.

El SNP 1 (Tabla 3.16, Figura 3.33) provoca un cambio de aminoácido de leucina a valina aunque según los valores obtenidos mediante SIFT (valor 1.00) es poco probable que afecte a la funcionalidad de la proteína. Las variedades que contienen el SNP 1 son PI 124112, PI 482429, CUM349, PI 436534 y PI 282448.

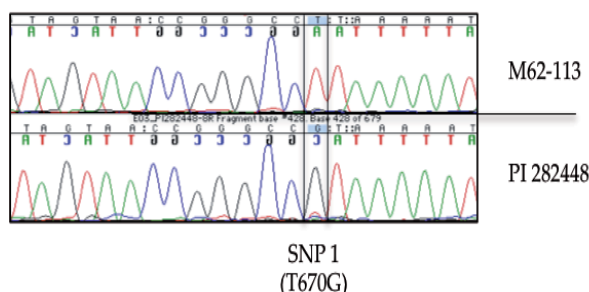


Figura 3.33: Cromatograma del SNP 1 del primer exón del gen *ETR1*, que provoca un cambio de aminoácido de leucina a valina en la variedad PI 282448.

Todas estas variedades pertenecen a grupos distintos (*momordica*, *agrestis*, *melo* y *flexuosus*) y son de orígenes geográficos distintos (India, Zimbawe, Irak, Senegal y Sudáfrica).

El SNP 2 no provoca ningún cambio de aminoácido. Se puede destacar que éste provoca un cambio en la variedad PI 124112 de A693G y, sin embargo, en las variedades PI 282448 y PI 482429 el cambio es de A693C.

Mediante la combinación de los dos SNPs se han identificado dos haplotipos distintos (Tabla 3.17). Según los resultados obtenidos se puede observar como el primer exón del gen *ETR1* presenta poca variabilidad puesto que sólo 5 accesiones de las 113 analizadas difieren en alguno de los dos polimorfismos respecto a la variedad estándar.

Tabla 3.16: Polimorfismos detectados en el primer exón del gen *ETR1* entre la variedad M62-113 y las

SNP	Posición	SNP (PS-Ac)*	aa	SIFT	Nombre Variedad	Tamaño fragmento canal IRD 700	Tamaño fragmento canal IRD 800
SNP1	Exón1	T670G	Leu224Val	1.00	PI 124112, PI 482429 CUM349, PI 436534 PI 282448	378 pb	272 pb
SNP2	Exón1	A693G A693C	Pro231Pro	-	PI 124112 (G), PI 282448 (C) PI 482429 (C)	401 pb	249 pb

113 accesiones analizadas. Se incluyen también los datos de las variedades que contienen cada SNP así como el tamaño del fragmento de digestión visualizado en ambos canales (IRD700 y IRD800). *PS: Piel de Sapo, Ac: Accesoión.

Tabla 3.17: Tabla con los haplotipos detectados para el gen *ETR1*. En sombreado gris se detalla el SNP 1 que provoca un cambio de aminoácido (leucina a valina) en la posición 224 de la secuencia proteica.

Nombre variedad	SNP1	SNP2	Haplotipo/SNP
CUM349	x		H1: 1
PI 436534	x		
PI 124112	x	x	H2:1+2
PI 482429	x	x	
PI 282448	x	x	

3.4 CATÁLOGO DE MUTANTES FENOTÍPICOS

Durante el segundo año de tesis, se obtuvieron las semillas M2 de las familias que representarían la población de TILLING. Conjuntamente con la empresa *Semillas Fitó* se decidió observar la población M2 a nivel fenotípico. Las poblaciones de TILLING se crean mayormente para aplicaciones de genética inversa. Así, se inicia la detección de mutantes en un gen de interés y posteriormente se analiza si el genotipo detectado es el responsable de un fenotipo observado. Pero también existe la posibilidad de evaluar fenotípicamente toda la población, o parte de ella, para detectar fenotipos interesantes y posteriormente caracterizarlos con fines comerciales o por su interés biológico, a pesar de no conocer la mutación responsable de dicho fenotipo.

El método óptimo es desarrollar una base de datos donde almacenar los fenotipos observados de un modo ordenado para poder hacer búsquedas exhaustivas en función de la morfología o el tipo de efecto que produce la mutación desconocida. Otro modo de usar los datos fenotípicos es elaborando una página web donde el usuario pueda elegir un mutante fenotípico que le interese. Un buen ejemplo es UTILLdb <http://urgv.evry.inra.fr/UTILLdb> donde se pueden solicitar semillas mutantes de guisante (*Pisum sativum*) y de tomate (*Solanum lycopersicum*). Se puede escoger el mutante en función únicamente de un fenotipo normalmente morfológico, en un gen de interés, o haciendo *Blast* con el fragmento de secuencia que nos pueda interesar en la base de datos. En este trabajo no se ha elaborado una base de datos pero si un catálogo fenotípico en función de las distintas etapas de crecimiento de la planta. Para los fenotipos identificados, se ha hecho una descripción de la mutación conjuntamente con la identificación de la familia que la contiene y la imagen del fenotipo.

Para la empresa *Semillas Fitó*, la parte más interesante se centraba en las características del fruto y en tratar de encontrar fenotipos interesantes a nivel comercial. Para ello, en el año 2007 se plantaron 400 familias M2, concretamente 10 individuos para cada familia (un total de 4000 plantas aproximadamente). Plantando diez individuos por familia se pretendía poder diferenciar los fenotipos debidos a efectos ambientales, de los efectos producidos por alguna mutación

inducida por EMS. También, en todos los estadios estuvieron presentes plantas control (plantas de la variedad M62-113) para poder comparar si el fenotipo observado era debido a una mutación o al ambiente

Desde la germinación de las semillas M2 se hizo un seguimiento exhaustivo de las plantas observando si aparecían fenotipos mutantes. En función del estadio de la planta se observaron distintos caracteres (plántula, planta adulta y fruto). Debe advertirse que en algunas de las familias presentadas en el catálogo, sólo uno de los diez individuos presentaba el fenotipo mutante. En estos casos deberían plantarse más individuos M2 para poder asegurar que el efecto fenotípico observado fue provocado por una mutación y no como consecuencia de las condiciones ambientales. El trabajo de crecimiento y mantenimiento de las plantas fue realizado por la empresa *Semillas Fitó*.

3.4.1 MUTANTES EN ESTADIO DE PLÁNTULA

A nivel de plántula y antes de trasplantar al invernadero se hizo un seguimiento detallado de características tales como el número de semillas germinadas por familia y el número de plantas viables por familia. Contabilizando estos resultados se obtuvieron los porcentajes presentados en la Tabla 3.18.

Tabla 3.18: Porcentajes de germinación en las 400 familias M2 observadas.

Familias no germinadas	Familias con baja germinación (menos de 4-6 /10)	Familias germinadas pero con letalidad
1,25%	11,80%	3,25%

Se observaron varios fenotipos anormales en los que las plántulas, sin llegar a ser adultas, presentaban malformaciones, desarrollo anormal en el crecimiento, clorosis u otros tipos de alteraciones que se pueden ver detalladas en la Tabla 3.19 y en las imágenes correspondientes (3.4.4 Catálogo de imágenes).

Tabla 3.19 Fenotipos encontrados en estadio de plántula.*Catálogo de imágenes 3.4.4. ** Algunas plantas han mostrado el fenotipo mutante en diferentes estadios de crecimiento, por lo tanto, las encontraremos tanto en el catálogo realizado para el estadio de plántula como en el de planta adulta.

Fenotipo	Familia	Plantas con fenotipo mutante/Familia	Imagen*
Gravitropismo. Desarrollo anormal	F22	3/10	1
	F55	5/10	2
	F798	1/10	7
	F1326	5/9	10
	F1505	1/2	11
	F1525	2/9	12
	F1646	3/10	14
	F1790	8/9	15
	F1876	3/7	16
	F1902	1/8	9
	F2193	1/9	18
	F2315	4/6	19
Manchas cloróticas. Clorosis parcial	F55	5/10	2
	F58**	2/10	3
	F81	7/9	5
	F137	7/10	6
	F798	1/10	7
	F979	4/10	8
	F1525	2/9	12
	F1550	1/2	13
	F1326	5/9	10
	F1902	1/8	9
	F1934	5/6	17
F2315	4/6	19	
Adhesión de las primeras hojas	F74	6/10	4
	F81	7/10	5

3.4.2 MUTANTES EN ESTADIO DE PLANTA ADULTA

En el estadio de planta adulta se observaron cambios en el desarrollo, así como cambios en la pigmentación, cambios en la forma de la hoja o en la flor, enanismo, y otros fenotipos (Tabla 3.20). Como el cribado era aleatorio no esperábamos un

fenotipo concreto, sino descubrir nuevos fenotipos, quizás no observados anteriormente en la especie.

Tabla 3.20 Fenotipos encontrados en estadio de planta adulta. *Catálogo de imágenes 3.4.4.

Fenotipo	Familia	Plantas con fenotipo mutante/Familia	Imagen*
Gravitropismo. Desarrollo anormal	F985	5/10	21
	F1086	1/9	22
	F2261	2/9	28
	F2284	2/6	29
	F2317	4/9	31
Deformaciones en hojas	F1086	1/9	22
	F1902	3/8	27
Clorosis	F58	2/10	20
	F985	5/10	21
	F1096	5/10	23
Esterilidad o malformación en flor	F1455	2/9	24
	F1902	3/8	27
Aspecto de infección viral	F1455	2/9	24
	F1550	2/6	25
	F1632	4/7	26
Letalidad	F2315	2/10	30

3.4.3 MUTANTES EN FRUTO

Se dedicó un mayor esfuerzo en el fenotipado del fruto, ya que el objetivo de *Semillas Fitó* era encontrar algún mutante que proporcionara uno o varios caracteres atractivos desde el punto de vista comercial. Los mejoradores en este caso buscaban caracteres que cambiaran el aspecto típico del melón “Piel de sapo” y que resultaran a su vez atractivos para el consumidor. Una forma más redondeada de fruto, o una piel sin manchas y poco escriturada podrían ser algunos de los caracteres que podrían ser de interés.

Se recogieron los frutos de aproximadamente 370 familias. Para cada una de ellas se crecieron diez plantas y para cada planta se recogió un melón. Se compararon los parámetros observados para todos los melones de la misma familia de modo que se seleccionaban aquellos que presentaran un rasgo diferencial respecto los demás. Todos los parámetros se introdujeron en una base de datos desarrollada por los

mejoradores de *Semillas Fitó* en la que se indicaba para cada melón, el vigor, la resistencia a oídio, el peso, el color exterior, el escriturado, la rugosidad, el cierre, la cavidad, el color de la carne, el brix (cantidad en azúcares), la dureza, el avinado, el sabor y las medidas de forma. Aquellos frutos que presentaban un carácter diferenciado y que pudiera resultar interesante para la mejora se seleccionaban y se almacenaban sus semillas. Estos resultados no se muestran en este trabajo ya que son datos de la empresa.

En la Figura 3.34 se muestra un ejemplo de cómo se comparaban todos los frutos de las plantas de una misma familia. Una vez analizados todos los caracteres se separaba aquél o aquellos frutos que se salían de los parámetros estándares, para producir más semilla e intentar fijar el fenotipo. En este ejemplo, se seleccionó el primero por la izquierda por ser más redondo que el resto y por lo tanto, un fenotipo interesante para poder fijar este carácter en el melón tipo “Piel de sapo”.



Figura 3.34: Procedimiento para la selección de melones con caracteres fenotípicos interesantes a nivel comercial. En este ejemplo, en la imagen superior vemos los melones recogidos de esta familia (2286) y

en la imagen inferior el melón seleccionado por su acentuada forma redonda en comparación a la forma elíptica típica del melón "Piel de sapo".

Se observaron las diferencias fenotípicas en los frutos, al igual que para los otros dos estadios, para poder catalogar algunos fenotipos mutantes. Las distintas mutaciones fenotípicas así como las familias M2 correspondientes se detallan en la Tabla 3.21.

Tabla 3.21: Fenotipos estudiados en frutos.*Catálogo de imágenes 3.4.4.

Fenotipo	Familia	Número de frutos con el fenotipo mutante/Familia	Imagen*
Variaciones en color/escriturado	F1094	2	32
	F1149	2	36
	F1505	1	38
	F1902	1	42
Deformaciones en cierre y malformación	1099	1	33
	F1344	1	37
	F1538	1	40
	F1753	1	41
	F1902	1	42
Variación en la forma	F1105	2	34
	F1133	1	35
	F1149	2	36
	F1506	1	39
	F1902	1	42
	F2286	1	43

3.4.4 CATÁLOGO DE IMÁGENES DE LOS MUTANTES

En el siguiente apartado se han reunido todas las imágenes correspondientes a los fenotipos detectados en los estadios de plántula, de planta adulta y de fruto.

-Mutantes en estadio de plántula: Imágenes de la 1 a la 19.

-Mutantes en planta adulta: Imágenes de la 20 a la 31.

-Mutantes en fruto: Imágenes de la 32 a la 43.

Estadio de plántula

1



F22: Desarrollo anormal, gravitropismo.

2



F55: Cotiledones deformados y pequeñas manchas cloróticas

3



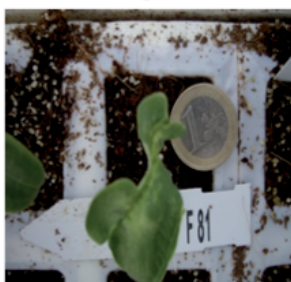
F58: Manchas cloróticas en el centro de las primeras hojas.

4



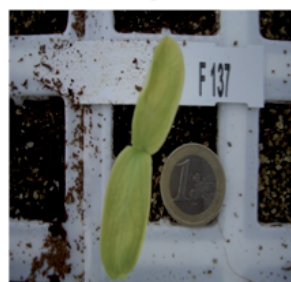
F74: Adhesión de las primeras hojas

5



F81: Adhesión de las primeras hojas y manchas cloróticas

6



F137: Cotiledones de color más claro.

7



F798: Desarrollo de las hojas pero no del tallo. Clorótica

8



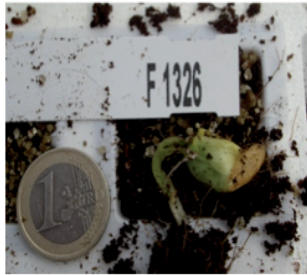
F979: Desarrollo alterado. Color verde claro.

9



F1902: Clorótica parcial

10



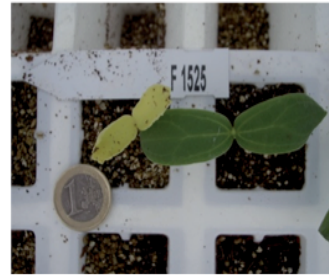
F1326: No se acaban de desarrollar los cotiledones, pigmentación alterada y la semilla no se desprende

11



F1505: Desarrollo totalmente alterado. Ni las hojas ni la estructura de planta pertenece a la típica del melón. Hojas puntiagudas y estructura similar a un cactus

12



F1525: Cotiledones más pequeños y claros.

13



F1550: Cotiledones con los nervios decolorados

14



F1646: Cotiledones enanos.

15



F1790: Familia con crecimiento decantado. Los cotiledones se orientan todos a un lado.

16



F1876: Cotiledones deformados.

17



F1934: Familia enana y ligeramente más clara.

18



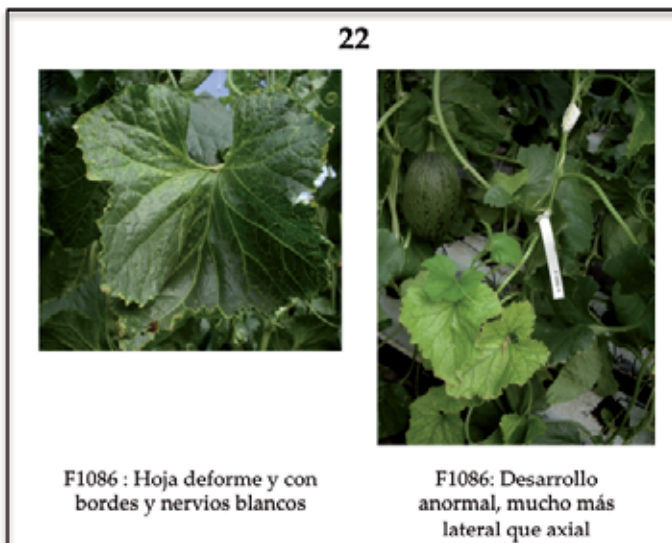
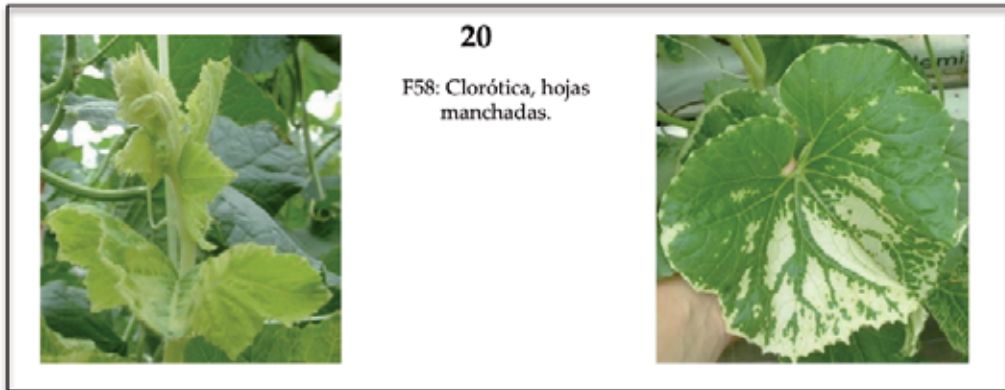
F2193: No desprende el caparazón de la semilla. Crecimiento anormal.

19



F2315: Planta enana, poco desarrollo de las hojas. Los nervios de las hojas son blancos y el color general es más claro.

Estadio de planta adulta



24



F1455: Fenotipo de planta que parece infectada por un virus, con hojas cloróticas en los bordes y en los nervios. Desarrollo anormal con menos formación foliar y la no producción de flor. Estéril.

25



F1550: Fenotipo de planta que parece infectada por un virus y estéril.



26

F1632: Fenotipo de planta que parece infectada por un virus y planta estéril.



27



F1902 (planta 5, 7 y 8):
Familia con malformación en las hojas y en la estructura foliar. Las hojas no crecen independientemente y forman un complejo de hojas. En el momento de floración también se pueden observar malformaciones, tanto en el número de pétalos (normalmente cinco y en este caso se presentan seis) como en el desarrollo de la flor.



28



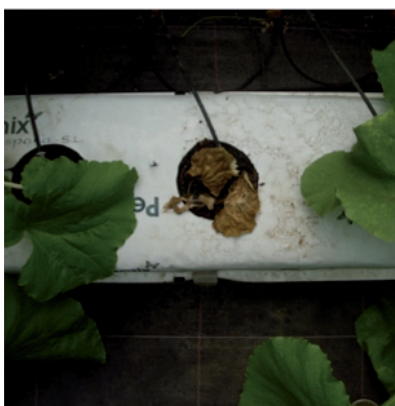
F2261: Toda la planta crece a partir del mismo nudo y no crece en el plano axial.

29



F2284: Enanas. En la floración todavía no han alcanzado ni un tercio de la altura del resto de la familia.

30



F2315: Letalidad de dos de los individuos de la familia, con aspecto de sequedad en la planta.

31



F2317: Enanismo y desarrollo extraño. Aspecto puntiagudo de las hojas y mucho más grosor. Aspecto de cactus

Estadio de Fruto

32



F1094: Color verde muy claro.

33



F1099 (planta 8 y 9): Malformación del fruto. Rotura del exocarpio.

34



F1105: Forma más elongada. Característico por el color liso sin manchas típicas del "Piel de sapo" ni escriturado.

35



F1133: Familia con forma de fruto más redondeada respecto a la forma más elíptica del "Piel de Sapo".

36



F1149: Forma elongada. Color claro, liso, sin manchas ni escriturado.

37



F1344: Deformación provocando una forma elongada y curvada. Con cierre muy abierto

38



F1505: Color exterior característico liso, sin manchas.

39



F1506: Forma muy alargada.

40



F1538: Cierre con deformación y claramente muy abierto

41



F1753: Deformación exterior.

42



F1902: Forma elongada y color claro y liso.

43



F2286: Forma redondeada

4. DISCUSIÓN

4.1 POBLACIÓN MUTAGENIZADA DE MELÓN Y PUESTA A PUNTO DEL MÉTODO TILLING

La eficiencia de mutación de la población de TILLING obtenida en este trabajo ha sido de 1 mutante/1,5 Mb. Ésta es una eficiencia menor que aquellas descritas en otras especies diploides como *Arabidopsis thaliana* (1/300 Kb) (Greene y col., 2003), maíz (1/500 Kb) (Till y col., 2004) o cebada (1/Mb) (Caldwell y col., 2004). En especies poliploides como el trigo tetraploide o el trigo hexaploide los valores todavía son más elevados 1/40 Kb y 1/24 Kb respectivamente (Slade y col., 2005).

En el momento de iniciar este trabajo no existía ninguna población de mutantes desarrollada en melón. El protocolo de mutagénesis de las semillas fue adaptado a partir de otras especies y aunque se hizo un ensayo previo para determinar la dosis óptima de EMS, los resultados obtenidos nos indican que la dosis usada no fue probablemente suficiente. En nuestro caso las concentraciones usadas de EMS estuvieron en el rango de 0,5% hasta 1,5%. En cambio, para el desarrollo de la población de TILLING del INRA, partieron con la ventaja de conocer los resultados de nuestro trabajo, y aplicaron concentraciones más altas de EMS (del 1% al 2%). Por otro lado, durante el protocolo de mutación de las semillas se realizó una hidratación de las mismas previa a la aplicación del mutágeno. Existe un proceso denominado *osmopriming* en el que la prehidratación de semillas desecadas proporciona una germinación muy homogénea dado que activa los sistemas de reparación del ADN (Ashraf y Bray, 1993). Es por ello que posiblemente la prehidratación realizada en las semillas pudo activar los sistemas de reparación del ADN y quizás, provocar la corrección de algunas mutaciones producidas por el EMS y en consecuencia bajar la eficiencia de la mutagénesis.

Teniendo en cuenta el efecto del mutágeno usado (EMS) el porcentaje de GCs presente en un gen es una de las variables que puede condicionar la eficiencia de mutación. Se han comparado los porcentajes en GCs en las zonas codificantes de los genes analizados. Para el gen *eIF4E* el contenido de GCs es de un 36,7 %, y su isoforma *eIF(iso)4E* tiene un contenido ligeramente superior (37,41 %). El gen de la

PDS tiene un contenido de GCs del 34,93 % (Tabla 4.1). Teniendo en cuenta el efecto del EMS esperábamos poder establecer una relación directamente proporcional entre el porcentaje de GCs y el número de mutantes detectados en cada gen, pero no ha sido así. El gen *elF(iso)4E* es el que tiene mayor porcentaje de bases susceptibles a ser mutadas y sin embargo, para este gen no hemos encontrado ningún mutante en la población IRTA y tan sólo uno en la población INRA. En cambio, para el gen *PDS* el porcentaje de GCs es menor y se han detectado dos mutantes en la población IRTA y cuatro en la del INRA. Por lo tanto, en este trabajo parece que la proporción de GCs del gen no se ha correlacionado con la obtención de mutantes (Tabla 4.1). Hay dos razones que podrían explicar este resultado. La primera, como ya se ha comentado anteriormente, es que no se ha analizado un número suficiente de genes para poder establecer una tasa de mutación más significativa a nivel estadístico. La segunda, es que debe tenerse en cuenta la importancia del gen en cuanto a la funcionalidad que realiza. Si un gen tiene una función vital para la especie, entonces la probabilidad de encontrar mutantes homocigotos en ese gen disminuye considerablemente puesto que el individuo con dicha mutación no sería viable, muriendo antes de la posible detección de la mutación.

Para poder desarrollar una población aún más efectiva en cuanto a tasa de mutación sería conveniente calcular las plantas necesarias para asegurar la obtención de un mutante para cada nucleótido analizado. Según Dalmais y col., (2007) conociendo la frecuencia de mutación obtenida en una población podemos estimar el número de plantas necesarias para conseguir una mutación deseada con una elevada probabilidad. Por consiguiente se podría aumentar el tamaño de la población hasta obtener un mutante por cada base analizada (Dalmais y col., 2007). Este sería un modo de aumentar la eficiencia de mutagénesis de la población, aunque el proceso de crecimiento de un número tan elevado de plantas sería muy largo y costoso en una especie como el melón.

A partir de los resultados de este trabajo, en MELONOMICS (proyecto 2010-2012 donde, entre otros resultados, se ha secuenciado el genoma del melón) se incluyó como objetivo la obtención de otra población de TILLING. La empresa *Semillas Fitó* juntamente con el IRTA-CRAG han desarrollado una nueva población de TILLING a partir de una línea pura de melón tipo "Piel de Sapo". En esta ocasión se han mutado las semillas con un porcentaje de 1,5% de EMS y se han obtenido aproximadamente 3100 individuos M2 en el momento de escribir esta tesis, en los

cuales, ya se han detectado 8 nuevos mutantes en el gen *eIF4E* y 3 en el gen *eIF(iso)4E* (estos resultados serán comentados con más detalle posteriormente).

Tabla 4.1: Porcentaje de bases susceptibles al mutágeno EMS, o contenido en GCs, en cada gen analizado, y la comparación del número de mutantes encontrados en cada población. *Debe tenerse en cuenta que la variedad utilizada para desarrollar la población del INRA (Charentais) era distinta a la utilizada para la población del IRTA (Piel de Sapo) por lo tanto, los porcentajes podrían variar ligeramente, ya que la secuencia utilizada para los cálculos es la de la variedad M62-113.

Gen	% GCs*	Mutantes IRTA	Mutantes INRA
<i>eIF4E</i>	36,7 %	4	14
<i>eIF(iso)4E</i>	37,41 %	0	1
<i>PDS</i>	34,93 %	2	4

En otras poblaciones se ha determinado la eficiencia de mutación mediante el análisis en un número de genes más elevado. Por ejemplo, para determinar la tasa de mutación obtenida en la población de melón desarrollada en el INRA se analizaron 11 genes (Dahmani-Mardas y col., 2010), de este modo la tasa de mutagénesis calculada es más fiable.

En este trabajo se han analizado cuatro genes en dos poblaciones distintas. Esto ha permitido que podamos establecer algunas comparaciones entre ambas. Según la Tabla 4.1 observamos que para el gen *eIF4E* se han detectado 4 mutantes en la población IRTA mientras que en la población INRA se han detectado 14, lo que representa una diferencia notable entre ambas poblaciones. El mismo caso se da en el gen *PDS* puesto que en la población del IRTA se han detectado 2 mutantes mientras que 4 en la del INRA.

Dado que se conocen los detalles del protocolo empleado para la mutagénesis y desarrollo de ambas poblaciones, podemos afirmar que la tasa de mutación obtenida en cada población ha sido condicionada básicamente por la eficiencia con que fue mutada, es decir, la dosis de mutágeno utilizada y el protocolo de mutagénesis empleado.

En la Tabla 4.2 se resumen todos los mutantes encontrados en este trabajo, tanto en la población IRTA como en la del INRA, así como toda la información detallada de cada uno de ellos. En la columna donde se detallan las frecuencias de mutación se evidencia que en todos los genes la frecuencia de mutación es mayor en la población INRA (1/395 Kb en INRA respecto 1/1,8 Mb en IRTA en el primer amplicón del gen *eIF4E*, 1/264 Kb respecto 1/923 Kb en el segundo amplicón del gen *eIF4E*, 1/2,52 Mb respecto a ningún mutante en el gen *eIF(iso)4E* y 1/737 Kb respecto 1/1,41 Mb en la *PDS*). Al mismo tiempo, a pesar de que en la población INRA la frecuencia de mutagénesis es mayor, en algunos casos como el del gen *eIF(iso)4E* sigue siendo muy baja. Este dato corrobora la hipótesis de que existen algunos genes que, por motivos todavía desconocidos, resultan más difíciles de mutagenizar.

Tabla 4.2: Tabla resumen de los mutantes encontrados en cada población y para cada gen, con detalle del tipo de cambio de base, el efecto de la mutación a nivel aminoacídico, el nombre de la familia, la segregación M2 para aquellos mutantes que se han podido estudiar, la frecuencia de mutación, el fenotipo en caso de haber sido estudiado, y finalmente el valor SIFT. (* Datos explicados en el apartado 4.2.2, Tabla 4.3 **Datos explicados en el apartado 3.2.1.3, Tabla 3.7.

Gen /Población	Tamaño amplificación (pb)	Número mutantes	Exón/ Intron	Posición	Cambio AA	M2 familia	M2 segregación (WT/H/M)	Frecuencia mutación	Fenotipo	Valor SIFT
<i>eIF4E</i> IRTA	500	1	Exón 1	G74A	Arg25Lys	F2036	09/07/08*	1/1,8Mb	-	0.00
	1169	3	Exón 2	G385A	Glu129Lys	F57	05/03/00*	1/923Kb	-	0.00
			Exón 2	G465A	Thr155Thr	F1880	-		-	-
			Intrón 2	G/C		F959	-		-	-
<i>eIF4E</i> INRA	500	3	Exón 1	G289A	Glu97Lys	53	-	1/395 Kb	-	0.02
			Exón 1	C294T	Phe98Phe	2532	-		-	-
			Intrón 1	G/A		1420	-		-	-
	1169	11	Exón 2	C401T	Ala134Val	1229	-	1/264 Kb	-	0.10
			Exón 2	G407A	Gly136Ala	444	-		-	0.00
			Intrón 2	C/A		1287	-		-	-
			Intrón 2	C/T		1115	-		-	-
			Intrón 3	G/A primer a base del intrón	<i>splicing</i>	1667	-		-	-
			Intrón 3	C/T		2646	-		-	-
			Intrón 3	C/T		1233	-		-	-
			Exón 4	G616A	Glu206Lys	2060	-		-	0.09
			Intrón 4	C/T		2035	-		-	-
			3' UTR	T/C		1368	-		-	-
			3' UTR	G/A		2163	-		-	-
<i>eIF(iso)</i> 4E IRTA	1014	0	-	-	-	-	-	0	-	-
<i>eIF(iso)</i> 4E INRA	1014	1	Exón 1	G128A	Arg43Lys	M1587	-	1/2.52Mb	-	1.00
<i>PDS</i> IRTA	1188	2	Intrón 1	C/T		F1151	04/07/04**	1/1.41Mb	-	-
			Exón 2	C560T	Ser540Phe	C384	06/09/00**		Planta albina en M3	0.00
<i>PDS</i>	1188	4	Exón 1	C395T	Ser485Phe	1182	-	1/737Kb	-	0.01

INRA			Intrón 1	C/T	-	2495	-	-	-	-
			Intrón 1	T/C	-	1460	-	-	-	-
			Intrón 1	C/T	-	1226	-	-	-	-

Una explicación para los resultados obtenidos en el gen *eIF(iso)4E*, podría ser el motivo ya comentado anteriormente, que la mutación provocara individuos inviables y por lo tanto nunca pudieran detectarse porque no llegan a desarrollarse. La letalidad o la falta de mutantes puede ser provocada tanto por el efecto de la mutación en el gen analizado como por el efecto de otras mutaciones que contiene el mismo mutante repartidas aleatoriamente por todo el genoma. Fijándonos en la frecuencia de la población más eficiente (INRA) este gen presenta una frecuencia de 1 mutante cada 2,52 Mb, claramente muy baja si la comparamos con la del gen *PDS* (1/737 Kb) o con la del segundo amplicón del gen *eIF4E* (1/264 Kb). Una de las funciones del gen *eIF(iso)4E* es la de ayudar a estabilizar el complejo proteico para la iniciación de la traducción de las proteínas (Browning y col., 1992), función básica para la célula. Por este motivo podríamos pensar que mutantes con un cambio que no permitiera la estabilidad de la proteína *eIF(iso)4E* en el complejo de iniciación de la traducción y por lo tanto, no permitiera la traducción de ciertas proteínas, haría que no fueran viables.

La eficiencia de mutación obtenida en una población de TILLING de tomate donde se analizaron 18 genes, uno de ellos el gen *eIF4E*, fue de 1/574 Kb (Piron y col., 2010). Sin embargo, para el gen *eIF(iso)4E* la eficiencia de mutación fue de 1/1110 Kb. Este dato coincide con los resultados obtenidos en este trabajo y demuestra como para ciertos genes la eficiencia de mutación puede ser más baja respecto la eficiencia media en la población, probablemente debido a que su función es más vital para el organismo que otros genes con una tasa de mutación mucho más alta.

4.2 EFECTOS DE LAS MUTACIONES IDENTIFICADAS EN LOS GENES ANALIZADOS

4.2.1 PDS (FITOENO DESATURASA)

Poder relacionar el fenotipo de un mutante con una mutación concreta permite confirmar la función del gen. Para ello, hace falta caracterizar bien los mutantes ya sea a nivel morfológico, metabólico, bioquímico o mediante otros análisis que caractericen con exactitud el efecto del gen mutado.

El gen *PDS* realiza una función determinante para la biosíntesis de carotenoides y en consecuencia, para que se acumulen pigmentos en la planta. La no funcionalidad o el silenciamiento del gen *PDS* provoca fenotipos albinos, como se ha demostrado en otras especies como *Arabidopsis* o arroz (Carol y Kuntz, 2001; Warthmann, 2008). Por ello, se escogió el gen *PDS* como control fenotípico de la población puesto que la detección de mutantes para este gen no requería de ensayos complejos. Encontrar una planta albina al mutar *PDS* demostraría no sólo la funcionalidad del gen, sino la eficiencia de la población de TILLING y el método usado para detectar los mutantes.

En la puesta a punto del método se detectaron 2 mutantes en el gen *PDS* en la población IRTA, uno de ellos en zona codificante (mutante C384) y con un valor SIFT de 0.00 que confirmaba que la mutación podía afectar a la funcionalidad de la proteína. Tras germinar 20 semillas M2 de la familia mutante C384 se obtuvieron plántulas con un fenotipo que mostraba los cotiledones alterados en su morfología, pero no se obtuvo ninguna planta albina. Sin embargo, al germinar semillas M3 obtenidas a partir de la autofecundación de un heterocigoto M2 para la mutación, si se obtuvo un individuo albino (Figura 4.1). Además, se comprobó mediante el análisis de la secuencia de *PDS* que era la única planta homocigota para la mutación. Por lo tanto, la detección de este mutante demostró que el método TILLING se estaba aplicando correctamente, y que se podían obtener fenotipos mutantes en nuestra población.



Figura.4.1: Plántula con fenotipo albino identificada en la familia C384, mutante en el gen *PDS* (Fitoeno desaturasa). Primer mutante detectado genotípica y fenotípicamente y que demuestra la eficiencia de la población de TILLING de M62-113.

Es importante destacar que en el momento de iniciar este trabajo no se conocía la secuencia completa del gen *PDS* en melón y por lo tanto, el diseño del amplicón de TILLING se hizo a partir de un unigen (MU9070). Actualmente, y gracias a que se dispone de la secuencia del genoma del melón, se ha podido determinar la estructura completa del gen *PDS*, que consta de 14 exones.

4.2.2 eIF4E (FACTOR DE INICIACIÓN DE LA TRADUCCIÓN 4E)

El gen eIF4E ha sido el gen en el que se han detectado más mutantes en este trabajo. En este gen, se han identificado mutaciones naturales responsables de impedir la interacción entre el factor de iniciación y la cadena infectiva de ARN positiva de algunos virus (Robaglia y Caranta, 2006). Robaglia y Caranta en 2006 revisaron los factores de iniciación de la traducción de algunas especies y su interacción con virus y cómo algunas mutaciones proporcionaban resistencia natural a diversos potivirus. Compararon la secuencia proteica de los genes *pvr2* de pepino (*Capsicum annuum*), *pot1* de tomate (*Solanum lycopersicum*), *mo1* de lechuga (*Lactuca sativa*) y *sbm1* de guisante (*Pisum sativum*). En este alineamiento señalaron las diferencias aminoacídicas entre los genotipos resistentes y los susceptibles (Figura 4.2, en fondo negro) en cada una de las especies citadas. También identificaron dos

regiones (Figura 4.2, zonas fucsia y azul turquesa) dónde se concentran un mayor número de polimorfismos entre los fenotipos susceptibles y resistentes. Además, ambas regiones conforman *loops* o lazos externos en la estructura 3D de la proteína. En este mismo alineamiento proteico hemos incluido el gen eIF4E de melón, con la histidina en la posición 228 marcada en blanco y fondo azul marino (Nieto y col., 2006) (Figura 4.2), responsable de la resistencia a MNSV (*nsv*), el gen ortólogo en cebada (*Hordeum vulgare*) *rym* responsable de la resistencia a BaYMV (Virus del Mosaico Amarillo en Cebada) y los mutantes encontrados mediante TILLING en este trabajo (Figura 4.2, en fondo rojo y verde). De este modo podemos comprobar que las mutaciones detectadas en este trabajo están localizadas en distintas regiones a las descritas anteriormente en otras especies. Al mismo tiempo, se ha podido determinar que las mutaciones Glu129Lys (familia F57), Ala134Val (familia 1229) y Gly136Glu (familia 444) provocan un cambio de aminoácido en regiones muy conservadas de la proteína, por lo tanto su efecto sea probablemente mayor que las mutaciones identificadas en zonas menos conservadas, las mutaciones Arg25Lys (familia F2036), Glu97Lys (familia 53) y la Glu206Lys (familia 2060). Las mutaciones silenciosas no provocan cambio de aminoácido y por lo tanto no afectarán a la funcionalidad de la proteína (Phe98Phe (familia 2532) y Thr155Thr (familia F1880)). Fijándonos en el alineamiento presentado en la Figura 4.2 encontramos el primer mutante detectado en la población IRTA en la posición 25 de la secuencia aminoacídica del gen eIF4E en melón (los mutantes IRTA están detallados en fondo rojo). Esta mutación pertenece a la familia F2036 y el cambio que provoca es de una arginina a una lisina. Ambos aminoácidos contienen el radical de cadena cargado positivamente pero la arginina contiene al final del radical dos grupos amino mientras que la lisina no. Para esta mutación el SIFT predice un valor 0.00 y por lo tanto un posible efecto en la funcionalidad de la proteína. En este caso se encuentra en una región muy poco conservada y por lo tanto quizás una región no tan imprescindible para la correcta funcionalidad de eIF4E.

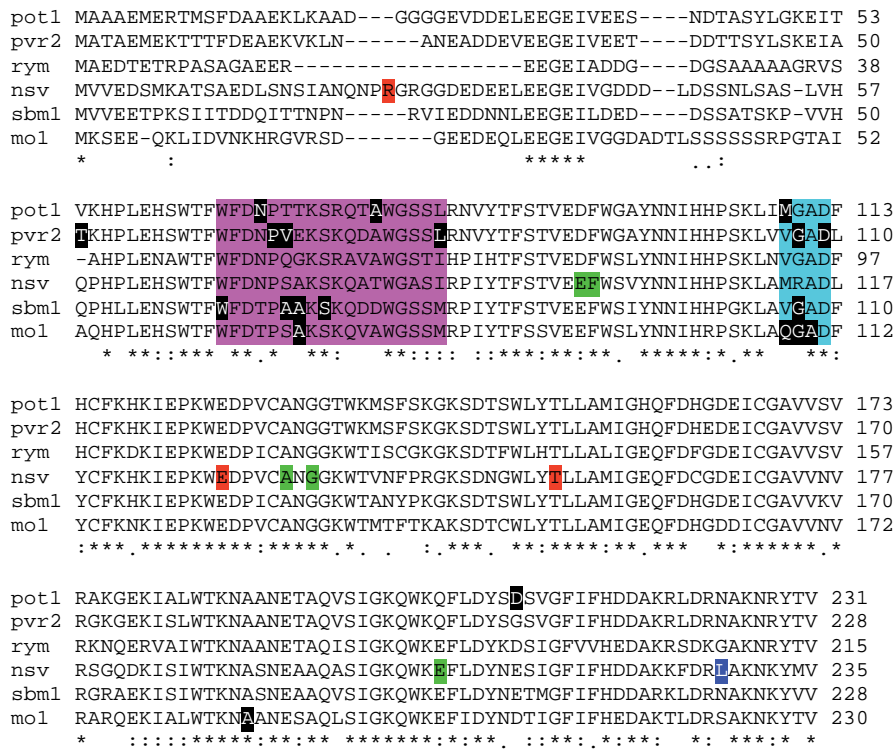


Figura 4.2: Alineamiento mediante *ClustalW* de la proteína eIF4E de distintas especies. Se destaca en fondo negro la localización de los aminoácidos que difieren entre los genotipos susceptibles y resistentes a potivirus de pimiento (*Capsicum annum*, *pvr2*), tomate (*Solanum lycopersicum*, *pot1*), lechuga (*Lactuca sativa*, *mo1*), guisante (*Pisum sativum*, *sbm1*), melón (*Cucumis melo*, *nsv*) y cebada (*Hordeum vulgare*, *rym*). Las partes coloreadas en fucsia y azul turquesa representan las zonas con más cambios de aminoácidos y a su vez zonas más expuestas al exterior tal y como determina la configuración en 3D de la proteína (Robaglia y Caranta, 2006). En rojo las mutaciones detectadas mediante TILLING en la población IRTA y en verde las mutaciones detectadas en la población del INRA.

Las siguientes mutaciones que encontramos pertenecen a las familias 53 y 2532 (en este caso en verde ya que se detectaron en la población del INRA). En la familia 53 la mutación se encuentra en la posición 97 y provoca un cambio de ácido glutámico a lisina. El SIFT en este caso predice una puntuación de 0.02 y por lo tanto un posible cambio en la funcionalidad de la proteína. Además, el cambio de ácido glutámico a lisina provoca también un cambio de carga (de negativo a positivo) que podría afectar en la interacción con otros aminoácidos y por lo tanto en el plegamiento de la proteína. Un posible cambio en la conformación de la proteína a nivel primario u otros niveles superiores podría afectar a la funcionalidad de la

proteína o a su interacción con su ligando. La familia 2532 tiene la mutación justo en la posición contigua a la anterior en el aminoácido número 98. No obstante, la mutación es silenciosa. Con este alineamiento hemos podido observar que hay ciertas regiones de la secuencia proteica más susceptibles a ser mutadas o que permiten la mutación, puesto que la probabilidad de mutar dos aminoácidos contiguos o muy cercanos es baja.

En la región delimitada por los aminoácidos 129 al 155 hemos detectado cuatro mutantes, dos en cada población. La mutación correspondiente al aminoácido 129 fue detectada en la familia F57 de la población IRTA. Este cambio es significativo (SIFT=0.00) puesto que provoca un cambio de ácido glutámico a lisina. Las mutaciones de las familias 1229 y 444 de la población INRA están localizadas con sólo un aminoácido de separación. Ambas producen cambio de aminoácido, la primera de alanina a valina (ambos hidrofóbicos) con un valor de SIFT de 0.10 mientras que la segunda de glicina a ácido glutámico con un valor de SIFT de 0.00. Este último parece que podría afectar más a la funcionalidad de la proteína puesto que el cambio estructural también podría ser más acentuado. La glicina es un aminoácido muy pequeño y su radical se basa en un átomo de hidrogeno, mientras que el ácido glutámico es un aminoácido cargado negativamente y con un radical mucho más grande que se estructura en una cadena de tres carbonos con un grupo carboxilo. Finalmente, la cuarta mutación en esta región fue identificada en la familia F1880 de la población IRTA, pero en este caso la mutación es silenciosa. Cabe destacar que las tres mutaciones (F57, 1229 y 444) están localizadas en una región muy conservada (Figura 4.2). La baja tasa de polimorfismos en esta región nos podría indicar que es importante para la funcionalidad de la proteína.

En la posición 206 encontramos el último mutante detectado en la familia 2060 perteneciente a la población INRA. En este caso el cambio producido es también de ácido glutámico a lisina. Esta última mutación se encuentra en una zona no tan conservada como puede verse en la Figura 4.2 y por ello su valor SIFT es también mayor (0.09).

De un total de ocho mutantes en región exónica del gen *eIF4E*, seis producen cambio de aminoácido y tres de ellos son de ácido glutámico a lisina. Aunque parece que la mutación mediante EMS es totalmente aleatoria, en nuestro caso ha habido una cierta tendencia a mutar algunas regiones más que otras y a su vez, ha habido tres casos de seis en que el cambio ha sido el mismo (E/K).

En la Figura 4.2 también se puede observar como ninguna de las mutaciones detectadas se localiza en las regiones de eIF4E con alto contenido en polimorfismos entre las variedades resistentes y las susceptibles de otras especies.

Actualmente existen muchas herramientas informáticas para poder estudiar con detalle como afectan las mutaciones a la funcionalidad de una proteína. En este trabajo se ha usado mayoritariamente el SIFT.

En 2006 Paulin y Henikoff publicaron un estudio donde presentaron una serie de métodos y herramientas para poder predecir los efectos de cambios de aminoácidos en la proteína, muchos de ellos disponibles via web. Uno de ellos es PMUT (*Phatology mutations*) (Rost y Sander, 1993) y aunque esté diseñado para detectar mutaciones en humanos resulta especialmente útil para la predicción de estructuras secundarias de proteínas.

En nuestro caso se han hecho también simulaciones en 3D con herramientas informáticas que se han presentado en el apartado de Resultados, utilizando PyMol. Sin embargo, resulta también interesante hacer un estudio detallado de la predicción de estructuras secundarias en la secuencia proteica. Estas estructuras son las primeras afectadas ya que se construyen a través de las interacciones entre aminoácidos. Por ello usamos esta herramienta (PMUT) en la que podemos observar la secuencia proteica de eIF4E con la predicción estructural correspondiente para cada aminoácido (Figura 4.3).

Las mutaciones correspondientes a las familias 1229 y 444 con cambios en las posiciones 134 y 136 están localizadas en una región con una alta probabilidad (valor 6) de asumir conformación de *loop* (Figura 4.3, detallado en fondo verde). Los *loops* son giros en la estructura secundaria de la proteína y por lo tanto definen en gran medida la estructura terciaria y cuaternaria. Es por ello que ambos mutantes podrían producir un cambio estructural de la proteína. Esta comparación no coincide del todo con los valores obtenidos con el SIFT puesto que a la primera mutación le asigna un valor de 0,10 y en cambio el PMUT le confiere un valor elevado en cuanto a su probabilidad estructural. En cualquier caso ambos resultados son complementarios puesto que cada programa usa unas matrices distintas y se basan en distintos parámetros iniciales.

```

SeqAA  MVVEDSMKATSAEDLSNSIANQNPGRGGDEDEEELEEIVGDDDLSSNLSASLVHQP
prH    000011111111112344444444322111111111111112222111112211111111221
prE    000000011111000000000000111111111233333333321111111122223333
prL    988877777666545544445566677765545444456666666666665444
SUBsec LLLLLLLLLLLLLL.....LLLLLLL.....LLLLLLL.....

SeqAA  PLEHSWTFWFDNPSAKSKQATWGASIRPIYTFSTVEEWSVYNNIHHPKSLAMRADLYCF
prH    21221122222222211111001122111111122222211112211100101011
prE    221111222211000000123443210111123334444333322111135676665543
prL    55666654455666667776544567666765544333344456666543212223344
SUBsec.....LL.....LLLLLLL.....LLLLLLL.....LL.....EEE.....

SeqAA  KHKIEPKWDPVCAAGKWTVNFPRGKSDNGWLYLLAMIGEQFDCGDEICGAVVNRSG
prH    11222212222211111122221222221222222111111111111211111112
prE    32223334333221112333332333333322233333321111123233344311
prL    555444444444566655444444444444444444333445667677654444433566
SUBsec.....LL.....LLLLL.....LL

SeqAA  QDKISIWTKNASNEAAQASIGKQWKEFLDYNESIGFIFHDDAKKFDRHAKNKYMV
prH    11111111112223322233334322112211111112357666553311100
prE    11112233321111123432121123332222235543100000000000100
prL    77766654456665654433444543334555455433356542323346678789
SUBsec LLLL.....L..H.....LLLLL

```

Figura 4.3: Output del programa PMUT con la secuencia proteica del gen eIF4E de *Cucumis melo* y los mutantes encontrados marcados en rojo (población IRTA) y en verde (población INRA). Las líneas inferiores corresponden a las probabilidades para cada aminoácido de contribuir en una conformación secundaria. prH: probabilidad de hélice, prE: probabilidad de cadena larga y prL: probabilidad de *loop*. Finalmente, la última línea (SUBsec) corresponde a la predicción de estructura secundaria en cada aminoácido según las probabilidades anteriores, donde H: hélice, L: *loop* y E: cadena larga. "." significa que no se puede asignar una estructura porque la probabilidad de cada conformación tiene puntuación igual o menor de 5.

Gracias a las predicción de la estructura 3D de la proteína eIF4E y a su modelaje mediante el programa PyMol se han podido analizar algunas de las mutaciones detectadas con mayor detalle. El estudio en el mutante 2 (familia F57) nos ha permitido observar que el cambio Glu129Lys podría afectar en la interacción con el sustrato (la molécula *Cap* del ARNm). Recordemos que tras la infección viral, si la interacción entre la proteína eIF4E y el ARNm del virus no se produce correctamente, el virus no podrá infectar la planta. En este mutante la sustitución es de un aminoácido cargado negativamente (ácido glutámico) a otro cargado positivamente (lisina). Este cambio podría provocar la desaparición o aparición de nuevos enlaces ocasionando una unión más débil entre el inicio del ARNm y el centro activo o de unión de eIF4E. Por este motivo, podría tratarse de un mutante menos eficiente en la interacción proteína-sustrato. Además, el valor de SIFT para

esta mutación es de 0.00, y por lo tanto confirma el posible efecto de la mutación en la funcionalidad de la proteína.

El mismo cambio pero en otra posición se ha producido en el mutante 5 (familia 53) que sustituye también el ácido glutámico por la lisina. El ácido glutámico en el genotipo silvestre crea un enlace de hidrógeno con la serina 93. Este enlace mantiene probablemente la hélice alfa y la hoja beta unidas dando estabilidad y cierta rigidez a ambas conformaciones. En cambio, la lisina que contiene el mutante no puede establecer este enlace puesto que su radical está cargado positivamente y no negativamente como el radical del ácido glutámico. La falta de este enlace podría proporcionar más flexibilidad en esta zona de la molécula con consecuencias impredecibles. Esta sustitución proporciona menor rigidez a esta parte estructural y en consecuencia menor estabilidad, además, el valor de SIFT para esta mutación es de 0.02.

Mediante la predicción en 3D se ha podido observar que la mutación del mutante 8 (familia 1229) prácticamente no provocaría cambios estructurales. El cambio de alanina a valina no altera la estructura puesto que ambos son aminoácidos no polares y con un radical muy pequeño. El valor de SIFT también confirma este dato dándole una puntuación de 0.10. Sin embargo, y como se ha comentado anteriormente, este resultado es contradictorio al resultado obtenido mediante PMUT. Esto es debido a que ambos programas se basan en distintos parámetros. En este caso, lo conveniente sería tener en cuenta ambos resultados y por lo tanto considerar que la mutación de la familia 1229 puede afectar a la estructura secundaria y quizás a la funcionalidad de la proteína.

Finalmente, la mutación del mutante 9 (familia 444), según el modelo 3D, no parece que provoque un cambio a nivel de giro en la estructura secundaria de la proteína. Sin embargo, el cambio de glicina a ácido glutámico introduce una carga negativa y un radical mucho mayor. Teniendo en cuenta la posición interna de este aminoácido y la introducción de la carga negativa, esta mutación podría afectar en las interacciones con los ligandos de la proteína eIF4E. Según el análisis con PMUT la mutación que contiene esta familia en la posición 136 de la secuencia proteica es una región con una alta probabilidad (valor 6) de asumir conformación de *loop* o giro.

En el protocolo de TILLING, a parte de la detección e identificación de mutantes, se debe realizar el estudio de segregación de la mutación y de su fenotipo. Para el gen

PDS pudo realizarse este estudio dado que el análisis fenotípico no requería mucho trabajo, tratándose de observar sólo el crecimiento de la planta y su estado de pigmentación. Para el resto de los genes analizados en este trabajo (*eIF4E*, *eIF(iso)4E*) el estudio fenotípico todavía queda pendiente por requerir más tiempo y recursos. Sin embargo, para el gen *eIF4E* se hizo la recolección de las semillas M2 y la obtención de las semillas M3 para poder fijar la mutación para, en un futuro, poder confirmar el efecto fenotípico mediante inoculación con una batería de virus (Tabla 4.3).

Tabla 4.3: Resultados del estudio de segregación de las familias mutantes del gen *eIF4E* encontrados en la población IRTA.*WT: Wild type o silvestre, H: Homocigoto, M: Mutante.

Amplicón	Exón	Posición	cambio AA	Familia M2	Plantas secuenciadas en M2	Segregación en M2 (WT/H/M)*
<i>eIF4E-1</i>	Exón 1	G74A	Arg25Lys	F2036	24	9/7/8
<i>eIF4E-2</i>	Exón 2	G385A	Glu129Lys	F57	8	5/3/0

Una vez identificadas las familias mutantes en el gen *eIF4E* de la población IRTA se creció la generación M2. Para la familia F2036 se pudieron crecer y secuenciar 24 plantas, mientras que para la familia F57 sólo se obtuvieron ocho debido a una baja tasa de germinación. Gracias a la secuenciación del fragmento del gen analizado se pudieron obtener los datos de segregación que aparecen en la Tabla 4.3.

Es relevante el hecho de que para la familia F57 no se obtuvo ningún homocigoto mutante. Este resultado podría estar relacionado con el hecho de que la mutación en homocigosis condicione el correcto crecimiento de la planta o su germinación. La mutación en esta familia tiene un valor en SIFT de 0.00 y por lo tanto la mutación afectaría a la funcionalidad de la proteína, como ya se ha comentado anteriormente.

En el caso de la familia F2036 la segregación fue de nueve silvestres homocigotos, siete heterocigotos y ocho homocigotos mutantes. Los homocigotos mutantes se podrían fenotipar para conocer y saber de qué modo afecta esta mutación a la funcionalidad de la proteína y si han adquirido alguna resistencia a algún virus que infecta a la planta del melón, sobretodo potivirus como WMV, CVYV o ZYMV, entre otros.

4.2.3 *eIF(iso)4E*

En el caso del gen *eIF(iso)4E* se obtuvo un mutante en la población INRA. La familia M1587 tiene una mutación que provoca un cambio de aminoácido de arginina a lisina. Algunos trabajos recientes han ayudado a entender mejor cuál es la diferencia de funcionalidad entre genes del complejo de iniciación de la traducción y sus isoformas. Un ejemplo es el de *eIF4F* y *eIF(iso)4F* en los cuales se ha demostrado que cada uno posee mayor afinidad por algún tipo de ARNm (Gingras y col., 1999). Otros estudios demuestran que los genes *eIF4E* y *eIF(iso)4E* pueden expresarse diferencialmente en según que tejidos y periodos de desarrollo de la planta. Mientras que en *Arabidopsis* se comprobó que ambos genes son expresados por igual en tejidos meristemáticos y en células en período de división de órganos inmaduros, se expresan en distinta proporción en tejido maduro (Rodríguez y col., 1998). Parece que las resistencias a virus descritas hasta ahora indican que cada virus tiene una cierta especificidad con uno de los dos genes y también esta especificidad parece variar en función de la especie de planta. El TEV (*Tobacco etch virus*) interacciona con el gen *eIF4E* en tomate, lechuga y pimiento mientras que en *Arabidopsis* interacciona con su isoforma (Lellis y col., 2012).

Es por ello que tanto el gen *eIF(iso)4E* como el *eIF4E* deben tenerse en cuenta en la búsqueda de posibles nuevas resistencias recesivas. Además, existen casos en los que sólo el gen *eIF(iso)4E* es el responsable de conferir la resistencia a un virus. Recientemente se ha demostrado que el virus PPV (*Plum pox virus*) sólo interacciona con el gen *eIF(iso)4E* mientras que no lo hace con el gen *eIF4E*. En el silenciamiento del gen *eIF(iso)4E* en plantas de *Prunus*, éstas se mostraron resistentes frente a la infección con el virus PPV mientras que en el silenciamiento del gen *eIF4E* no se obtuvo resistencia. Además, se comprobó la especificidad de interacción entre el Vpg del virus y el gen *eIF(iso)4E* mediante ensayo de doble híbrido, corroborando el mismo resultado (Wang y col., 2013).

En melón existen estudios que confirman que el silenciamiento del gen *eIF4E* aporta resistencia frente algunos virus. Más concretamente el experimento realizado por Rodríguez-Hernández y col., 2012 demostró que el silenciamiento de este gen confería resistencia a cuatro (CVYV, MNSV, MWMV, ZYMV) de los ocho virus con los que infectaron las plantas de melón. Sin embargo, no se comprobó el silenciamiento del gen *eIF(iso)4E* y por lo tanto no se puede saber si este conferiría resistencia para algunos de estos virus.

variedad CharMono de la población INRA los cebadores no fueron suficientemente específicos. Probablemente los cebadores diseñados no amplifican esa región porque la región complementaria no existe o contiene demasiados polimorfismos respecto a la región original de M62-113. Por ello se decidió no seguir adelante con el análisis de este gen en la población INRA.

4.2.5 FUTUROS EXPERIMENTOS

Con los resultados obtenidos en este trabajo queda pendiente la fijación de algunos de los mutantes en homocigosis y la obtención de suficiente semilla de los mismos, para poder hacer los experimentos de fenotipado adecuados para cada gen.

También es indispensable obtener el mutante aislado de otras mutaciones, puesto que una de las desventajas del TILLING es que las plantas seleccionadas no sólo contienen la mutación de interés sino también muchas otras que podrían afectar al fenotipo en estudio.

En este trabajo esta labor debería hacerse con los mutantes obtenidos del gen eIF4E (F2036 ya se ha fijado y faltaría fijar las familias 1229, 2060, 444, F57 y 53). A continuación se deberían infectar con una batería de virus para poder comprobar si alguna de las mutaciones confiere resistencia a la planta. Sin embargo, es posible que para algún mutante no sea posible obtener el homocigoto. En el caso de la familia F57 ya se ha intentado varias veces a partir de la autofecundación de un heterocigoto M2 y no ha sido posible.

Respecto a la batería de virus con la cual deberían infectarse los mutantes, ya existe un estudio reciente (Rodríguez-Hernández y col., 2012) que ha demostrado como mediante el silenciamiento (ARNi) del ARNm de Cm-eIF4E las plantas de melón transformadas resultaban resistentes a 4 virus distintos (*Cucumber vein yellowing virus* (CVYV), *Melon necrotic spot virus* (MNSV), *Zucchini yellow mosaic virus* (ZYMV) y *Moroccan watermelon mosaic virus* (MWMV)). Durante este trabajo se probaron también otros virus, como *Cucumber mosaic virus* (CMV) (género *Cucumovirus*), *Watermelon mosaic virus* (WMV) (género *Potyviridae*), *Cucurbit yellow stunting disorder virus* (CYSDV) (género *Crimivirus*), *Cucurbit aphid-borne yellows virus* (CABYV) (género *Polerovirus*), *Papaya ringspot virus* (PRSV) (género *Potyviridae*) y *Beet pseudoyellows virus* (BPYV) (género *Closteroviridae*), pero en estos casos el silenciamiento de eIF4E no confirmó resistencia (Rodríguez-Hernández y col., 2012). Gracias a los resultados presentados en este trabajo se han podido determinar las condiciones para desarrollar posteriores plataformas de TILLING, un ejemplo es la

desarrollada por la empresa *Semillas Fitó* juntamente con el IRTA-CRAG en el marco del proyecto MELONOMICS. Tal y como se ha comentado anteriormente, esta plataforma se ha desarrollado a partir de una línea pura de melón tipo "Piel de Sapo". Se han mutado las semillas con un porcentaje de 1,5% de EMS y se han obtenido aproximadamente 3100 individuos M2 (en el momento de escribir esta tesis), en los cuales ya se han detectado 8 mutantes en el gen *eIF4E* y 3 en el gen *eIF(iso)4E*. De los 8 mutantes en el gen *eIF4E*, 6 tienen la mutación en región exónica mientras que 2 en región intrónica. De los mutantes exónicos, 2 tienen un valor SIFT menor de 0.05 y por lo tanto no serían tolerados. Para el gen *eIF(iso)4E*, el número de mutantes encontrados sigue siendo mucho menor y de los tres mutantes detectados sólo uno se localiza en región exónica con valor 0.02 en SIFT. Esta proporción confirmaría la hipótesis de que *eIF(iso)4E* es más indispensable para la célula (Wang y col., 2013).

Todos estos mutantes, juntamente con los obtenidos en esta tesis serán en un futuro infectados con algunos de los virus mencionados anteriormente y de este modo se probará si la mutación que ha adquirido cada uno confiere alguna resistencia a la planta, de modo que pudiera ser utilizable en programas de mejora genética de la especie.

Debe tenerse en cuenta que para poder comercializar nuevas líneas obtenidas a partir de la mejora genética, la tecnología a utilizar en Europa en estos momentos no puede ser la transformación genética. Por ejemplo, la tecnología ARNi no se podría utilizar en productos comerciales, en cambio, la técnica de TILLING no presenta esta limitación ya que se obtienen nuevas líneas mediante una tecnología que no emplea transgenesis.

4.3 EcoTILLING

El EcoTILLING ofrece la posibilidad de identificar eficientemente mutaciones naturales en las variedades analizadas. Por ello resulta muy interesante analizar distinto germoplasma para la detección de haplotipos presentes en la naturaleza y que puedan aportar variabilidad natural e información funcional en genes de interés. En el marco del proyecto MELOGEN se decidió hacer el análisis de EcoTILLING en los genes ya analizados mediante TILLING. En esta tesis se ha realizado el análisis de los genes *eIF4E* y *ETR1* y en el COMAV se realizó el análisis en los genes *eIF(iso)4E* y *PDS*.

El análisis se realizó comparando la misma variedad con la que se había desarrollado la población de TILLING (M62-113) y 113 accesiones proporcionadas por el banco de germoplasma del COMAV. Las 113 accesiones están caracterizadas para ciertos caracteres, entre ellos caracteres de fruto como el color de la pulpa, el aroma y, de algunas de ellas, se dispone de imágenes del fruto (esta información está disponible en el COMAV). Sin embargo, desafortunadamente no están caracterizadas para resistencia a patógenos. Por ello, el análisis de EcoTILLING con el gen *eIF4E* genera un primer cribado para seleccionar aquellas variedades que pudieran tener un haplotipo que confiriera resistencia a potivirus y en una segunda etapa, se deberían valorar fenotípicamente las accesiones de interés.

Con el análisis del gen *ETR1* se pretendía relacionar algún haplotipo detectado con algún carácter de maduración de fruto teniendo en cuenta que la variedad de referencia M62-113 es una variedad no climatérica.

4.3.1 *eIF4E* (FACTOR DE INICIACIÓN DE LA TRADUCCIÓN)

Como se ha comentado a lo largo de este trabajo, se ha descrito que el gen *eIF4E* es responsable de algunas resistencias a potivirus y a otros virus, ya que la presencia de mutaciones específicas en este gen provocan que el virus no pueda interaccionar con el complejo de iniciación de la traducción y por lo tanto, traducirse como primer paso de la infección de la planta. Por este motivo encontrar variabilidad natural en este gen podría implicar identificar algún nuevo haplotipo que confiriera resistencia a algún virus a la planta.

Se han identificado un total de 12 polimorfismos en el análisis de EcoTILLING con el gen eIF4E. De los 12, seis SNPs se encuentran en zonas exónicas y uno en el 5' UTR del gen. Sólo uno de ellos provoca un cambio de aminoácido (SNP 1.5, Asp71Gly) (Tabla 4.4) con un valor 0.00 en SIFT y que por lo tanto, podría afectar en la funcionalidad de la proteína.

Tabla 4.4: Tabla comparativa de los polimorfismos detectados por EcoTILLING en el estudio de Nieto y col. (2007) y los identificados en este trabajo. * Posición en la secuencia nucleotídica respecto al codón de inicio de la traducción y numeración asignada en este trabajo.

Estudio	Exón/SNP*	Cambio aminoácido	Accesión	Resistencia a virus
EcoTILLING IRTA	1/T81A (1.2)	Silencioso		
	1/C123G (1.3)	Silencioso		
	1/G186T (1.4)	Silencioso		
	1/A212G (1.5)	Asp71Gly	PI 505602	No hay datos
	1/C243T (1.6)	Silencioso		
	5/G690A (2.5)	Silencioso		
EcoTILLING Nieto y col., 2007	1/T81A (1.2)	Silencioso		
	1/C123G (1.3)	Silencioso		
	1/G186T (1.4)	Silencioso		
	5/T683A	Leu228His	C-046	MNSV
			C-178	MNSV
			C-512	MNSV
			PI 161375	MNSV
5/G690A (2.5)	Silencioso			

El análisis mediante EcoTILLING del gen eIF4E en melón ya se había realizado anteriormente (Nieto y col., 2007) con el melón climatérico Vedrantais como referencia y otras 147 variedades de melón obtenidas de la colección de germoplasma de la Estación Experimental de "La Mayora"- CSIC (Málaga). En este trabajo previo se detectaron cinco polimorfismos, tres de ellos en el primer exón y dos en el quinto (Tabla 4.4). Todos los SNPs que se detectaron resultaron ser

silenciosos excepto el cambio T683A del quinto exón que corresponde al cambio de leucina a histidina en la posición 228 y el cuál confiere resistencia a MNSV en PI 161375. Las accesiones probadas en este trabajo fueron sometidas a estudios previos de resistencia frente a una batería de virus pudiendo identificar variedades resistentes para algunos de ellos, como las 4 accesiones resistentes a MNSV (indicadas en la Tabla 4.4). Los amplicones utilizados en ambos trabajos son los mismos.

En la Figura 4.4 se muestran todos los SNPs detectados en este trabajo y por Nieto y col., (2007) para el gen eIF4E.

a)

```

atggtagttgaagattcgcgatgaaagctacatcgccggaagatcttttctaattccattgct
M V V E D S M K A T S A E D L S N S I A
aatcaaaaccctagaggacgtggcggtgacgaagatgaggaacttgaggaaggtgagatc
N Q N P R G R G G D E D E E L E E G E I
gtcggcgacgacgacctcgactcctccaatttgccgcgtccctagtcacatcagcctcac
V G D D D L D S S N L S A S L V H Q P H
cctctggagcactcttggaccttttggttcgataacccatctgccaatccaagcaagcc
P L E H S W T F W F D N P S A K S K Q A
acctgggggtgcgtctattcgaccgatctataccttctctaccgtcgaggagtctggagt
T W G A S I R P I Y T F S T V E E F W S
gtttacaacaacattcatcatccaagcaaatggcgatgagggcagatttgtactgcttc
V Y N N I H H P S K L A M R A D L Y C F
aaacataaaattgagcctaaatgggaagatcccgtttgtgctaattggagggaaatggact
K H K I E P K W E D P V C A N G G K W T
gtgaactttccaaggggaaatctgataaatggctggtgtacaccctgcttgcctatgatc
V N F P R G K S D N G W L Y H I L A M I
ggagaacagtttgactgtggtgatgaaatttgggagcagttgtaatgtaggtctggg
G E Q F D C G D E I C G A V V N V R S G
caggataaaatcaatttggacgaagaatgcttccaatgaagctgcgagggcagcatt
Q D K I S I W T K N A S N E A A G A S I
ggaaaacagtggaaggagtcttcttgattacaatgagagcattggctttatattccacgat
G K Q W K E F L D Y N E S I G F I F H D
gacgcaagaaattcgcgatagacatgcaagaataaatatattggtgga
D A K K F D R H A K N K Y M V
    
```

Rojo: final exones
Verde: inicio exones

b)

E1: (297)
 ATG GTAGTTGAAGATTCCGATGAAAGCTACATCGGCGGAAGATCTTTCTAATCCATTGCTAATCAAACCCCTAGAGGACG
 GGC CGGTGACGAAGATGAGGAACTTGAGGAAGGTGAGATCGT GGGC GACGACGACCTCGACTCCTCCAATTTGTCCCGCT
 CCTAGTGATCAGCCTCACCCCTT GAGCACTCTTGGACCTTTTGGTTCC TAACCCATCTGCCAAATCCAAGCAAGCC
 AC TGGGGTGGCTTATTCCGACCGATCTATACCTTCTCTACCGTCCGAGGAGTTC TGG

E2: (168)
 AGT GTTTACAACAACATTCATCATCCAAGCAAATTGGCGATGAGGGCAGATTTGTACTGCTTCAAACATAAAATTGAGCC
 TAAATGGGAAGATCCCGTTTGTGCTAATGGAGGGAAATGGACTGTGAACTTTCCAAGGGGAAAATCTGATAATGGCTGGT
 TGTAC TGG

E3: (132)
 CTG CTTGCTATGATCGGAGAACAGTTTGACTGTGGTGATGAAATTTGTGGAGCAGTTGTTAATGTTAGGTCTGGGCAGGA
 TAAAATATCAATTTGGACGAAGAATGCTTCCAATGAAGCTGCGCAGGCG TGG

E4: (66)
 ATT GGAACAACAGTGAAGGAGTTTCTTGATTACAATGAGAGCATTGGCTTTATATCCACGAT DAC

E5: (45)
 CCA AAGAAATTCGATAGACA TCCGAA AATAAATATATGCTC TGA

Figura 4.5: a) Traducción a proteína del ARNm del gen eIF4E. En rojo se señala el último codón de cada exón y en verde el inicio de cada exón. b) Exones del gen eIF4E separados indicando el tamaño de cada uno (en pb) y las variaciones encontradas: en fucsia las detectadas en este trabajo y en el trabajo de Nieto y col., (2007). En azul, el C123G (SNP 1.3) y el A212G (SNP 1.5) detectados en este trabajo pero no en el de Nieto y col., (2007). En gris, el SNP detectado en el trabajo de Nieto y col. (2007) pero no detectado en este trabajo.

Cuatro de los SNPs detectados en este trabajo coinciden con los que fueron detectados mediante EcoTILLING en el estudio de Nieto y col. 2007 (Tabla 4.4 y Figura 4.5b, en fucsia). Esto demuestra que dentro de la variedad natural encontramos algunos SNPs con más frecuencia y por lo tanto, que se detectan en distinto germoplasma. En cambio, otros SNPs se pueden encontrar con menor frecuencia como es el caso del SNP 1.3 (C123G) y el SNP 1.5 (A212G) que se han detectado sólo en el germoplasma analizado en este trabajo y no se detectaron en las variedades analizadas en el estudio de Nieto et al. El SNP 1.5 está presente sólo en la variedad PI 505602 (Tabla 4.4) y provoca un cambio de ácido aspártico a glicina. Para comprobar si este SNP provoca alguna modificación en la conformación de la proteína y en consecuencia, confiere resistencia frente algún patógeno debería infectarse esta variedad (Tabla 4.4) con una batería de virus, como ya se ha comentado anteriormente. De todos modos, debe tenerse en cuenta que esta variedad presenta el polimorfismo en heterocigosis y debería primero fijarse la mutación.

Por otro lado, el SNP detectado en el trabajo de Nieto localizado en el exón 5 y que confiere resistencia a MNSV no se ha identificado en ninguna de las variedades analizadas en este trabajo (Tabla 4.4). En este caso el SNP se presentaba en homocigosis en las variedades que resultaron ser resistentes a MNSV.

En ambos estudios el exón que presenta más variabilidad ha sido el primero mientras que en el segundo, tercero y cuarto no se han identificado SNPs entre las variedades comparadas.

4.3.2 *ETR1* (RECEPTOR DE ETILENO 1)

En el gen *ETR1* se han identificado dos polimorfismos mediante la técnica del EcoTILLING entre las 113 variedades analizadas y la variedad de referencia M62-113. El SNP 1 provoca un cambio de aminoácido de leucina a valina en la posición 224 de la secuencia aminoacídica con un valor SIFT de 1.00 mientras que el SNP 2 es silencioso.

La proteína *ETR1* en *Arabidopsis* está bien descrita. Esta proteína contiene tres dominios (Figura 4.6). El dominio amino terminal (aminoácidos 1-325) contiene tres segmentos transmembrana. Expresando la proteína *ETR1* en levadura se ha demostrado que este primer dominio es el responsable de la unión con el etileno (Schaller y Bleecker, 1995). Además, la mutación *etr1-1* localizada en este dominio pierde la capacidad de unión con la hormona (Schaller y Bleecker, 1995).

El segundo dominio (aminoácidos 326-609) se encuentra en una región que presenta homología con dominios tipo histidina quinasa y de la cual se ha demostrado su actividad quinasa *in vitro* (Gamble y col., 1998, 2002). El último dominio (aminoácidos 610-738) es el responsable de recibir el fosfato proveniente de la actividad de la histidina del dominio anterior (Chang y col., 1993) y por lo tanto responsable de la regulación de la respuesta hacia el interior celular de la proteína *ETR1*.

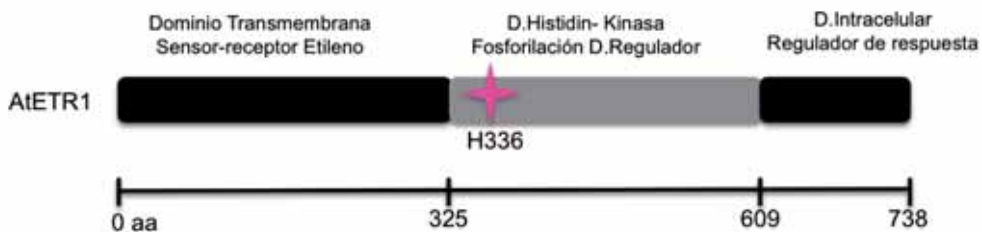


Figura 4.6: Estructura de los dominios identificados en el gen *ETR1* de *Arabidopsis*. Cada dominio tiene una función distinta y en fucsia se detalla la posición de la histidina 336 de la secuencia aminoacídica, responsable de la fosforilación del dominio regulador.

En *Arabidopsis* se han identificado mutaciones en el primer dominio que producen insensibilidad al etileno en las posiciones Ala31Val, Ile62Phe, Cys65Tyr y Ala102Thr (Figura 4.7 en amarillo) (Wilkinson y col., 1995).

En tomate se han podido aislar los genes *LeETR1*, *LeETR2*, *LeETR3 (LeNR)*, *LeETR4* y *LeETR5*. El estudio de estos genes ha sido muy útil para poder comprender mejor la actividad del gen *ETR1* y su familia génica. Al comparar la secuencia de estos cuatro genes se ha demostrado que la región más conservada es la responsable de la unión con el etileno (Tieman y col., 2000). Además, el gen *LeETR5* ha perdido la histidina 336 del dominio histidina quinasa (Figura 4.6) aparentemente responsable de la fosforilación del tercer dominio, demostrando que la actividad de la histidina no parece indispensable para la respuesta al etileno (Tieman y col., 2000).

Existen otras mutaciones relevantes que han ayudado en la comprensión de la funcionalidad del gen *ETR1*. Por ejemplo, en tomate se identificó el mutante *Never ripe (NR, LeNR)*. Este mutante produce frutos que no maduran (Rick y Bulter, 1956), ya que es insensible al etileno en todos los tejidos (Lanahan y col., 1994). La mutación responsable de este fenotipo es un cambio de prolina a leucina en la posición 36 (Wilkinson y col., 1995) (Figura 4.7).

Una vez identificadas las mutaciones descritas en *Arabidopsis* y en tomate que afectan a la funcionalidad del gen produciendo insensibilidad a etileno, procedimos a compararlas con los polimorfismos que se habían identificado mediante EcoTILLING en melón. Con un alineamiento múltiple (Figura 4.7) de los genes *ETR1* de *Arabidopsis (AtETR1)*, el *ETR1* de tomate (*LeETR1*), *ETR4* y *5* de tomate (*LeETR4, LeETR5*), el *NR* de tomate (*LeNR*) y el *ETR1* de melón (*CmETR1*) hemos podido identificar y comparar las posiciones de los polimorfismos y discutir su efecto.

En verde en la Figura 4.7 se detalla el polimorfismo encontrado mediante EcoTILLING en melón. Este SNP provoca una sustitución de Leu224Val (Figura 4.7). Este cambio se localiza en la región del dominio transmembrana (sensor-receptor del etileno) y por lo tanto coincide en la región en que se han descrito mutaciones que afectan a la funcionalidad de la proteína en *Arabidopsis* y en tomate (Wilkinson y col., 1995). Esta sustitución (Leu224Val) será probablemente tolerada ya que si nos fijamos en la secuencia, tanto en tomate como en *Arabidopsis* existen receptores que contienen una valina en esta posición, lo que favorece pensar que el mutante de melón no provocaría ningún cambio en la funcionalidad de la proteína (Figura 4.7), dato corroborado mediante SIFT.

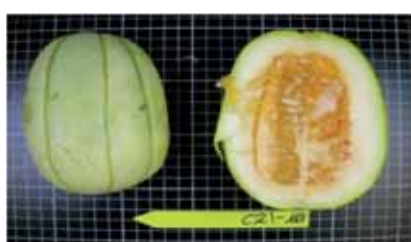
LeNR	LPILAVGDAKRLIQTLLNVAGNAVKFTKEGHISIEASVAKPEYARDCHPPEMFPMPDGO	507
LeETR5	LPNQVMGDEKRTFQVLLHMVGHLLNVSIGKGSVIFRVVLETGAETGNDKVGTRRPSSTD	534
LeETR4	LPNHVLDGERRVFQVILHMVGNLLKDPNG-GLLTFRVLPESVSRREGIGAWRTRRSNSR	532
	** .:*** :* :* :*:..* : : . : : . : .	
AtETR1	FYLRVKVKDSGAGINPDIPKIFTKFAQTQSLATRSSGGSGGLGLAISKRFFVNLMEGNIWI	564
CmETR1	FYLRVQVKDTGSGISPDIPKLPKTKFAQT-VGPRNSGGSGGLGLAICKRFFVNLMEGHIWL	588
LeETR1	FYLRVQIKDGTGIGITPDIPNLFSKFTQSALATNSGGTGLGLAICKRFFVNLMEGHIWI	582
LeNR	FYLRVQVRDTGCGISPDIPLVFTKFAESRPTSNRSTGGEGGLGLAICRRFIQLMKGNWI	567
LeETR5	EYVTIKFEIEVSLGEGSDSSISTIHFGGRRHNSKEVT-EGLSFNMCKKLVQMMQGNIM	593
LeETR4	DNAYIRFEVGTSNHSQPEGTMPLPHYRPKR--CSKEMD-EGLSFTVCRKLVQLMQGDIW	589
	: : . . . * : : . . . * : : : : : : : * : * : *	
AtETR1	ESDGLGKCTAIFDVKLGISERSNES-KQSGIPKVPALPRHSNFTGLKVLVMDENGVSRM	623
CmETR1	ESEGLGKCTATFIVKLGIAQSNES-KLPYTSKIHENSIHTSFPGLKVLVMDNDG----	643
LeETR1	ESEGLGKCTAIFIIKLGIPGRANES-KLPFVTKLPANHTQMSFQGLKVLVMDENGVSRM	641
LeNR	ESEGPQKGTTFVVKLGIHHPNALPLLPMPRGRNLKGSDDLFRYRQFRGDDGMS--	625
LeETR5	SSNAQGHAGMTLILRFQKQSSFRKR-MFEYRNPLEQPISSTMFRGLHVLLTDDDDVNRL	652
LeETR4	IPNPEGFDQSMVVVLGLQLRFSIAIG-IPEYGESSDHSHPHLLQGVKVLVLDYDDVNRA	648
	: : * : : : : : : : : : : : : * : . .	
AtETR1	VTKGLLVHLGCEVTTVSSNEECLRVVSHHEKVVFMDCMPGVENYQIALRIHEKFTKQRH	683
CmETR1	-----	
LeETR1	VTKGLLTHLGCDDVTVGSRDECLRVVTHEHKVVIMDVSMQGIDCYEVAVVIHERFG-KRH	700
LeNR	-----VNAQRY	631
LeETR5	VTRKLEKLGCVTAVSTGFCLSALGPSLTFQVLLLDLQMPMDGYEVALRVKFRSR	712
LeETR4	VTSKLEKLGCSVSAVSSGRDCIGVLSPAVSSFQIVLLDLHLPLDLDFEVTMRIRKFGSH	708
AtETR1	QRPLLVALSGNTDKSTKEKCMSFGLDGVLLKPVSLDNIRDVLSDLLEPRVLYEGM	738
CmETR1	-----	
LeETR1	GRPLIVALTGNTDRVTKENCMRVGMDGVILKPVSVYKMRSVLSLELLEHGVVLES-	754
LeNR	QRSL-----	635
LeETR5	SWPLIIALTASSEEQVWEKCLQVGMNGLIRKPVLLQGLADELQRLLRGGGGDGL	767
LeETR4	NWPLIVGLTATADENVTRCLQIGMNLIRKPVLLPGIADELQRVLLRGSRM--	761

Figura 4.7: Alineamiento múltiple de los genes *ETR1* en *Arabidopsis*, tomate y melón y los genes *ETR4*, *ETR5* y *NR* de tomate. En color amarillo las mutaciones en *Arabidopsis* que provocan insensibilidad en etileno, en rojo la mutación del fenotipo *Nr* (*Never Ripe*), en fucsia la histidina con actividad fosforilante en el gen *ETR1* de *Arabidopsis* y en verde el cambio detectado mediante EcoTILLING en melón.

El COMAV dispone de información fenotípica de caracteres morfológicos como el color de la pulpa, el aroma y el sabor para algunas de las variedades. En la Tabla 4.5 se detallan algunos de los caracteres que parecen estar relacionados con la actividad del etileno durante el proceso de la maduración del fruto, como la consistencia de la pulpa, la intensidad del aroma o la maduración climatérica. Las 4 accesiones descritas en la Tabla 4.5 presentan el SNP 1 que causa la sustitución de Leu224Val y se caracterizan por tener la carne blanca (Figura 4.8) y además, un aroma y dulzura menos intensos que en otras variedades. Una de ellas presenta maduración tipo climatérico, dos de ellas no, y para dos más no existen datos sobre tipo de maduración.

Tabla 4.5: Características fenotípicas de las accesiones que presentan el SNP 1 (de la variedad PI 282448, que también presenta el SNP 1, no se dispone de datos). ND: No hay datos disponibles.

Nombre variedad	spp	Origen	Color y consistencia de la pulpa	Intensidad del aroma	Dulzura	Maduración climatérica
CUM349	melo	Irak	Blanca/medio-blanda	Ácida	no dulce	Si
PI 124112	agrestis	India	ND	ND	ND	ND
PI 482429	agrestis	Chakonda (Zimbawe)	Blanca/dura	no aroma	no dulce	No
PI 436534	agrestis	Tiguinchor (Senegal)	Blanca/medio-blanda	poco aromático	poco dulce	No



PI 124112 (C21)



CUM349 (C24)



PI 482429 (C99)



PI 436534 (C25)

Figura 4.8: Imagen de las accesiones en las que se han detectado el SNP 1 que provoca un cambio de aminoácido en el primer amplicón del gen *ETR1* (de PI 282448 no se dispone de imagen).

Por tanto, la información de que se dispone no permite afirmar que la mutación detectada en el gen *ETR1* esté implicada en cambios en el tipo de maduración del fruto.

4.4 ESTUDIO FENOTÍPICO DE LA POBLACIÓN DE TILLING

Inicialmente el objetivo del estudio fenotípico de 400 familias M2 de la población de TILLING creada en este trabajo era encontrar mutantes con características de interés comercial. Sobretudo, en caracteres del fruto cómo la forma, el color, el escriturado y también en caracteres de la planta como algún tipo de resistencia, por ejemplo, resistencia al oídio (*Podosphaera xanthii*). Se crecieron diez plantas por familia para poder observar si había más de una planta por familia que mostraba el fenotipo mutante y de este modo confirmar que era un fenotipo producto de una mutación heredada y no un efecto ambiental.

Otras plataformas de TILLING han desarrollado bases de datos donde se ofrece además un servicio que proporciona las semillas del mutante que se desee estudiar. Las bases de datos facilitan métodos de búsqueda basados en el gen de interés para el cual queremos encontrar mutantes o basadas en la búsqueda de un fenotipo concreto. Por ejemplo, en el caso de la base de datos de *UTILLdb* (*URGV TILLING database*) (<http://urgv.evry.inra.fr/cgi-bin/projects/Tilling/index.pl>) se puede acceder a la información recopilada de tres poblaciones distintas de TILLING: *Lycopersicum esculentum*, *Pisum sativum* y *Brachypodium distachyon* (Figura 4.9, 1). Para cada una de ellas, se puede realizar una búsqueda en función del gen en el que podamos estar interesados (Figura 4.9, 2a) o bien un fenotipo ya caracterizado (Figura 4.9, 2b).

En esta base de datos tienen caracterizados 31 genes con 559 mutaciones y 2.405 fenotipos, sumando los resultados obtenidos para las tres especies. Si se selecciona algún gen, la base de datos proporciona información sobre las familias mutantes que se han detectado para éste (Figura 4.9, 3a, como ejemplo se ha seleccionado el gen *eIF(iso)4E*). Posteriormente, también se puede obtener información más detallada (Figura 4.9, 4a) edonde se muestra la secuencia del gen y la posición concreta de la mutación tanto a nivel nucleotídico como aminoacídico. Por otro lado, se puede seleccionar un fenotipo (Figura 4.9, 3b) y a partir de éste ver imágenes y, en algunos casos, si se ha detectado alguna mutación en la familia de la imagen seleccionada (Figura 4.9, 4b).

En definitiva, esta base de datos es un ejemplo de como se puede explotar la información obtenida en una o varias plataformas de TILLING. De este modo, las

mutaciones están bien identificadas y son de uso público de modo que toda la comunidad científica puede acceder a ellas.

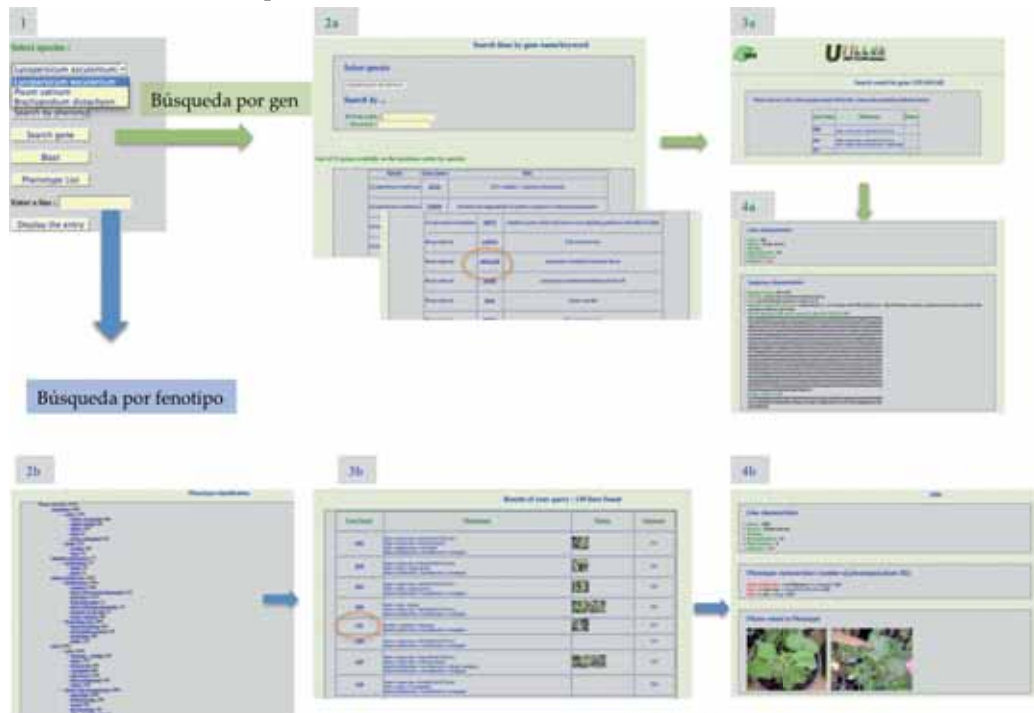


Figura 4.9 : Información disponible en la base de datos UTILLdb. Se puede realizar la búsqueda para un gen concreto (vía a, en verde) o en función de un fenotipo (vía b, en azul).

En nuestro caso no se ha realizado una base de datos pero en un futuro debería diseñarse quizás conjuntamente, con las otras plataformas de TILLING desarrolladas en esta misma especie. Por el momento, en este trabajo si se ha hecho un primer catálogo fenotípico ya que durante el año 2007, se siguió con detalle el crecimiento de 400 familias M2 desde la germinación hasta la obtención de fruto. En este período se tomaron imágenes para poder registrar aquellos mutantes con nuevos fenotipos. El objetivo era visualizar fenotipos y determinar el número y la diversidad de los mismos, también como dato confirmativo de que la población estaba mutada con una cierta eficiencia. Observamos fenotipos a nivel de pigmentación con plantas cloróticas o parcialmente cloróticas. También a nivel de crecimiento de la planta observando fenotipos enanos. A nivel de desarrollo observamos en algunas familias crecimiento sin dominancia apical y con un desarrollo internodal mucho mayor. Finalmente se observaron familias mutantes

en la forma y características del fruto. Los fenotipos descritos en el catálogo pueden ser un inicio para la creación de una base de datos más completa.

Por ejemplo, los fenotipos descritos con pérdida de la dominancia apical podrían poseer una deficiencia en la biosíntesis o transporte de las auxinas o citoquininas, ya que estas hormonas son las que regulan el crecimiento con dominancia apical (Bohn-Courseau, 2010). Para otros fenotipos es más complicado determinar que genes pueden ser responsables de los caracteres mutantes. Sin embargo, algunos fenotipos pueden ser interesantes a nivel agronómico y comercial sin necesidad de saber cuáles son los genes responsables de ese efecto. Por ejemplo, se han encontrado familias con características diferentes como la forma y el color de fruto respecto la variedad M62-113. En alguna familia los frutos son mucho más redondeados (Figura 4.10) y dado que el melón típico de "Piel de sapo" es elíptico este nuevo carácter podría ser interesante a nivel comercial. Respecto al color, se han encontrado familias con un color liso, sin escriturado (Figura 4.10) y también en este caso podría ser un carácter interesante a nivel comercial puesto que podría resultar atractivo para el consumidor. La elección de estos fenotipos para su inclusión en programas de mejora genética debe ser valorada por las empresas de semillas siempre que se confirmen los fenotipos en generaciones más avanzadas.



Figura 4.10: Izquierda: Imagen del fruto de la familia F1902 con la piel lisa, sin manchas ni escriturado. Derecha: Imagen del fruto de la familia F2286 con una forma redondeada totalmente distinta a la elíptica, típica del melón "Piel de sapo".

Una vez caracterizados los fenotipos encontrados en la población de TILLING de melón podemos compararlos con los fenotipos detectados en las poblaciones

desarrolladas en otras especies. En la Figura 4.11 se muestran fenotipos encontrados en otras poblaciones de TILLING de otras especies: *Solanum lycopersicum* (Minoia y col., 2010), *Pisum sativum* (Dalmais y col., 2008) o en melón en la variedad tipo Charentais, CharMono (Dahmani-Mardas y col., 2010) y que muestran cierta semejanza con fenotipos encontrados en la población mutante desarrollada en este trabajo. La deformación de cotiledones (Figura 4.11, A), la falta de pigmentación o presencia de clorosis en hojas (Figura 4.11, B), así como la malformación en el desarrollo apical de la planta (Figura 4.11, C), son algunos de los fenotipos que más frecuentemente se repiten en las poblaciones de TILLING.

La deformación en cotiledones que observamos en la familia F1876 es prácticamente igual al fenotipo que mostraba el mutante en la variedad CharMono que presentaba cierta deformación y enroscamiento en los cotiledones. El tipo de mutación con coloración albina detectada en la familia (C384) también se ha detectado en las dos poblaciones con las que se comparan los fenotipos, así como la malformación generalizada no sólo en la morfología de las hojas sino también en el crecimiento apical de la familia F1505 (Figura 4.11).

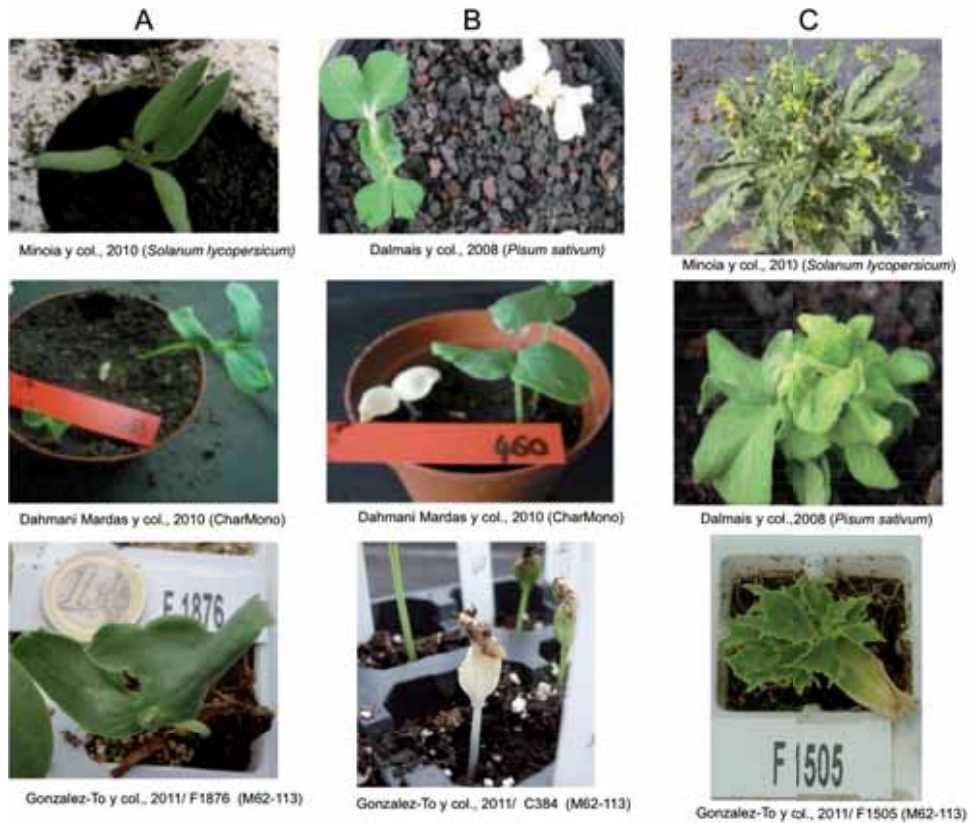


Figura 4.11: Ejemplos de fenotipos encontrados en otras poblaciones de TILLING. A) Malformación de cotiledones (F1876) B) Albinismo, falta de pigmentación (C384) C) Malformación en el desarrollo apical (F1505). Las imágenes se han obtenido de las poblaciones de *Solanum lycopersicum* (Minoia y col., 2010), *Pisum sativum* (Dalmais y col., 2008) y melón variedad tipo Charentais CharMono (Dahmani-Mardas y col., 2010). Finalmente se han añadido algunas imágenes de mutantes de la población de melón (M62-113) obtenidas en este trabajo para poder establecer la comparación (Gonzalez-To y col., 2010).

También podemos encontrar paralelismos entre los fenotipos encontrados en fruto en estas poblaciones de TILLING desarrolladas en otras especies. En la población de melón desarrollada por Dahmani y colaboradores fenotiparon frutos del melón CharMono con un tamaño menor (Figura 4.12, A3) , o con un número mayor de carpelos (Figura 4.12, A2) respecto al control (Figura 4.12, A1). También en la población de TILLING de tomate (Minoia y col., 2010) se fenotiparon frutos que presentaban una forma más alargada (Figura 4.12, B1) o una pigmentación alterada (Figura 4.12, B2). Algunos de estos fenotipos mutantes también se han encontrado en la población desarrollada en este trabajo (Figura 4.6, C1-2).

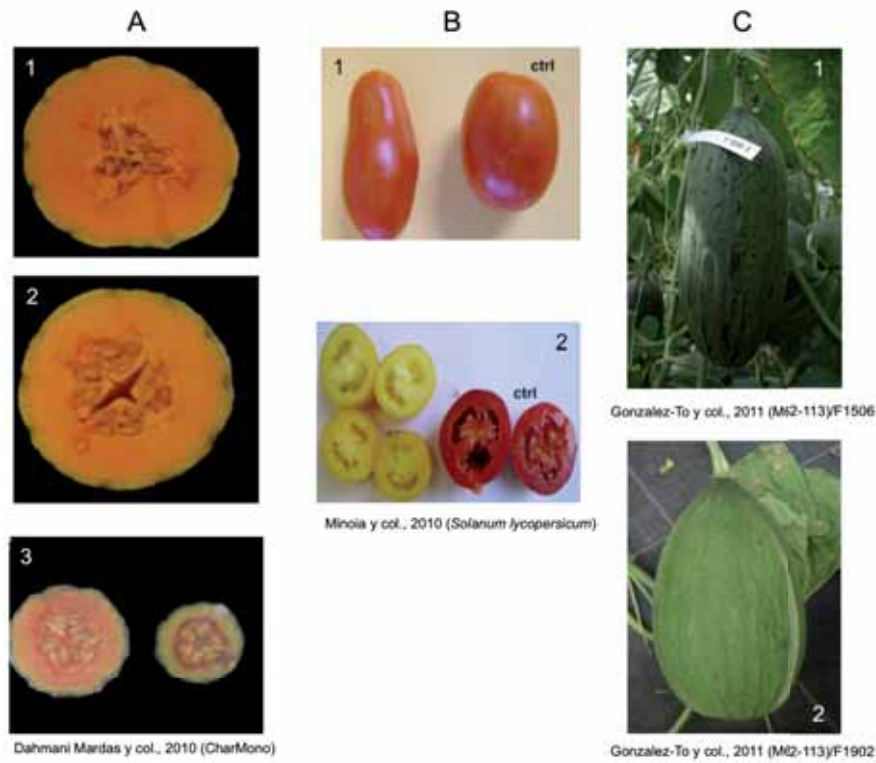


Figura 4.12: Fenotipos identificados en fruto en otras poblaciones de TILLING elaboradas en melón (Dahmani Mardas y col., 2010) y en tomate (Minoia y col., 2010). Se muestran dos ejemplos de mutantes encontrados en la población de TILLING desarrollada en este trabajo (González-To y col., 2011); uno de fruto muy alargado (C1) y otro que no contiene la pigmentación o manchas características del Piel de Sapo (C2).

Por lo tanto, la comparación respecto a los fenotipos mutantes encontrados en otras poblaciones confirma la eficiencia del método empleado para mutagenizar en melón y a su vez, confirma la tendencia a encontrar algunas mutaciones específicas que ya se han encontrado en otras poblaciones, como pueden ser la malformación de cotiledones, la alteración de pigmentación tanto en hojas como en fruto y las malformaciones en el desarrollo apical o axial de la planta.

CONCLUSIONES

1- Se ha desarrollado una población de TILLING en *Cucumis melo* L. con una tasa de mutación de 1/1.5 Mb. Esta nueva herramienta ha demostrado ser útil para la detección de mutantes en genes de interés agronómico.

2-La prehidratación de las semillas así como la utilización de bajas dosis de mutágeno (0,5% al 1,5%) durante el protocolo de mutagenización han determinado, probablemente, la baja tasa de mutación encontrada. Para desarrollar la población del INRA no se prehidrataron las semillas y se usaron dosis más altas del agente mutágeno EMS, obteniendo una tasa de mutación más elevada que en la población IRTA.

3-En la nueva población desarrollada se han analizado cuatro genes: *PDS* y los genes *eIF4E*, *eIF(iso)4E* y *ETR1*. Los mismos genes han sido analizados en la población de TILLING del INRA para poder establecer una comparación entre ambas.

4-En la población IRTA se han detectado dos mutantes en el gen de la *PDS*, cuatro en el *eIF4E*, ninguno en su isoforma (*eIF(iso)4E*) y ninguno en el gen *ETR1*. Cuatro de estos mutantes se encuentran en zona codificante (4/6) y tres de ellos podrían afectar a la funcionalidad de la proteína según los valores obtenidos en SIFT.

5-En la población INRA se han detectado cuatro mutantes en el gen de la *PDS*, catorce en el *eIF4E* y uno en *eIF(iso)4E*. Siete de estos mutantes se encuentran en zona codificante (7/19) y tres de ellos podrían afectar a la funcionalidad de la proteína según los valores obtenidos en SIFT.

6- Se ha realizado el estudio de segregación y de fenotipado para los mutantes detectados en el gen de la *PDS* en la población IRTA. En la familia C384 no se han obtenido homocigotos mutantes en M2 y sólo se ha podido obtener un homocigoto mutante en la generación M3, que presenta un fenotipo albino. Por lo tanto, hemos relacionado este fenotipo con la mutación del gen *PDS*, que representaría una ineficiencia en la vía de biosíntesis de carotenoides.

7-Para los mutantes en eIF4E se han obtenido estudios de segregación en M2 pero no se ha podido hacer un estudio fenotípico. Sin embargo, si se ha hecho un estudio de predicción de estructura 3D, sugiriendo que algunas mutaciones pueden afectar la funcionalidad de la proteína, como por ejemplo la familia F57 con el cambio Glu129Lys.

8-Durante este trabajo también se realizó un estudio de EcoTILLING comparando 113 variedades de melón con la variedad de referencia M62-113. Este ensayo se hizo con los genes *eIF4E* y *ETR1*. En ambos genes se encontraron diferentes haplotipos con SNPs que provocan un cambio de aminoácido y pueden afectar la funcionalidad de ambas proteínas según los valores obtenidos en SIFT. Mediante EcoTILLING se han definido nuevos haplotipos o variaciones naturales que no habían sido descritas con anterioridad.

9- Se ha elaborado un catálogo de mutantes fenotípicos en tres estadios distintos de crecimiento a partir de 400 familias M2: estadio de plántula, de planta adulta y de fruto. En las tres etapas se han descrito y catalogado variaciones fenotípicas que no habían sido descritas con anterioridad en esta especie. Este puede ser un inicio para diseñar una base de datos más compleja y específica conjuntamente con las poblaciones de TILLING de melón desarrolladas posteriormente a este trabajo.

Bibliografía

Agudelo-Romero, P., Carbonell, P., de la Iglesia, F., Carrera, J., Rodrigo, G. y col. (2008). "Changes in the gene expression profile of *Arabidopsis thaliana* after infection with Tobacco etch virus." *Virology* 5: 92.

Akasaka-Kennedy, Y., Tomita, K., Ezura, H. (2004). "Efficient plant regeneration and *Agrobacterium*-mediated transformation via somatic embryogenesis in melon (*Cucumis melo* L.)." *Plant Science Elsevier* 166(3): 763-769.

Al-Babili C., Hartung, W. (1999). "CPTA Modulates Levels of Carotenogenic Proteins and their mRNAs and Affects Carotenoid and ABA Content as well as Chromoplast Structure in *Narcissus pseudonarcissus* Flowers." *Plant Biology* 1(6): 607-612.

Allard, R.W. (1999). "Principles of Plant Breeding." Printed by John Wiley & Sons, Inc.

An, S., Park, S., Jeong, D.H., Lee, D.Y., Kang, H.G. y col. (2003). "Generation and analysis of end sequence database for T-DNA tagging lines in rice." *Plant Physiology* 133(4): 2040-7.

Arumuganathan, E. (1991). "Nuclear DNA content of some important plant species." *Plant Mol Bio Rep* 9: 211-215.

Ashraf, M., Bray, M. (1993). "DNA synthesis in osmoprimed leek (*Allium porrum* L.) seeds and evidence for repair and replication." *Seed Science Research* 3(15-23).

Auerbach, C., Robson, J.M. (1946). "Chemical production of mutations." *Nature* 157: 302.

Boch, J. y Bonas, U. (2010). "Xanthomonas AvrBs3 family-type III effectors: Discovery and function. *Annu. Rev. Phyto.* 48: 419-436.

Baltimore, D., Smoler, D. (1971). "Primer requirement and template specificity of the DNA polymerase of RNA tumor viruses." *Proc Natl Acad Sci U S A* 68(7): 1507-11.

Barkley, N.A., Wang, M.L. (2008). "Application of TILLING and EcoTILLING as Reverse Genetic Approaches to Elucidate the Function of Genes in Plants and Animals." *Curr Genomics* 9(4): 212-26.

Barnes, W.M. (1994). "PCR amplification of up to 35-kb DNA with high fidelity and high yield from lambda bacteriophage templates." *Proc Natl Acad Sci U S A* 91(6): 2216-2220.

Bartley, G.E., Scolnik, P.A., Beyer, P. (1999). "Two *Arabidopsis thaliana* carotene desaturases, phytoene desaturase and zeta-carotene desaturase, expressed in *Escherichia coli*, catalyze a poly-cis pathway to yield pro-lycopene." *Eur J Biochem* 259(1-2): 396-403.

- Baudracco-Arnas, S., Pitrat, M. (1996). "A genetic map of melon (*Cucumis melo* L.) with RFLP, RAPD, isozyme, disease resistance and morphological markers." *Theor Appl Genet* 93: 57-64.
- Becker, J., Vos, P., Kuiper, M., Salamini, F., Heun, M. (1995). "Combined mapping of AFLP and RFLP markers in barley." *Mol Gen Genet* 249(1): 65-73.
- Bent, A.F. (1996). "Plant Disease Resistance Genes: Function Meets Structure." *Plant Cell* 8(10): 1757-1771.
- Bentley, A., MacLennan, (2000). "Targeted recovery of mutations in *Drosophila*." *Genetics* 156(3): 1169-1173.
- Birnboim, H.C., Doly, J. (1979). "A rapid alkaline extraction procedure for screening recombinant plasmid DNA." *Nucleic Acids Res* 7(6): 1513-23.
- Blanca, J., Canizares, J., Roig, C., Ziarsolo, P., Nuez, F. y col. (2011). "Transcriptome characterization and high throughput SSRs and SNPs discovery in *Cucurbita pepo* (Cucurbitaceae)." *BMC Genomics* 12: 104.
- Blanca J, Cañizares J, Roig C, Ziarsolo P, Nuez F, Picó B. (2011). "Transcriptome characterization and high throughput SSRs and SNPs discovery in *Cucurbita pepo* (Cucurbitaceae)." *BMC Genomics*. 12:104.
- Blanca, J., Esteras, C., Ziarsolo, P., Pérez, D., Fernã Ndez-Pedrosa, V., y col. (2012). "Transcriptome sequencing for SNP discovery across *Cucumis melo*". *BMC Genomics* 13:280.
- Bleecker, A.B. (1999). "Ethylene perception and signaling: an evolutionary perspective." *Trends Plant Sci* 4(7): 269-274.
- Bleecker, A.B., Kende, H. (2000). "Ethylene: a gaseous signal molecule in plants." *Annu Rev Cell Dev Biol* 16: 1-18.
- Bohn-Courseau, I. (2010). "Auxin: a major regulator of organogenesis." *C R Biol* 333(4): 290-6.
- Bono, C., Nuzzo, D., Albeggiani, G., Zizzo, C., Francofonte, D. y col. (2011). "Genetic screening of Fabry patients with EcoTILLING and HRM technology." *BMC Res Notes* 4: 323.
- Borevitz, J.O., Liang, D., Plouffe, D., Chang, H.S., Zhu, T. y col. (2003). "Large-scale identification of single-feature polymorphisms in complex genomes." *Genome Res* 13(3): 513-23.
- Boualem, A., Fergany, M., Fernandez, R., Troadec, C., Martin, A. y col. (2008). "A conserved mutation in an ethylene biosynthesis enzyme leads to andromonoecy in melons." *Science* 321(5890): 836-8.
- Brandwagt, B.F., Kneppers, T.J., Nijkamp, H.J., Hille, J. (2002). "Overexpression of the tomato *Asc-1* gene mediates high insensitivity to AAL toxins and fumonisin B1 in tomato hairy roots and confers resistance to *Alternaria alternata* f. sp. *lycopersici* in *Nicotiana umbratica* plants." *Mol Plant Microbe Interact* 15(1): 35-42.
- Brookes, A.J. (1999). "The essence of SNPs." *Gene* 234(2): 177-186.

Brotman, Y., Silberstein, L., Kovalski, I., Perin, C., Dogimont, C. y col. (2002). "Resistance gene homologues in melon are linked to genetic loci conferring disease and pest resistance." *Theor Appl Genet* 104(6-7): 1055-1063.

Browning, K.S., Webster, C., Roberts, J.K., Ravel, J.M. (1992). "Identification of an isozyme form of protein synthesis initiation factor 4F in plants." *J Biol Chem* 267(14): 10096-100.

Bruggemann, E., Handwerger, K., Essex, C., Storz, G. (1996). "Analysis of fast neutron-generated mutants at the *Arabidopsis thaliana* HY4 locus." *Plant J* 10(4): 755-60.

Brummell, D.A., Harpster, M.H. (2001). "Cell wall metabolism in fruit softening and quality and its manipulation in transgenic plants." *Plant Mol Biol* 47(1-2): 311-40.

Budiman, M.A., Mao, L., Wood, T.C., Wing, R.A. (2000). "A deep-coverage tomato BAC library and prospects toward development of an STC framework for genome sequencing." *Genome Res* 10(1): 129-36.

Bulow, L., Schindler, M., Choi, C., Hehl, R. (2004). "PathoPlant: a database on plant-pathogen interactions." *In Silico Biol* 4(4): 529-36.

Buschges, R., Hollricher, K., Panstruga, R., Simons, G., Wolter, M. y col. (1997). "The barley Mlo gene: a novel control element of plant pathogen resistance." *Cell* 88(5): 695-705.

Bush, S.M., Krysan, P.J. (2010). "iTILLING: a personalized approach to the identification of induced mutations in *Arabidopsis*." *Plant Physiol* 154(1): 25-35.

Caldwell, D.G., McCallum, N., Shaw, P., Muehlbauer, G.J., Marshall, D.F. y col. (2004). "A structured mutant population for forward and reverse genetics in Barley (*Hordeum vulgare* L.)." *The Plant Journal* 40(1): 143-150.

Carol, P., Kuntz, M. (2001). "A plastid terminal oxidase comes to light: implications for carotenoid biosynthesis and chlororespiration." *Trends Plant Sci* 6: 31-36.

Castagna, A., Ederli, L., Pasqualini, S., Mensuali-Sodi, A., Baldan, B. y col. (2007). "The tomato ethylene receptor LE-ETR3 (NR) is not involved in mediating ozone sensitivity: causal relationships among ethylene emission, oxidative burst and tissue damage." *New Phytol* 174(2): 342-56.

Cecchini, E., Mulligan, B.J., Covey, S.N., Milner, J.J. (1998). "Characterization of gamma irradiation-induced deletion mutations at a selectable locus in *Arabidopsis*." *Mutat Res* 401(1-2): 199-206.

Cermak, T., Doyle, E., Christian, M., Waig, L., Zhang, Y., Schmidt, J.A., y col. (2011). Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Research* 39:82.

Chakravorty, D., Trusov, Y., Botella, J.R. (2011). "Site-directed mutagenesis of the *Arabidopsis* heterotrimeric G protein beta subunit suggests divergent mechanisms of effector activation between plant and animal G proteins." *Planta* 235(3): 615-27.

- Chang, C., Kwok, S.F., Bleecker, A.B., Meyerowitz, E.M. (1993). "Arabidopsis ethylene-response gene ETR1: similarity of product to two-component regulators." *Science* 262(5133): 539-44.
- Chawade, A., Sikora, P., Brautigam, M., Larsson, M., Vivekanand, V. y col. (2010). "Development and characterization of an oat TILLING-population and identification of mutations in lignin and beta-glucan biosynthesis genes." *BMC Plant Biol* 10: 86.
- Chen, J.F., Gallie, D.R. (2010). "Analysis of the functional conservation of ethylene receptors between maize and Arabidopsis." *Plant Mol Biol* 74(4-5): 405-21.
- Chisholm, S.T., Coaker, G., Day, B., Staskawicz, B.J. (2006). "Host-microbe interactions: shaping the evolution of the plant immune response." *Cell* 124(4): 803-14.
- Choi, I.Y., Hyten, D.L., Matukumalli, L.K., Song, Q., Chaky, J.M. y col. (2007). "A soybean transcript map: gene distribution, haplotype and single-nucleotide polymorphism analysis." *Genetics* 176(1): 685-96.
- Chovelon, V., Restier, V., Giovinazzo, N., Dogimont, C., Aarouf, J. (2011). "Histological study of organogenesis in Cucumis melo L. after genetic transformation: why is it difficult to obtain transgenic plants?" *Plant Cell Rep* 30(11): 2001-11.
- Clark, K.A., Krysan, P.J. (2007). "Protocol: An improved high-throughput method for generating tissue samples in 96-well format for plant genotyping (Ice-Cap 2.0)." *Plant Methods* 3: 8.
- Colbert, T., Till, B.J., Tompa, R., Reynolds, S., Steine, M.N. y col. (2001). "High-Throughput Screening for Induced Point Mutations." *Plant Physiol* 126: 480 - 484.
- Collinge, M., Boller, T. (2001). "Differential induction of two potato genes, *Stprx2* and *StNAC*, in response to infection by *Phytophthora infestans* and to wounding." *Plant Mol Biol* 46(5): 521-9.
- Comai, L., Henikoff, S. (2006). "TILLING: practical single-nucleotide mutation discovery." *Plant J* 45(4): 684-94.
- Comai, L., Young, K., Till, B.J., Reynolds, S.H., Greene, E.A. y col. (2004). "Efficient discovery of DNA polymorphisms in natural populations by Ecotilling." *Plant J* 37(5): 778-86.
- Cooper, J.L., Till, B.J., Laport, R.G., Darlow, M.C., Kleffner, J.M. y col. (2008). "TILLING to detect induced mutations in soybean." *BMC Plant Biol* 8: 9.
- Corbacho, J., Romojaro, F., Pech, J-C., Latché, A., Gomez-Jimenez, MC. (2013). "Transcriptomic Events Involved in Melon Mature-Fruit Abscission Comprise the Sequential Induction of Cell-Wall Degrading Genes Coupled to a Stimulation of Endo and Exocytosis". *PLoS ONE* 8(3): e58363.
- Dahmani-Mardas, F., Troadec, C., Boualem, A., Leveque, S., Alsadon, A.A. y col. (2010). "Engineering Melon Plants with Improved Fruit Shelf Life Using the TILLING Approach." *PLoS One* 5(12): e15776.

- Dalmais, M., Schmidt, J., Le Signor, C., Moussy, F., Burstin, J. y col. (2008). "UTILdb, a *Pisum sativum* in silico forward and reverse genetics tool." *Genome Biol* 9(2): R43.
- Danin-Poleg, Y., Reis, N., Baudracco-Arnas, S., Pitrat, M., Staub, J.E. y col. (2000). "Simple sequence repeats in *Cucumis* mapping and map merging." *Genome* 43(6): 963-74.
- Deleu, W., González, V., Monfort, A., Bendahmane, A., Puigdomenech, P. y col. (2007). "Structure of two melon regions reveals high microsynteny with sequenced plant species." *Mol Genet Genomics* 278(6): 611-22.
- Deleu, W., Esteras, C., Roig, C., González-To, M., Fernández-Silva, I. y col. (2009). "A set of EST-SNPs for map saturation and cultivar identification in melon." *BMC Plant Biol* 9: 90.
- Denny, T.P. (1995). "Involvement of bacterial polysaccharides in plant pathogenesis." *Annu Rev Phytopathol* 33: 173-97.
- Diaz, A., Fergany, M., Formisano, G., Ziarsolo, P., Blanca, J. y col. (2011). "A consensus linkage map for molecular markers and quantitative trait loci associated with economically important traits in melon (*Cucumis melo* L.)." *BMC Plant Biol* 11: 111.
- Diaz, C., Purdy, S., Christ, A., Morot-Gaudry, J.F., Wingler, A. y col. (2005). "Characterization of markers to determine the extent and variability of leaf senescence in *Arabidopsis*. A metabolic profiling approach." *Plant Physiol* 138(2): 898-908.
- Dodgson, J.B., Wells, R.D. (1977). "Action of single-strand specific nucleases on model DNA heteroduplexes of defined size and sequence." *Biochemistry* 16(11): 2374-9.
- Dong, C., Vincent, K., Sharp, P. (2009). "Simultaneous mutation detection of three homoeologous genes in wheat by High Resolution Melting analysis and Mutation Surveyor." *BMC Plant Biol* 9: 143.
- Doyle, J., Doyle, J. (1990). "Isolation of plant DNA from fresh tissue." *Focus* 12: 13-15.
- Draper, B.W., McCallum, C.M., Stout, J.L., Slade, A.J., Moens, C.B. (2004). "A high-throughput method for identifying N-ethyl-N-nitrosourea (ENU)-induced point mutations in zebrafish." *Methods Cell Biol* 77: 91-112.
- ENCODE Consortium (2004). "The ENCODE (ENCyclopedia Of DNA Elements) Project." *Science* 306(5696): 636-40.
- Ezura H, Yuhashi, K., Yasuta T. (2000). "Effect of ethylene on *Agrobacterium tumefaciens*-mediated gene transfer to melon." *Plant Breed* 119: 75-79.
- Fang, G., Grumet, R. (1990). "*Agrobacterium tumefaciens* mediated transformation and regeneration of muskmelon plants." *Plant Cell Rep* 9: 160-164.
- Fernández-Aparicio (2010). "Parasitic plant infection is partially controlled through symbiotic pathways." *Weed Research* 50: 1.
- Fernandez-Silva, I., Eduardo, I., Blanca, J., Esteras, C., Pico, B. y col. (2008). "Bin mapping of genomic and EST-derived SSRs in melon (*Cucumis melo* L.)." *Theor Appl Genet* 118(1): 139-50.

- Flor, H.H. (1971). "Current status gene-for-gene concept." *Ann. Rev. Phytopathol.* 9: 275-296.
- Flores, F., El Yahyaoui, F., de Billerbeck, G., Romojaro, F., Latche, A. y col. (2002). "Role of ethylene in the biosynthetic pathway of aliphatic ester aroma volatiles in Charentais Cantaloupe melons." *J Exp Bot* 53(367): 201-6.
- Flores, F.B. (2001). Fisiología y maduración del melón Cantaloupe Charentais con un gen antisentido de la enzima ACC oxidasa. . Tesis, Universidad de Murcia. Murcia.
- Fraser (1986). "Genes for resistance to plant viruses." *CRC Crit. Rev. Plant Sci.*: 257-294.
- Fu, D.Q., Zhu, B.Z., Zhu, H.L., Jiang, W.B., Luo, Y.B. (2005). "Virus-induced gene silencing in tomato fruit." *Plant J* 43(2): 299-308.
- Fuhrmann, M., Oertel, W., Berthold, P., Hegemann, P. (2005). "Removal of mismatched bases from synthetic genes by enzymatic mismatch cleavage." *Nucleic Acids Res* 33(6): e58.
- Gady, A.L., Hermans, F.W., Van de Wal, M.H., Van Loo, E.N., Visser, R.G. y col. (2009). "Implementation of two high through-put techniques in a novel application: detecting point mutations in large EMS mutated plant populations." *Plant Methods* 5: 13.
- Gamble, R.L., Coonfield, M.L., Schaller, G.E. (1998). "Histidine kinase activity of the ETR1 ethylene receptor from Arabidopsis." *Proc Natl Acad Sci U S A* 95(13): 7825-9.
- Gamble, R.L., Qu, X., Schaller, G.E. (2002). "Mutational analysis of the ethylene receptor ETR1. Role of the histidine kinase domain in dominant ethylene insensitivity." *Plant Physiol* 128(4): 1428-38.
- Gao, Z., Johansen, E., Evers, S., Thomas, C.L., Noel Ellis, T.H. y col. (2004). "The potyvirus recessive resistance gene, sbm1, identifies a novel role for translation initiation factor eIF4E in cell-to-cell trafficking." *Plant J* 40(3): 376-85.
- Garcia-Mas J, Oliver, M., Gómez H y Vicente MC (2000). "AFLP, RAPD and RFLP markers to measure genetic diversity in melon." *Theor Appl Genet* 101: 860-864.
- Garcia-Mas, J., Benjak, A., Sanseverino, W., Bourgeois, M., Mir, G. y col. (2012). "The genome of melon (*Cucumis melo* L.)." *Proc Natl Acad Sci U S A* 109(29): 11872-7.
- Genard, M., Gouble, B. (2005). "ETHY. A theory of fruit climacteric ethylene emission." *Plant Physiol* 139(1): 531-45.
- Gepts (2002). "A comparison between Crop Domestication, Classical Plant Breeding and Genetic Engineering." *CS* 42(6): 1780-1790.
- Gilchrist, E.J., Haughn, G.W. (2005). "TILLING without a plough: a new method with applications for reverse genetics." *Current Opinion in Plant Biology* 8: 211 - 215.
- Gilchrist, E., O'Neil, N., Rose, A., Zetka, M., Haughn, G. (2006). "TILLING is an effective reverse genetics technique for *Caenorhabditis elegans*." *BMC Genomics* 7(1): 262.

- Gingras, A.C., Raught, B., Sonenberg, N. (1999). "eIF4 initiation factors: effectors of mRNA recruitment to ribosomes and regulators of translation." *Annu Rev Biochem* 68: 913-63.
- Giovannoni, J. (2001). "Molecular Biology of Fruit Maturation and Ripening." *Annu Rev Plant Physiol Plant Mol Biol* 52: 725-749.
- Giovannoni, J. (2007). "Fruit ripening mutants yield insights into ripening control." *Curr Opin Plant Biol* 10(3): 283-9.
- Giuliano, G., Bartley, G.E., Scolnik, P.A. (1993). "Regulation of carotenoid biosynthesis during tomato development." *Plant Cell* 5(4): 379-87.
- González, V.M., Benjak, A., Henaff, E.M., Mir, G., Casacuberta, J.M. y col. (2010 a). "Sequencing of 6.7 Mb of the melon genome using a BAC pooling strategy." *BMC Plant Biol* 10: 246.
- González, V.M., Garcia-Mas, J., Arús, P., Puigdomenech, P. (2010 b). "Generation of a BAC-based physical map of the melon genome." *BMC Genomics* 11: 339.
- González-Ibeas, D., Blanca, J., Roig, C., Gonzalez-To, M., Pico, B. y col. (2007). "MELOGEN: an EST database for melon functional genomics." *BMC Genomics* 8: 306.
- Gonzalez-Ibeas, D., Canizares, J., Aranda, M.A. (2012). "Microarray analysis shows that recessive resistance to Watermelon mosaic virus in melon is associated with the induction of defense response genes." *Mol Plant Microbe Interact* 25(1): 107-18.
- González-To, M., Xu, M., Esteras, C., Roig, C., Monforte, A.J. y col. (2011). "Towards a TILLING platform for functional genomics in Piel de Sapo melons." *BMC Res Notes* 4: 289.
- Gonzalo, M.J., Oliver, M., Garcia-Mas, J., Monfort, A., Dolcet-Sanjuan, R. y col. (2005). "Simple-sequence repeat markers used in merging linkage maps of melon (*Cucumis melo* L.)." *Theor Appl Genet* 110(5): 802-11.
- Greenberg, J.T., Yao, N. (2004). "The role and regulation of programmed cell death in plant-pathogen interactions." *Cell Microbiol* 6(3): 201-11.
- Greene, E.A., Codomo, C.A., Taylor, N.E., Henikoff, J.G., Till, B.J. y col. (2003). "Spectrum of chemically induced mutations from a large-scale reverse-genetic screen in *Arabidopsis*." *Genetics* 164(2): 731-40.
- Grotewold, E. (2006). "The genetics and biochemistry of floral pigments." *Annu Rev Plant Biol* 57: 761-80.
- Gruenbaum, Y., Naveh-Many, T., Cedar, H., Razin, A. (1981). "Sequence specificity of methylation in higher plant DNA." *Nature* 292(5826): 860-2.
- Guis M, Rousan, J., Dogimont, C., Pitrat, M., Pech, JC. (1998). "Melon biotechnology." *Biotechnol Gent Eng Rev* 15: 289-311.
- Guzman, P., Ecker, J.R. (1990). "Exploiting the triple response of *Arabidopsis* to identify ethylene-related mutants." *Plant Cell* 2(6): 513-23.

- Hamada H., Petrino, M.G., Kakunaga, T. (1982). "A novel repeated element with Z-DNA- forming potential is widely found in evolutionary diverse eukariotic genomes." *Proc Natl Acad Sci U S A* 79: : 6465-6469.
- Handel, S. (1982). "Dynamics of gene flow in an experimental population of *Cucumis melo* " *Amer. J. Bot* 69(10): 1538-1546.
- Hase, I., Imaoka, S., Oda, Y., Hiroi, T., Nakamoto, T. y col. (2000). "Area under the plasma concentration-time curve of inorganic fluoride following sevoflurane anesthesia correlates with CYP2E1 mRNA level in mononuclear cells." *Anesthesiology* 92(6): 1661-6.
- Henikoff, S., Comai, L. (2003). "Single-nucleotide mutations for plant functional genomics." *Annu Rev Plant Biol* 54: 375-401.
- Hirochika, H., Guiderdoni, E., An, G., Hsing, Y.I., Eun, M.Y. y col. (2004). "Rice mutant resources for gene discovery." *Plant Mol Biol* 54(3): 325-34.
- Hirschberg, J. (2001). "Carotenoid biosynthesis in flowering plants." *Current Opinion in Plant Biology* 4(3): 210-218.
- Hospital, F. (2009). "Challenges for effective marker-assisted selection in plants." *Genetica* 136(2): 303-10.
- Hua, J., Chang, C., Sun, Q., Meyerowitz, E.M. (1995). "Ethylene insensitivity conferred by Arabidopsis ERS gene." *Science* 269(5231): 1712-4.
- Iehisa, J.C., Shimizu, A., Sato, K., Nasuda, S., Takumi, S. (2012). "Discovery of High-Confidence Single Nucleotide Polymorphisms from Large-Scale De Novo Analysis of Leaf Transcripts of *Aegilops tauschii*, A Wild Wheat Progenitor." *DNA Res.*
- Isaacson, T., Ronen, G., Zamir, D., Hirschberg, J. (2002). "Cloning of tangerine from tomato reveals a carotenoid isomerase essential for the production of beta-carotene and xanthophylls in plants." *Plant Cell* 14(2): 333-42.
- Izquierdo, M. (2001). "Ingeniería genética y transferencia génica". Ediciones Pirámide S.A. Madrid.
- Jeffrey, D. (1990.). Capítulo: An outline classification of the Cucurbitaceae. En: Bates, D.M., Robinson, R.W., Jeffrey, C. "Biology and utilization of the Cucurbitaceae". Ithaca and London Cornell University: p.449-463. 485p.
- Joobeur, T., King, J.J., Nolin, S.J., Thomas, C.E., Dean, R.A. (2004). "The Fusarium wilt resistance locus Fom-2 of melon contains a single resistance gene with complex features." *Plant J* 39(3): 283-97.
- Katzir, N., Danin-Poleg, Y., Tzuri, G., Karchi, Z. (1996). "Length polymorphism and homologies of microsatellites in several Cucurbitaceas species." *Theor Appl Genet* 93: 1282-1290.
- Keen, N.T. (1990). "Gene-for-gene complementarity in plant-pathogen interactions." *Annu Rev Genet* 24: 447-63.
- Kendall, T.J.N. (1988). "Genetic-variation of ethylene production in harvested muskmelon fruits." *HortScience* 23: 759-761.

- Klee, H., Tieman, D. (1999). "Differential Expression of Two Novel Members of the Tomato Ethylene-Receptor Family." *Plant Physiol* 120: 165-172.
- Knoll, J.E., Ramos, M.L., Zeng, Y., Holbrook, C.C., Chow, M. y col. (2011). "TILLING for allergen reduction and improvement of quality traits in peanut (*Arachis hypogaea* L.)." *BMC Plant Biol* 11: 81.
- Kodym, A., Afza, R. (2003). "Physical and chemical mutagenesis." *Methods Mol Biol* 236: 189-204.
- Konieczny, A., Ausubel, FM. (1993). "A procedure for mapping *Arabidopsis* mutations using co-dominant ecotype-specific PCR-based markers." *Plant Journal* 4: 403-410.
- Koornneef, M., Dellaert, L.W., Van der Veen, J.H. (1982). "EMS- and radiation-induced mutation frequencies at individual loci in *Arabidopsis thaliana* (L.) Heynh." *Mutat Res* 93(1): 109-23.
- Krysan, P. (2004). "Ice-cap. A high-throughput method for capturing plant tissue samples for genotype analysis." *Plant Physiol* 135(3): 1162-9.
- Kuang, H., Padmanabhan, C., Li, F., Kamei, A., Bhaskar, P.B. y col. (2009). "Identification of miniature inverted-repeat transposable elements (MITEs) and biogenesis of their siRNAs in the Solanaceae: new functional implications for MITEs." *Genome Res* 19(1): 42-56.
- Kumar, S., Lamarche, B.J., Tsai, M.D. (2007). "Use of damaged DNA and dNTP substrates by the error-prone DNA polymerase X from African swine fever virus." *Biochemistry* 46(12): 3814-25.
- Lanahan, M.B., Yen, H.C., Giovannoni, J., Klee, H.J. (1994). "The never ripe mutation blocks ethylene perception in tomato." *Plant Cell* 6(4): 521-30.
- Le Signor, C., Savoio, V., Aubert, G., Verdier, J., Nicolas, M. y col. (2009). "Optimizing TILLING populations for reverse genetics in *Medicago truncatula*." *Plant Biotechnol J* 7(5): 430-41.
- Lehman, A., Black, R., Ecker, J.R. (1996). "HOOKLESS1, an ethylene response gene, is required for differential cell elongation in the *Arabidopsis* hypocotyl." *Cell* 85(2): 183-94.
- Lellis, A.D., Kasschau, K.D., Whitham, S.A., Carrington, J.C. (2002). "Loss-of-susceptibility mutants of *Arabidopsis thaliana* reveal an essential role for eIF(iso)4E during potyvirus infection." *Curr Biol* 12(12): 1046-51.
- Leonard, S., Plante, D., Wittmann, S., Daigneault, N., Fortin, M.G. y col. (2000). "Complex formation between potyvirus VPg and translation eukaryotic initiation factor 4E correlates with virus infectivity." *J Virol* 74(17): 7730-7.
- Lewis, R., Roberts W., Alberts (2006). "La Célula." Ediciones Omega. Madrid.
- Li, X., Song, Y., Century, K., Straight, S., Ronald, P. y col. (2001). "A fast neutron deletion mutagenesis-based reverse genetics system for plants." *Plant J* 27(3): 235-42.
- Li, X., Zhang, Y. (2002). "Reverse genetics by fast neutron mutagenesis in higher plants." *Funct Integr Genomics* 2(6): 254-8.

- Li, Z.H., Matthews, P.D., Burr, B., Wurtzel, E.T. (1996). "Cloning and characterization of a maize cDNA encoding phytoene desaturase, an enzyme of the carotenoid biosynthetic pathway." *Plant Mol Biol* 30(2): 269-79.
- Linden, H., Misawa, N., Chamovitz, D., Pecker, I., Hirschberg, J. y col. (1991). "Functional complementation in *Escherichia coli* of different phytoene desaturase genes and analysis of accumulated carotenenes." *Z Naturforsch C* 46(11-12): 1045-51.
- López-Lara, I.M., Sohlenkamp, C., Geiger, O. (2006). "The post-genomic era-- what comes next?" *Rev Latinoam Microbiol* 48(1): 44-5.
- López-Raez, J.A., Flors, V., Garcia, J.M., Pozo, M.J. (2011). "AM symbiosis alters phenolic acid content in tomato roots." *Plant Signal Behav* 5(9): 1138-40.
- Lu, R., Martin-Hernandez, A.M., Peart, J.R., Malcuit, I., Baulcombe, D.C. (2003). "Virus-induced gene silencing in plants." *Methods* 30(4): 296-303.
- Lundberg, K.S., Shoemaker, D.D., Adams, M.W., Short, J.M., Sorge, J.A. y col. (1991). "High-fidelity amplification using a thermostable DNA polymerase isolated from *Pyrococcus furiosus*." *Gene* 108(1): 1-6.
- Luo, M., Wang, Y.H., Frisch, D., Joobeur, T., Wing, R.A. y col. (2001). "Melon bacterial artificial chromosome (BAC) library construction using improved methods and identification of clones linked to the locus conferring resistance to melon *Fusarium* wilt (Fom-2)." *Genome* 44(2): 154-62.
- Manzoor, B., Cheema, A. (2003). "Radiosensitivity studies in basmati rice." *Pak. J. Bot* 35(2): 197-207.
- Martin, A., Troadec, C., Boualem, A., Rajab, M., Fernandez, R. y col. (2009). "A transposon-induced epigenetic change leads to sex determination in melon." *Nature* 461(7267): 1135-8.
- Martin, G.B., Bogdanove, A.J., Sessa, G. (2003). "Understanding the functions of plant disease resistance proteins." *Annu Rev Plant Biol* 54: 23-61.
- Mascarell-Creus, A., Canizares, J., Vilarrasa-Blasi, J., Mora-Garcia, S., Blanca, J. y col. (2009). "An oligo-based microarray offers novel transcriptomic approaches for the analysis of pathogen resistance and fruit quality traits in melon (*Cucumis melo* L.)." *BMC Genomics* 10: 467.
- Mashal, R.D., Koontz, J., Sklar, J. (1995). "Detection of mutations by cleavage of DNA heteroduplexes with bacteriophage resolvases." *Nat Genet* 9(2): 177-83.
- McCallum, C.M., Comai, L., Greene, E.A., Henikoff, S. (2000 a). "Targeted screening for induced mutations." *Nat Biotechnol* 18(4): 455-7.
- McCallum, C.M., Comai, L., Greene, E.A., Henikoff, S. (2000 b). "Targeting induced local lesions IN genomes (TILLING) for plant functional genomics." *Plant Physiol* 123(2): 439-42.
- McCarty, D.R., Settles, A.M., Suzuki, M., Tan, B.C., Latshaw, S. y col. (2005). "Steady-state transposon mutagenesis in inbred maize." *Plant J* 44(1): 52-61.
- McDowell, J.M., Woffenden, B.J. (2003). "Plant disease resistance genes: recent insights and potential applications." *Trends Biotechnol* 21(4): 178-83.

McMillin, D.E., Scandalios, J.G. (1983). "Genetic analysis of the duplicated mitochondrial and cytosolic malate dehydrogenase isozymes in maize." *Isozymes Curr Top Biol Med Res* 8: 67-90.

Meinke, D.W., Cherry, J.M., Dean, C., Rounsley, S.D., Koornneef, M. (1998). "Arabidopsis thaliana: a model plant for genome analysis." *Science* 282(5389): 662, 679-82.

Meyer, J.D., Deleu, W., Garcia-Mas, J., Havey, M.J. (2008). "Construction of a fosmid library of cucumber (*Cucumis sativus*) and comparative analyses of the eIF4E and eIF(iso)4E regions from cucumber and melon (*Cucumis melo*)." *Mol Genet Genomics* 279(5): 473-80.

Meyers, B.C., Dickerman, A.W., Michelmore, R.W., Sivaramakrishnan, S., Sobral, B.W. y col. (1999). "Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide-binding superfamily." *Plant J* 20(3): 317-32.

Minoia, S., Petrozza, A., D'Onofrio, O., Piron, F., Mosca, G. y col. (2010). "A new mutant genetic resource for tomato crop improvement by TILLING technology." *BMC Res Notes* 3: 69.

Monforte, A.J., Oliver, M., Gonzalo, M.J., Alvarez, J.M., Dolcet-Sanjuan, R. y col. (2004). "Identification of quantitative trait loci involved in fruit quality traits in melon (*Cucumis melo* L.)." *Theor Appl Genet* 108(4): 750-8.

Montelone, B.A. (1998). "Mutation, Mutagens, and DNA Repair." Editorial: *Human Genetics* Editor. Kansas State University. Kansas.

Moore, S., Vrebalov, J., Payton, P., Giovannoni, J. (2002). "Use of genomics tools to isolate key ripening genes and analyse fruit maturation in tomato." *J Exp Bot* 53(377): 2023-30.

Moreno, E., Obando, J.M., Dos-Santos, N., Fernandez-Trujillo, J.P., Monforte, A.J. y col. (2008). "Candidate genes and QTLs for fruit ripening and softening in melon." *Theor Appl Genet* 116(4): 589-602.

Morgante, M., Olivieri, A.M. (1993). "PCR-amplified microsatellites as a markers in plant genetics." *Plant J* 3: 175-185.

Nambara, E., Keith, K., McCourt, P., Naito, S. (1994). "Isolation of an internal deletion mutant of the Arabidopsis thaliana ABI3 gene." *Plant Cell Physiol* 35(3): 509-13.

Nicaise, V., German-Retana, S., Sanjuan, R., Dubrana, M.P., Mazier, M. y col. (2003). "The eukaryotic translation initiation factor 4E controls lettuce susceptibility to the Potyvirus Lettuce mosaic virus." *Plant Physiol* 132(3): 1272-82.

Nieto, C., Morales, M., Orjeda, G., Clepet, C., Monfort, A. y col. (2006). "An eIF4E allele confers resistance to an uncapped and non-polyadenylated RNA virus in melon." *Plant J* 48(3): 452-62.

Nieto, C., Piron, F., Dalmais, M., Marco, C.F., Moriones, E. y col. (2007). "EcoTILLING for the identification of allelic variants of melon eIF4E, a factor that controls virus susceptibility." *BMC Plant Biol* 7: 34.

- Nishiyama, K., Guis, M., Rose, J.K., Kubo, Y., Bennett, K.A. y col. (2007). "Ethylene regulation of fruit softening and cell wall disassembly in Charentais melon." *J Exp Bot* 58(6): 1281-90.
- Nonaka, S., Yuhashi, K., Takada, K., Sugawara, M., Minamisawa, K. (2008). "Ethylene production in plants during gene transformation suppresses vir gene expression in *Agrobacterium tumefaciens*." *New Phytologist* 178: 647-656.
- Nunez-Palenius, H.G., Cantliffe, D.J., Huber, D.J., Ciardi, J., Klee, H.J. (2006). "Transformation of a muskmelon 'Galia' hybrid parental line (*Cucumis melo* L. var. *reticulatus* Ser.) with an antisense ACC oxidase gene." *Plant Cell Rep* 25(3): 198-205.
- Nunez-Palenius, H.G., Gomez-Lim, M., Ochoa-Alejo, N., Grumet, R., Lester, G. y col. (2008). "Melon fruits: genetic diversity, physiology, and biotechnology features." *Crit Rev Biotechnol* 28(1): 13-55.
- Oleykowski, C.A., Bronson Mullins, C.R., Godwin, A.K., Yeung, A.T. (1998). "Mutation detection using a novel plant endonuclease." *Nucleic Acids Res* 26(20): 4597-602.
- Oliver, M., Garcia-Mas, J., Cardús, M., Pueyo, N., Lopez-Sese, A.L. y col. (2001). "Construction of a reference linkage map for melon." *Genome* 44(5): 836-45.
- Oppenheimer, J.H., Levy, M.L., Sinha, U., el-Kadi, H., Apuzzo, M.L. y col. (1992). "Radionecrosis secondary to interstitial brachytherapy: correlation of magnetic resonance imaging and histopathology." *Neurosurgery* 31(2): 336-43.
- Ossowski, S., Schneeberger, K., Lucas-Lledo, J.I., Warthmann, N., Clark, R.M. y col. (2010). "The rate and molecular spectrum of spontaneous mutations in *Arabidopsis thaliana*." *Science* 327(5961): 92-4.
- Parker, J.E., Coleman, M.J. (1997). "Molecular intimacy between proteins specifying plant-pathogen recognition." *Trends Biochem Sci* 22(8): 291-6.
- Pauline, C., Henikoff, S. (2001). "Predicting deleterious amino acid substitutions on protein function." *Genome Res* 11(5): 863-74.
- Pauline, C., Henikoff, S. (2003). "SIFT: predicting amino acid changes that affect protein function." *Nucleic Acids Res* 31(13).
- Pauquet, J., Burguet, E., Hagen, L. (2004). Map-based cloning of the Vat gene from melon conferring resistance to both aphid colonization and aphid transmission of several viruses. En: Lebeda, A., Paris, HS. Editorial. Palack University, Olomouc, Czech Republic.
- Pecker, I., Chamovitz, D., Linden, H., Sandmann, G., Hirschberg, J. (1992). "A single polypeptide catalyzing the conversion of phytoene to zeta-carotene is transcriptionally regulated during tomato fruit ripening." *Proc Natl Acad Sci U S A* 89(11): 4962-6.
- Perin, C., Gómez-Jiménez, M., Hagen, L., Dogimont, C., Pech, J.C. y col. (2002). "Molecular and genetic characterization of a non-climacteric phenotype in melon reveals two loci conferring altered ethylene response in fruit." *Plant Physiol* 129(1): 300-9.

Perry, J.A., Wang, T.L., Welham, T.J., Gardner, S., Pike, J.M. y col. (2003). "A TILLING reverse genetics tool and a web-accessible collection of mutants of the legume *Lotus japonicus*." *Plant Physiol* 131: 866 - 871.

Piron, F., Nicolai, M., Minoia, S., Piednoir, E., Moretti, A. y col. (2010). "An induced mutation in tomato eIF4E leads to immunity to two potyviruses." *PLoS One* 5(6): e11313.

Pitrat, M., Hanelt, P., Hammer, (2000). "Some comments on infraspecific classification of cultivars of melon." *Acta Hort. (ISHS)* 510: 29-3.

Portnoy, V., Diber, A., Pollock, S., Karchi, H., Lev, S. (2011). "Use of non-normalized, Non-Amplified cDNA for 454-Based RNA Sequencing of Fleshy Melon Fruit". *Plant Gen. (4)*: 36-46.

Provvidenti, R., Hampton, R.O. (1992). "Sources of resistance to viruses in the Potyviridae." *Arch Virol Suppl* 5: 189-211.

Qin G, Gu, H., Ma, L., Peng, Y., Deng, XW., y col. (2007). "Disruption of phytoene desaturase gene results in albino and dwarf phenotypes in *Arabidopsis* by impairing chlorophyll, carotenoid, and gibberellin biosynthesis." *Cell Research*. 17(5): 471-82.

Quadrana, L., Rodríguez, M.C., López, M., Bermúdez, L., Nunes-Nesi, A. y col. (2011). "Coupling virus-induced gene silencing to exogenous green fluorescence protein expression provides a highly efficient system for functional genomics in *Arabidopsis* and across all stages of tomato fruit development." *Plant Physiol* 156(3): 1278-91.

Rédei, G.P., Koncz, C (1992). "Classical mutagenesis." *Methods in Arabidopsis research*. 23(8) 134.

Rick, CM., Bulter, L. (1956). "Cytogenetics of the tomato." *Adv Genet VIII*: 267-382.

Rigola, D., Van Oeveren, J., Janssen, A., Bonne, A., Schneiders, H. y col. (2009). "High-throughput detection of induced mutations and natural variation using KeyPoint technology." *PLoS One* 4(3): e4761.

Robaglia, C., Caranta, C. (2006). "Translation initiation factors: a weak link in plant RNA virus infection." *TRENDS in Plant Science* 11(1).

Rodríguez, C.M., Freire, M.A., Camilleri, C., Robaglia, C. (1998). "The *Arabidopsis thaliana* cDNAs coding for eIF4E and eIF(iso)4E are not functionally equivalent for yeast complementation and are differentially expressed during plant development." *Plant J* 13(4): 465-73.

Rodríguez-Hernández, A.M., Gosalvez, B., Sempere, R.N., Burgos, L., Aranda, M.A. y col. (2012). "Melon RNA interference (RNAi) lines silenced for Cm-eIF4E show broad virus resistance." *Mol Plant Pathol* 13(7): 755-63.

Romer, P., Hahn, S., Jordan, T., Strauss, T., Bonas, U. y col. (2007). "Plant pathogen recognition mediated by promoter activation of the pepper Bs3 resistance gene." *Science* 318(5850): 645-8.

- Rose J.K.C, Hadfield, K.A. (1998). "Temporal sequence of cell wall disassembly in rapidly ripening melon fruit." *Plant Physiol* 117: 345-361.
- Rost, B., Sander, C. (1993). "Prediction of protein secondary structure at better than 70% accuracy." *J Mol Biol* 232(2): 584-99.
- Ruffel, S., Dussault, M.H., Palloix, A., Moury, B., Bendahmane, A. y col. (2002). "A natural recessive resistance gene against potato virus Y in pepper corresponds to the eukaryotic initiation factor 4E (eIF4E)." *Plant J* 32(6): 1067-75.
- Ruffel, S., Caranta, C., Palloix, A., Lefebvre, V., Caboche, M. y col. (2004). "Structural analysis of the eukaryotic initiation factor 4E gene controlling potyvirus resistance in pepper: exploitation of a BAC library." *Gene* 338(2): 209-216.
- Ruffel, S., Gallois, J.L., Lesage, M.L., Caranta, C. (2005). "The recessive potyvirus resistance gene pot-1 is the tomato orthologue of the pepper pvr2-eIF4E gene." *Mol Genet Genomics* 274(4): 346-53.
- Sabetta, W., Alba, V., Blanco, A., Montemurro, C. (2011). "SunTILL: a TILLING resource for gene function analysis in sunflower." *Plant Methods* 7(1): 20.
- Sakamoto, A., Lan, V.T., Hase, Y., Shikazono, N., Matsunaga, T. y col. (2003). "Disruption of the AtREV3 gene causes hypersensitivity to ultraviolet B light and gamma-rays in Arabidopsis: implication of the presence of a translesion synthesis mechanism in plants." *Plant Cell* 15(9): 2042-57.
- Saladié, M., Rose, J., Vicente, AR. (2007). "The linkage between cell wall metabolism and fruit softening: looking to the future." *J Sci Food Agric* 87: 1435-1448.
- Sanseverino, W., Roma, G., De Simone, M., Faino, L., Melito, S. y col. (2009). "PRGdb: a bioinformatics platform for plant resistance gene analysis." *Nucleic Acids Res* 38(Database issue): D814-21.
- Sasaki, T., Matsumoto, T., Antonio, B.A., Nagamura, Y. (2005). "From mapping to sequencing, post-sequencing and beyond." *Plant Cell Physiol* 46(1): 3-13.
- Sato-Nara Kumi, K.-I.Y. (1999). "Ethylene Receptors and Genetic Engineering of Ethylene Sensitivity " *Plant Biotechnology* 16(5): 321-334.
- Schaller, G.E., Bleecker, A.B. (1995). "Ethylene-binding sites generated in yeast expressing the Arabidopsis ETR1 gene." *Science* 270(5243): 1809-11.
- Schneider, D.S. (2002). "Plant immunity and film Noir: what gumshoe detectives can teach us about plant-pathogen interactions." *Cell* 109(5): 537-40.
- Sebastian, P.S., Telford, H., Ian, R., Renner, S. (2010). Cucumber (*Cucumis sativus*) and melon (*C. melo*) have numerous wild relatives in Asia and Australia, and the sister species of melon is from Australia. *Proceedings of the National Academy of Sciences of the United States of America*. 107: 14269-73.
- Seffernick, J.L., Samanta, S.K., Louie, T.M., Wackett, L.P., Subramanian, M. (2009). "Investigative mining of sequence data for novel enzymes: a case study with nitrilases." *J Biotechnol* 143(1): 17-26.
- Sega, G.A. (1984). "A review of the genetic effects of ethyl methanesulfonate." *Mutat Res* 134(2-3): 113-42.

Seymour, GB., MacGlasson, W. ((1993)). "Melons. in Biochemistry of Fruit Ripening." Editorial: Seymour GB, Taylor JE, Tucker GA (Chapman & Hall). Londres. p. 273-290.

Shaun, J., Voytas, F., Stupar, M., (2012). "Genome Engineering of crops with designer nucleases. *The Plant Genome*. 5 (2). 42-50.

Shellie, K., Mikal, E. (1993). "The Lack of a Respiratory Rise in Muskmelon Fruit Ripening on the Plant Challenges the Definition of Climacteric Behaviour." *Journal of Experimental Botany* 44(8): 1403-1406.

Shen, K.A., Chin, D.B., Arroyo-Garcia, R., Ochoa, O.E., Lavelle, D.O. y col. (2002). "Dm3 is one member of a large constitutively expressed family of nucleotide binding site-leucine-rich repeat encoding genes." *Mol Plant Microbe Interact* 15(3): 251-61.

Shikazono, N., Yokota, Y., Tanaka, A., Watanabe, H., Tano, S. (1998). "Molecular analysis of carbon ion-induced mutations in *Arabidopsis thaliana*." *Genes Genet Syst* 73(3): 173-9.

Shikazono, N., Tanaka, A., Watanabe, H., Tano, S. (2001). "Rearrangements of the DNA in carbon ion-induced mutants of *Arabidopsis thaliana*." *Genetics* 157(1): 379-87.

Shikazono, N., Yokota, Y., Kitamura, S., Suzuki, C., Watanabe, H. y col. (2003). "Mutation rate and novel tt mutants of *Arabidopsis thaliana* induced by carbon ions." *Genetics* 163(4): 1449-55.

Shirley, B.W., Hanley, S., Goodman, H.M. (1992). "Effects of ionizing radiation on a plant genome: analysis of two *Arabidopsis* transparent testa mutations." *Plant Cell* 4(3): 333-47.

Skelton, R.L., Yu, Q., Srinivasan, R., Manshardt, R., Moore, P.H. y col. (2006). "Tissue differential expression of lycopene beta-cyclase gene in papaya." *Cell Res* 16(8): 731-9.

Slade, A.J., Fuerstenberg, S.I., Loeffler, D., Steine, M.N., Facciotti, D. (2005). "A reverse genetic, nontransgenic approach to wheat crop improvement by TILLING." *Nat Biotechnol* 23(1): 75-81.

Smith, B.L., Lu, C.P., Alvarado Bremer, J.R. (2012). "Methodological streamlining of SNP discovery and genotyping via high-resolution melting analysis (HRMA) in non-model species." *Mar Genomics*.

Soberónn J.R, Quiroga, E.N, Sampietro, A.R. (2005). "Etileno". Publicado por el Instituto de Estudios Vegetales. Facultad de Bioquímica y Farmacia. Universidad Nacional de Tucumán.

Song, W.Y., Wang, G.L., Chen, L.L., Kim, H.S., Pi, L.Y. y col. (1995). "A receptor kinase-like protein encoded by the rice disease resistance gene, Xa21." *Science* 270(5243): 1804-6.

Spieth, J., Lawson, D. (2005). " Overview of gene structure" *WormBook Journal*: <http://www.wormbook.org/>.

- Sreelakshmi, Y., Gupta, S., Bodanapu, R., Chauhan, V.S., Hanjabam, M. y col. (2009). "NEATTILL: A simplified procedure for nucleic acid extraction from arrayed tissue for TILLING and other high-throughput reverse genetic applications." *Plant Methods* 6(1): 3.
- Stebbins, W.C., Moody, D.B. (1989). "Concerning the hallmark of a discrimination." *Behav Anal* 12(2): 203-11.
- Stepanova, A.N., Ecker, J.R. (2000). "Ethylene signaling: from mutants to molecules." *Curr Opin Plant Biol* 3(5): 353-60.
- Stephenson, P., Baker, D., Girin, T., Perez, A., Amoah, S. y col. (2010). "A rich TILLING resource for studying gene function in *Brassica rapa*." *BMC Plant Biol* 10: 62.
- Strachan, T., Read, A.P. (1999). "Human Molecular Genetics." Editorial: Scientific Publishers. Oxford.
- Streisinger, G., Owen, J. (1985). "Mechanisms of spontaneous and induced frameshift mutation in bacteriophage T4." *Genetics* 109(4): 633-59.
- Subbaiyan, G.K., Waters, D.L., Katiyar, S.K., Sadananda, A.R., Vaddadi, S. y col. (2012). "Genome-wide DNA polymorphisms in elite indica rice inbreds discovered by whole-genome sequencing." *Plant Biotechnol J*.
- Tadmor, Y., Katzir, N., Meir, A., Yaakov, Y., Sa'ar, U. (2007). "Induced mutagenesis to augment the natural variability of melon (*Cucumis melo* L.)." *Israel J Plant Sci* 55: 159-169.
- Takumi, S., Koyama, K., Fujiwara, K., Kobayashi, F. (2011). "Identification of a large deletion in the first intron of the *Vrn-D1* locus, associated with loss of vernalization requirement in wild wheat progenitor *Aegilops tauschii* Coss." *Genes Genet Syst* 86(3): 183-95.
- Talame, V., Bovina, R., Sanguineti, M.C., Tuberosa, R., Lundqvist, U. y col. (2008). "TILLMore, a resource for the discovery of chemically induced mutants in barley." *Plant Biotechnol J* 6(5): 477-85.
- Tanaka, A., Shikazono, N., Yokota, Y., Watanabe, H., Tano, S. (1997). "Effects of heavy ions on the germination and survival of *Arabidopsis thaliana*." *Int J Radiat Biol* 72(1): 121-7.
- Taylor, N.E., Greene, E.A. (2003). "PARSESNP: a tool for the analysis of nucleotide polymorphisms." *Nucl Acids Res* 31: 3808 - 3811.
- Thanh, P.T., Phan, P.D., Ishikawa, R., Ishii, T. (2010). "QTL analysis for flowering time using backcross population between *Oryza sativa* Nipponbare and *O. rufipogon*." *Genes Genet Syst* 85(4): 273-9.
- Tieman, D.M., Taylor, M.G., Ciardi, J.A., Klee, H.J. (2000). "The tomato ethylene receptors *NR* and *LeETR4* are negative regulators of ethylene response and exhibit functional compensation within a multigene family." *Proc Natl Acad Sci U S A* 97(10): 5663-8.
- Till, B.J., Colbert, T., Tompa, R., Enns, L.C., Codomo, C.A. y col. (2003). "High-throughput TILLING for functional genomics." *Methods Mol Biol* 236: 205-20.

- Till, B.J., Reynolds, S.H., Weil, C., Springer, N., Burtner, C. y col. (2004 a). "Discovery of induced point mutations in maize genes by TILLING." *BMC Plant Biol* 4: 12.
- Till, B.J., Burtner, C., Comai, L., Henikoff, S. (2004 b). "Mismatch cleavage by single-strand specific nucleases." *Nucl Acids Res* 32: 2632 - 2641.
- Till, B.J., Cooper, J., Tai, T.H., Colowit, P., Greene, E.A. y col. (2007). "Discovery of chemically induced mutations in rice by TILLING." *BMC Plant Biol* 7: 19.
- Triques, K., Sturbois, B., Gallais, S., Dalmais, M., Chauvin, S. y col. (2007). "Characterization of *Arabidopsis thaliana* mismatch specific endonucleases: application to mutation discovery by TILLING in pea." *Plant J* 51(6): 1116-25.
- Tromas, N., Elena, S.F. (2010). "The rate and spectrum of spontaneous mutations in a plant RNA virus." *Genetics* 185(3): 983-9.
- Truniger, V., Nieto, C., González-Ibeas, D., Aranda, M. (2008). "Mechanism of plant eIF4E-mediated resistance against a Carmovirus (Tombusviridae): cap-independent translation of a viral RNA controlled in cis by an (a)virulence determinant." *Plant J* 56(5): 716-27.
- Tsai, H., Howell, T., Nitcher, R., Missirian, V., Watson, B. y col. (2011). "Discovery of rare mutations in populations: TILLING by sequencing." *Plant Physiol* 156(3): 1257-68.
- Upadhyaya, N.M., Zhu, Q.H., Bhat, R.S. (2010). "Transposon insertional mutagenesis in rice." *Methods Mol Biol* 678: 147-77.
- Van der Biezen, E.A., Jones, J.D. (1998). "Plant disease-resistance proteins and the gene-for-gene concept." *Trends Biochem Sci* 23(12): 454-6.
- Van der Hoorn, R.A., De Wit, P.J., Joosten, M.H. (2002). "Balancing selection favors guarding resistance proteins." *Trends Plant Sci* 7(2): 67-71.
- Van der Plank, J.E. (1965). "Dynamics of Epidemics of Plant Disease: Population bursts of fungi, bacteria, or viruses in field and forest make an interesting dynamical study." *Science* 147(3654): 120-4.
- Van Ooijen, G., Van den Burg, H.A., Cornelissen, B.J., Takken, F.L. (2007). "Structure and function of resistance proteins in solanaceous plants." *Annu Rev Phytopathol* 45: 43-72.
- Vigouroux, Y., Jaqueth, J.S., Matsuoka, Y., Smith, O.S., Beavis, W.D. y col. (2002). "Rate and pattern of mutation at microsatellite loci in maize." *Mol Biol Evol* 19(8): 1251-60.
- Villanueva, B., Pong-Wong, R., Woolliams, J.A. (2002). "Marker assisted selection with optimised contributions of the candidates to selection." *Genet Sel Evol* 34(6): 679-703.
- Vos, P., Hogers, R., Bleeker, M., Reijans, M., Van de Lee, T. y col. (1995). "AFLP: a new technique for DNA fingerprinting." *Nucleic Acids Res* 23(21): 4407-14.

- Wang, D.K., Sun, Z.X., Tao, Y.Z. (2006). "Application of TILLING in plant improvement." *Yi Chuan Xue Bao* 33(11): 957-64.
- Wang, J. (2001). "Optimal marker-assisted selection to increase the effective size of small populations." *Genetics* 157(2): 867-74.
- Wang, X., Kohalmi, S.E., Svircev, A., Wang, A., Sanfacon, H. y col. (2013). "Silencing of the host factor eIF(iso)4E gene confers plum pox virus resistance in plum." *PLoS One* 8(1): e50627.
- Warthmann, N., Chen, H., Ossowski, S., Weigel, D., Herve, P. (2008). "Highly specific gene silencing by artificial miRNAs in rice." *PLoS One* 3(3): e1829.
- Weil, CF., Monde, R.A. (2007). "Getting to the point - Mutations in maize." *Crop Sci* 47: S60-67.
- Whitaker, T.W. (1976). "Cucurbits, botany, cultivation and utilization." Editorial: Interscience Publishers. Nueva York.
- Wienholds, E., Van Eeden, F., Kusters, M., Mudde, J., Plasterk, R.H. y col. (2003). "Efficient target-selected mutagenesis in zebrafish." *Genome Res* 13(12): 2700-7.
- Wilkinson, J.Q., Lanahan, M.B., Yen, H.C., Giovannoni, J., Klee, H.J. (1995). "An ethylene-inducible component of signal transduction encoded by never-ripe." *Science* 270(5243): 1807-9.
- Williams (1990). "DNA polymorphism amplified by arbitrary primers are useful as genetic markers." *Nucl Acids Res* 18: 6531-6535.
- Winkler, S., Schwabedissen, A., Backasch, D., Bokel, C., Seidel, C. y col. (2005). "Target-selected mutant screen by TILLING in *Drosophila*." *Genome Res* 15: 718 - 723.
- Wittmann, S., Chatel, H., Fortin, M.G., Laliberte, J.F. (1997). "Interaction of the viral protein genome linked of turnip mosaic potyvirus with the translational eukaryotic initiation factor (iso) 4E of *Arabidopsis thaliana* using the yeast two-hybrid system." *Virology* 234(1): 84-92.
- Wolfsberg, T.G. (2011). "Using the NCBI Map Viewer to browse genomic sequence data." *Curr Protoc Hum Genet* Chapter 18: Unit18 5.
- Wright, D.A., Voytas, D.F. (1998). "Potential retroviruses in plants: Tat1 is related to a group of *Arabidopsis thaliana* Ty3/gypsy retrotransposons that encode envelope-like proteins." *Genetics* 149(2): 703-15.
- Wu, J.L., Wu, C., Lei, C., Baraoidan, M., Bordeos, A. y col. (2005). "Chemical- and irradiation-induced mutants of indica rice IR64 for forward and reverse genetics." *Plant Mol Biol* 59(1): 85-97.
- Yang, B., Wen, X., Kodali, N.S., Oleykowski, C.A., Miller, C.G. y col. (2000). "Purification, cloning, and characterization of the CEL I nuclease." *Biochemistry* 39(13): 3533-41.
- Yen, H.C., Lee, S., Tanksley, S.D., Lanahan, M.B., Klee, H.J. y col. (1995). "The tomato Never-ripe locus regulates ethylene-inducible gene expression and is linked to a homolog of the *Arabidopsis* ETR1 gene." *Plant Physiol* 107(4): 1343-53.

Yeung, A.T., Hattangadi, D., Blakesley, L., Nicolas, E. (2005). "Enzymatic mutation detection technologies." *Biotechniques* 38(5): 749-58.

Youil, R., Kemper, B.W., Cotton, R.G. (1995). "Screening for mutations by enzyme mismatch cleavage with T4 endonuclease VII." *Proc Natl Acad Sci U S A* 92(1): 87-91.

Yu, T.A., Chiang, C.H., Wu, H.W., Li, C.M., Yang, C.F. y col. (2010). "Generation of transgenic watermelon resistant to Zucchini yellow mosaic virus and Papaya ringspot virus type W." *Plant Cell Rep* 30(3): 359-71.

Zarembinski, T.I., Theologis, A. (1994). "Ethylene biosynthesis and action: a case of conservation." *Plant Mol Biol* 26(5): 1579-97.

Zheng, S.J., Snoeren, T.A., Hogewoning, S.W., Van Loon, J.J., Dicke, M. (2010). "Disruption of plant carotenoid biosynthesis through virus-induced gene silencing affects oviposition behaviour of the butterfly *Pieris rapae*." *New Phytol* 186(3): 733-45.

Zhong, S., Lin, Z., Grierson, D. (2008). "Tomato ethylene receptor-CTR interactions: visualization of never-ripe interactions with multiple CTRs at the endoplasmic reticulum." *J Exp Bot* 59(4): 965-72.

ANEXO I

En este anexo se presenta el artículo publicado fruto, en gran parte, del trabajo realizado en esta tesis doctoral. Se incluyen también dos artículos que no están relacionados con la temática de la tesis, en los que se colaboró paralelamente al desarrollo de la población de TILLING de melón.

RESEARCH ARTICLE

Open Access

Towards a TILLING platform for functional genomics in Piel de Sapo melons

Mireia González¹, Meihong Xu^{1,5}, Cristina Esteras², Cristina Roig², Antonio J Monforte^{1,3}, Christelle Troadec⁴, Marta Pujol¹, Fernando Nuez², Abdelhafid Bendahmane⁴, Jordi Garcia-Mas¹ and Belén Picó^{2*}

Abstract

Background: The availability of genetic and genomic resources for melon has increased significantly, but functional genomics resources are still limited for this crop. TILLING is a powerful reverse genetics approach that can be utilized to generate novel mutations in candidate genes. A TILLING resource is available for *cantalupensis* melons, but not for *inodorus* melons, the other main commercial group.

Results: A new ethyl methanesulfonate-mutagenized (EMS) melon population was generated for the first time in an andromonoecious non-climacteric *inodorus* Piel de Sapo genetic background. Diverse mutant phenotypes in seedlings, vines and fruits were observed, some of which were of possible commercial interest. The population was first screened for mutations in three target genes involved in disease resistance and fruit quality (*Cm-PDS*, *Cm-elf4E* and *Cm-elf1(iso)4E*). The same genes were also tilled in the available monoecious and climacteric *cantalupensis* EMS melon population. The overall mutation density in this first Piel de Sapo TILLING platform was estimated to be 1 mutation/1.5 Mb by screening four additional genes (*Cm-ACO1*, *Cm-NOR*, *Cm-DET1* and *Cm-DHS*). Thirty-three point mutations were found for the seven gene targets, six of which were predicted to have an impact on the function of the protein. The genotype/phenotype correlation was demonstrated for a loss-of-function mutation in the *Phytoene desaturase* gene, which is involved in carotenoid biosynthesis.

Conclusions: The TILLING approach was successful at providing new mutations in the genetic background of Piel de Sapo in most of the analyzed genes, even in genes for which natural variation is extremely low. This new resource will facilitate reverse genetics studies in non-climacteric melons, contributing materially to future genomic and breeding studies.

Keywords: *Cucumis melo*, *inodorus*, TILLING, mutant, disease resistance, fruit quality

Background

Melon (*Cucumis melo* L.) is an important vegetable crop. Genetic and genomic information for this crop is increasing significantly due to several national and international projects [1]. A broad range of genomic tools are available today [2-7]. An effort is also in progress, through a Spanish initiative, to obtain the whole genome sequence of this crop [8]. These tools are generating a lot of information about genes involved in various biological processes, such as plant resistance and fruit quality [9,10]. However, the tools necessary for reverse genetic

studies to conduct the functional validation of candidate genes in melons are still limited.

The TILLING (Targeting Induced Local Lesions in Genomes) method may represent an effective means of addressing limitations in melon research [11]. The application of TILLING has proved to be useful in identifying novel alleles in genes controlling traits of agronomic interest in legumes [12-15], cereals [16-23], solanaceous crops [24-26] and *Brassica* spp [27,28].

C. melo is a highly variable species divided into two subspecies, *melo* and *agrestis*. Most of the commercial cultivars belong to the *inodorus* and *cantalupensis* groups of the subspecies *melo* [29]. Cultivars of these two commercial groups are quite different in plant, flowering and fruit traits. *Cantalupensis* cultivars are

* Correspondence: MPICOSI@BTC.UPV.ES

²Institute for the Conservation and Breeding of Agricultural Biodiversity (COMAV-UPV), Universitat Politècnica de València, Camino de Vera s/n, 46022 Valencia, Spain

Full list of author information is available at the end of the article



early flowering, frequently monoecious, and produce climacteric fruits (aromatic, medium sugar, soft fleshed and with a short shelf life), whereas *inodorus* are late flowering, mostly andromonoecious, and non-climacteric, producing non-aromatic, high sugar, firm-fleshed fruits with a long shelf life, and which are preferred in certain markets. These differences make breeding strategies and objectives quite different for the two groups of melons.

A melon TILLING platform generated from a monoecious climacteric *cantalupensis* genotype has recently become available [30], and has proven to be useful for improving the shelf life of climacteric melons. A TILLING platform generated in an *inodorus* background might represent a useful resource for functional studies and for breeding non-climacteric melons.

In this paper, we present the development of an EMS-mutagenized population from an andromonoecious non-climacteric *inodorus* melon type (*Cucumis melo* L. subspecies *melo* cv Piel de Sapo). In total, 2,368 M2 families were obtained. A diversity of interesting mutant phenotypes were observed in the field. In addition, several mutant families have been identified in several genes that were selected based on their potential contribution to fruit quality and disease resistance.

Methods

EMS mutagenesis of melon seeds

The double haploid line, M62-113 (from Semillas Fito S.A.), belonging to the Piel de Sapo commercial type (*Cucumis melo* subsp *melo* var *inodorus*), was used as the starting cultivar. This is the parent of the melon genetic map and is one of the parents of the DHL line selected for the sequencing of the melon genome. Different batches of a total of ~12,000 M62-113 seeds were treated with the alkylant mutagen ethyl methane-sulfonate (EMS). Different EMS concentrations were tested, and 1% was finally selected and tested on a larger batch of seeds. Seeds were first hydrated with tap water (16h), and were then soaked in tap water containing EMS concentrations (18 h). The treated seeds were thoroughly rinsed in tap water twice for 4 hours. After washing, seeds were placed on trays over wet paper, and were kept in a germination chamber overnight. ~5,000 mutant plants of this first (M1) generation were transplanted at the greenhouse. Plants were grown in two localities, in Barcelona (by Semillas Fito S.A. and IRTA) and in Valencia (by COMAV-UPV). Each M1 plant was selfed (1 fruit per plant) giving rise to the M2 seed.

TILLING procedure

Twenty seeds were sown per M2 family, and the germination percentage was scored. Finally, a total of 2,368

mutagenized M2 families, which produced at least 10 viable seedlings in at least two plantings, were sampled for DNA extraction. 1-cm discs of tissue from 10 individual plants per M2 were pooled, and total DNA was extracted, quantified, diluted and organized into twenty-five 96-well plates. Using equivalent amounts of DNA from individual M2 families, the samples were pooled four-fold and organized into a 96-well format. Seeds of all pooled M2 populations are maintained at the Genebank of the COMAV-UPV, coded as F (S. Fito), I (IRTA) or C (COMAV-UPV) plus the corresponding number.

Primer sets were designed from 3 genes that were selected based on their potential contribution to fruit quality and disease resistance: *Cm-eIF4E* (translation initiation factor 4E) and *Cm-eIF(iso)4E* (translation initiation factor E, Isoform) which are involved in resistance to viruses [31], and *Cm-PDS* (Phytoene Desaturase), which is involved in carotenoid synthesis [32]. The CODDLE program (Codons Optimized to Deliver Deleterious Lesions) [33] was used to select the regions most likely to harbor deleterious changes induced by EMS (Table 1). These 3 genes were tilled in the Piel de Sapo population as well as in a set of 2,483 available lines of the *cantalupensis* TILLING population reported in [30]. Four additional genes involved in fruit ripening were tilled in the Piel de Sapo population: *Cm-ACO1* (ACC oxidase 1), involved in the conversion of ACC into ethylene [34]; *Cm-NOR* (non-ripening), a transcription factor related to ethylene-sensitive/insensitive phenotypes [35]; *Cm-DETI* (de-tiolated-1), a negative regulator of light-mediated responses that affects carotenoid and flavonoid pathways in tomato and other crops [36,37]; and *Cm-DHS* (Deoxyhypusine Synthase), mutations of which delay fruit softening in tomato [38]. Primer sets used for these additional genes are the same as those used in [30]. The amplicons analyzed for each gene are indicated in Figures 1 and 2.

Nested PCR was used to improve the specificity of amplification. The PCR reactions were performed in a 25 μ l volume consisting of dH₂O, 1 \times PCR buffer (Promega Corp, Madison, WI), 2.5 mM MgCl₂, 0.2 mM dNTPs, 1 U Taq polymerase, 0.4 mM forward and reverse primers, unlabeled in the first reaction, and 700 nm and 800 nm 5' labeled (MWG Biotech AG, Germany) in the second reaction, and 10 ng DNA. The thermocycling conditions were 95°C for two minutes for initial denaturing, followed by 35 cycles of 95°C for 20 seconds, 63-69°C (specific for each primer pair) for one minute, 72°C from 30 seconds to one minute and one cycle of 72°C for five minutes. PCR products (~0.2 μ l) were separated in 1% agarose gels. The PCR products were heated and cooled in a thermocycler (a gradient starting at 94°C and decreasing 0.1°C per second

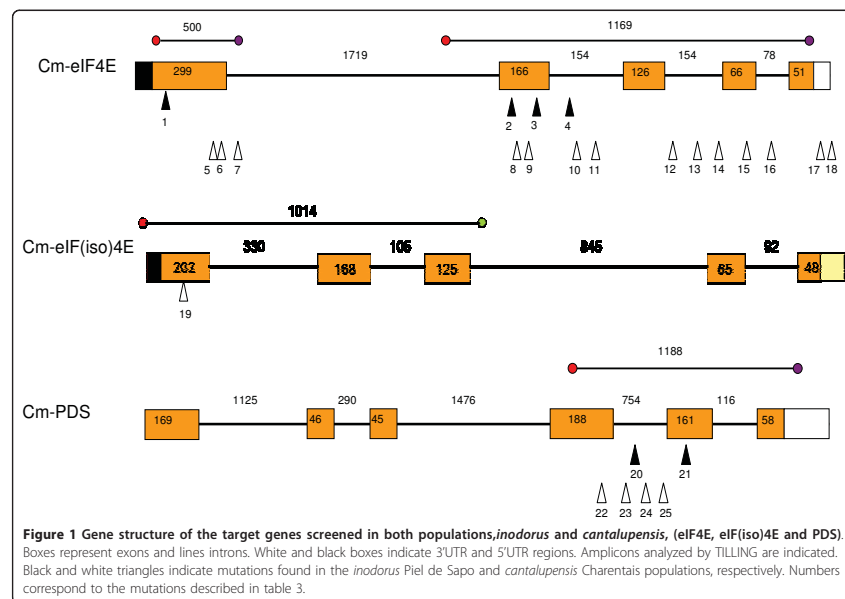
Table 1 Primer pairs used to amplify the 3 target genes tiled in both EMS-mutant populations

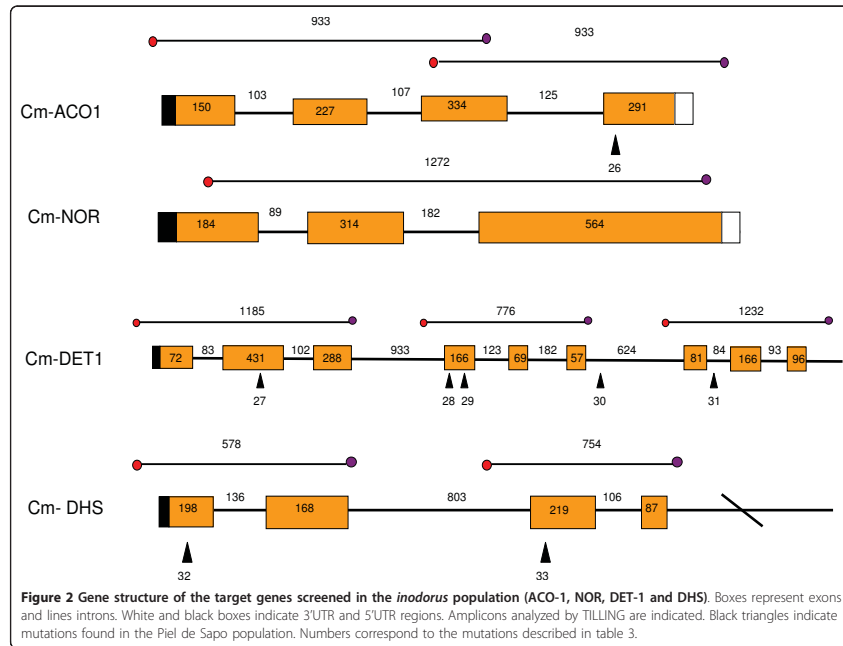
Gene	Forward (*LabelledIRD700)	Reverse (*LabelledIRD800)
Cm-eIF4E 1 st amplicon	GACTCAAACGCCTAACAGAAAATC	CTATCTCACCAAGTTTCCTAAATT
	*GAGGGCGGTGCCATTCTCTCCGG	*TCCCTAAATCGAACCAAGAAACGCC
Cm-eIF4E 2 nd amplicon	TTGTTTCGTGTGACATGCCATC	AGCATGCATTACCTCATTGGCT
	*TGCTTGGCTGTTAATTTATCTCTGC	*GTCAGTACAGAACAAGAATCTGAG
Cm-eIF(iso)4E	GATCGATAATTCCCTTTTC	TGACAGGCTTAAGCACTATG
	*ACCCCAATTCATTCTAGGG	*GTGCAAGCGCAACAAGGTAC
Cm-PDS	CAATGGGATGCTTAGATCAAC	TCCTTCCAAAATTACCCTAG
	*GGATTCTCGAGTGACTCGG	*GCTCTCGAGCAACTAGCACT

Nested PCR was used in all cases with labeled primers in the second PCR.

to 4°C) to form the heteroduplexes. Once the heteroduplexes were formed, the products were treated with Endo I. Digestions were performed in 30µl of final volume with dH₂O, 3µl digestion buffer, 150 ng of PCR product and 3µl of Endo I (diluted 1/5000). The products were incubated at 45°C for 20 minutes to digest mismatches in the heteroduplex. Digestion was stopped by adding 5µl of 0.15 M EDTA and placing the reaction on ice. After the digestion was completed, the products were filtered through a Millipore MultiScreen filter plate that was

packed with hydrated Sephadex G-50 medium beads (Amersham Biosciences). Samples were concentrated in the Speed Vac at 65°C for 45 minutes. Loading dye was added and the PCR products were denatured and loaded onto a polyacrylamide gel attached to a LI-COR 4300 DNA Analyzer (LICOR, Lincoln, NE) for separation. A total of 903 pools were assayed (592 four-fold pools of the Piel the Sapó population and 311 eight-fold pools of the *cantalupensis* population). Once a mutation was revealed in a pool, the corresponding families were





analyzed individually to select the mutant M2 family. Mutation was confirmed by sequencing in the pooled DNA, in the individual M2 families and in several individual plants used in the segregation studies. The effect of the mutations was analyzed with SIFT (Sorting Intolerant from Tolerant, <http://blocks.fhcrc.org/sift/SIFT.html>) [39], which predicts whether an amino acid substitution affects protein function. Scores below 0.05 are predicted to affect protein function.

Phenotyping of M2 and M3 individuals

Seedling phenotypes were systematically evaluated in the M2 generation. Additionally, in order to increase the number of seeds in our mutagenized population, we obtained the M3 generation. M2 populations that are being reproduced are also being inspected for phenotypes that are distinct from the wild type in plant, flowering and fruit traits. To date, about 800 M2 plants have been reproduced by S. Fito. M3 plants are also being analyzed for ethylene-response mutants by growing seedlings in darkness with and without an air flux containing 10 microL L-1 ethylene, and analyzing the "triple response".

Results

Generation of an EMS mutant population in the Piel de Sapo background

We selected 0.5%, 1 and 1.5% for generating a pilot population of ~600 lines, and studied the effect of EMS dosage on M2 seed viability and seedling vigor. Increased EMS concentration led to an increased percentage of M2 fruits in which seeds were non-viable or in which the number of viable seeds was too low to perform TILLING analysis (13.24%, 27.75% and 31.38%, respectively). Only a slight reduction in the average germination rate was observed in M2 families obtained from 0.5% M1 (93.51%), whereas higher reductions were obtained in 1% and 1.5% (82.73 and 75.94%, respectively). In addition, seedling growth was more restricted at the 1.5% concentrations. In contrast, the progeny from the 1%-treated M1 were robust. Based on these results, 1% EMS was used in the end to complete the mutagenized melon population. This EMS dose produces an acceptable level of M1 seed survival, fertility in M1 plants to set viable M2 seeds, and vigor in M2 seedlings. From a total of 3,140 M1 fruits finally

obtained, 2,368 M2 families were sampled for DNA and used for TILLING purposes.

Phenotyping of the mutagenized population

Ten plants per M2 line were scored for seedling phenotypes. A large number of the mutations affected cotyledons, with 5% of the M2 families showing variation in their number position and shape. About 4.3% of the M2 families segregated for "dwarf" or "semi-dwarf" plants. Albinism and chlorophyll deficiency occurred in 2.1% of the M2 families. Over 2% of the families showed alterations in leaves and shoot morphology. Gravitropic mutants and necrosis also appeared, but were less frequent. Figure 3 shows several seedling phenotypes. M3 families were characterized for plant flowering and fruit traits. Fruit traits of interest for melon breeding were identified: longer and rounder fruits than the standard Piel de Sapo fruits, with a lighter rind color, without the typical rind spots, and with different degrees of netting. Some lines have a subtle aroma and showed a color change and the formation of an abscission layer near maturity,

traits that are more similar to climacteric *cantalupensis* types than to the standard non-climacteric *inodorus* Piel de Sapo. Differences in the dark-grown seedlings' response to ethylene were also observed. Wild type M62-113 showed the typical triple response: exaggeration of the apical hook, together with the inhibition of hypocotyl and root growth as well as the radial swelling of the shoot [40], whereas some M3 families showed responses of varying intensities (Figure 3).

Identification of mutations in target genes by TILLING

The TILLING approach allowed us to identify nucleotide changes in two of the three gene targets first screened in the Piel de Sapo population (Table 2). EMS-induced changes were mostly G:C to A:T transitions (67%) (Table 3), as was expected due to the frequent alkylation of guanine residues by EMS, thus forcing mispairing with T [41]. Four mutations were exonic, most of which were determined to be missense mutations, only one being silent. A graph of the target genes marking the location of the induced polymorphisms is shown in Figures 1 and 2.

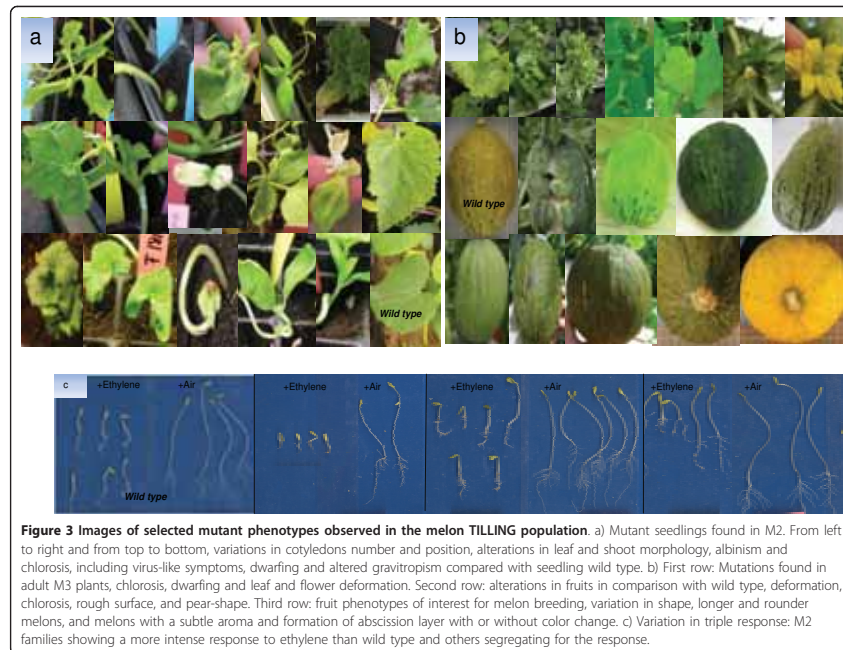


Figure 3 Images of selected mutant phenotypes observed in the melon TILLING population. a) Mutant seedlings found in M2. From left to right and from top to bottom, variations in cotyledons number and position, alterations in leaf and shoot morphology, albinism and chlorosis, including virus-like symptoms, dwarfing and altered gravitropism compared with seedling wild type. b) First row: Mutations found in adult M3 plants, chlorosis, dwarfing and leaf and flower deformation. Second row: alterations in fruits in comparison with wild type, deformation, chlorosis, rough surface, and pear-shape. Third row: fruit phenotypes of interest for melon breeding, variation in shape, longer and rounder melons, and melons with a subtle aroma and formation of abscission layer with or without color change. c) Variation in triple response: M2 families showing a more intense response to ethylene than wild type and others segregating for the response.

Table 2 Comparison of the mutations found by TILLING in the gene targets screened in the two EMS-mutagenized melon populations, Piel de Sapo (PS) and Charentais (CH)

Gene	PS Population	CH population
	Exon/Intron-UTR	
<i>Cm-eIF4E</i>	3/1	5/9
<i>CmelF(iso)4E</i>	0/0	1/0
<i>Cm-PDS</i>	1/1	1/3
TOTAL	4/2	7/12

Four new alleles were reported for *Cm-eIF4E*, defined by 4 mutations. Two (in exons 1 and 2) were predicted to affect protein function according to SIFT ($p < 0.05$). The segregation of the missense mutations was analyzed in each mutant M2 family (Table 3) and homozygous individuals for each mutation are being produced for phenotypic analysis. Two new alleles were reported for *Cm-PDS*. Mutations were detected in the first intron and in the second exon of the amplicon (Figure 1). The exonic mutation was predicted to affect protein function.

The genotype/phenotype segregation was verified for the *Cm-PDS* mutations in 15 plants of each M2. The intronic mutation found in M2 family F1151 (mutation 20 in Figure 1) presented a Mendelian segregation of 4 wild type/7 heterozygous/4 homozygous mutant (fitting to a 1:2:1 ratio) (Table 3). No associated phenotype was observed, which agrees with the position of the mutation in an intronic region. The M2 family C384 carried a missense mutation in exon 2 (mutation 21 in Figure 1). The segregation of M2 was 9 heterozygous/6 wild type. In this family, we observed that some plants showed delayed growth and even death at the cotyledon stage, and therefore could not be sampled. M3 seeds from a heterozygous M2 plant were used to repeat segregation analysis. Fourteen plants germinated out of the 20 seeds sown. Again, plants with delayed growth were observed, but we also found a seedling with albino cotyledons and hypocotyl (Figure 4). This seedling was carefully grown until tissue could be sampled for DNA extraction. The albino was the only plant homozygous for the mutant allele; the remaining M3 plants were wild type or heterozygous. This mutant line showed a similar albino phenotype comparable to the previously reported phenotypes for *PDS* disruption [32].

To further evaluate the mutation rate of the Piel de Sapo population, we extended the TILLING screen to 4 additional genes involved in fruit quality. Mutations were found for all genes, except for *Cm-NOR*, with *Cm-DET1* being the most mutated gene. Six mutations were exonic, most determined to be missense, only one silent, but all were predicted to be tolerated (Figure 2, Table 3).

Mutation efficiency

We used these results to calculate the mutation rate of the population. A total of 14 point mutations were detected out of the total 11,534 bp analyzed. Previous studies reported the difficulty in tracking mutations on the ends of the fragments (~100 bp) [42]. Therefore, 200 bp was subtracted off each amplicon, and we considered the length screened in this study to be 9,134 bp. The overall mutation density was then calculated by dividing the total base pairs screened, which includes the sum of the total length of the 12 amplicon sizes \times the total number of individuals screened, by the total number of mutations revealed by TILLING $((9,134 \times 2,368)/14 = 1,544,950)$. It was calculated to be ~1/1.5 Mb. The mutation frequency as determined by TILLING is coherent with the moderate level of phenotypic mutants found in our population.

In order to compare the *inodorus* and the *cantalupensis* populations, the same genes were screened in a set of 2,483 lines of the Charentais population [30]. The number of mutations was higher, and new alleles were found for the three genes (Table 2 Figure 1). The spectrum of observed changes was similar to that found in the Piel de Sapo population (Table 3). Only one mutation was found for *Cm-eiF(iso)4E*. *Cm-eIF4E* was also the most mutated gene with 14 mutations: seven intronic, two in the 3'-UTR, and five exonic. Of the exonic mutations, two, located in exons 1 and 2, were not tolerated ($p < 0.05$). One intronic mutation was located in the first nucleotide of intron 3, and may affect a splicing site. TILLING analysis also revealed four mutations in *Cm-PDS*, one of which was not tolerated. The 19 point mutations discovered in the 3,871 bp analyzed in the Charentais population make a mutation rate of ~1/401 Kb $((3,071 \times 2,483)/19 = 401,331)$. This is about 3 times higher than that found in the Piel de Sapo population. This high mutation rate in the *cantalupensis* population has also been observed by screening additional genes in the final version of the population, which consists of 4,023 lines [30].

Discussion

The TILLING platform reported here is the first ever developed in an andromonoecious non-climacteric *inodorus* melon genetic background. The mutation rate in the Piel de Sapo population was moderate. It is less mutated than the *cantalupensis* population that was developed using a different genotype and a different protocol for treating the seeds, which resulted in increased seed viability at higher EMS doses (1 to 3%). Our attempt was the first with Piel de Sapo melons, and results suggest that our EMS treatment causes a high level of seed lethality. Regarding the protocol, Charentais seeds were not hydrated before being treated. Seed hydration has been reported to activate cellular division and induce the

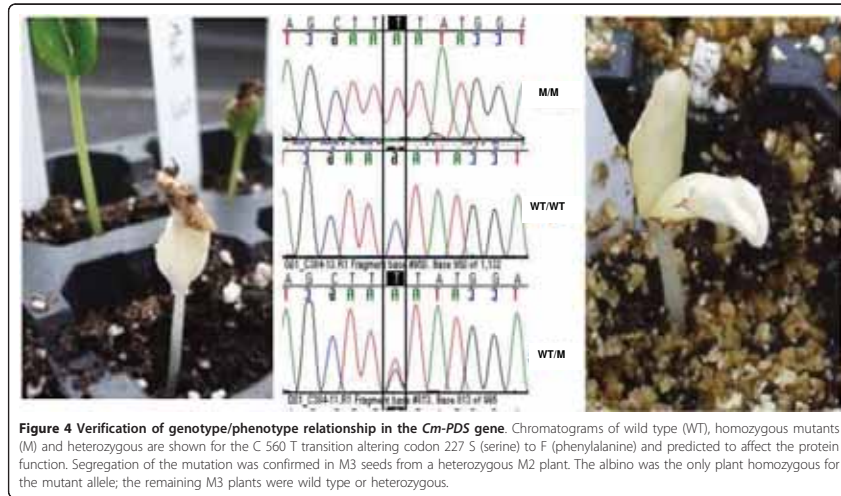
Table 3 Details of the mutations found by TILLING in the gene targets screened in the two EMS-mutagenized melon populations, Piel de Sapo (PS) and Charentais (CH)

Gene/ population	Amplicon size (bp)	Identified mutants	N ^o 1	Exon/ Intron	Position	AA change	M2 Family	M2 segregation (WT/ H/M) ²	SIFT score			
Cm-elF4E/PS	500	1	1	Exon 1	A 74 G	R 25 K	F2036	9/7/8	0.00			
	1169	3	2	Exon 2	G 385 A	E 129 K	F57	5/3/0	0.00			
			3	Exon 2	G 465 A	T 155 T	F1880					
			4	Intron 2	G/C	F959						
Cm-elF4E/CH	500	3	5	Exon 1	G 289 A	E 97 K	53		0.02			
	1169	11	6	Exon 1	C 294 T	F 98 F	2532					
			7	Intron 1	G/A		1420					
			8	Exon 2	C 401 T	A 134 V	1229		0.10			
			9	Exon 2	G 407 A	G 136 E	444		0.00			
			10	Intron 2	C/A		1287					
			11	Intron 2	C/T		1115					
			12	Intron 3	G/A first nt of intron	splicing	1667					
			13	Intron 3	C/T		2646					
			14	Intron 3	C/T		1233					
			15	Exon 4	G 616 A	E 206 K	2060		0.09			
			16	Intron 4	C/T		2035					
			17	3'UTR	T/C		1368					
			18	3'UTR	G/A		2163					
							222 TGA					
			Cm-elF(iso)4E/ PS	1014	0	-						
			Cm-elF(iso)4E/ CH	1014	1	19	Exon 1	G 128 A	R 43 K	M1587		1.00
			Cm-PDS/PS	1188	2	20	Intron 1	C/T		F1151	4/7/4	
21	Exon 2	C 560 T				S 227 F	C384	6/9/0	0.00			
Cm-PDS/CH	1188	4	22	Exon 1	C 395 T	S 172 F	1182		0.01			
			23	Intron 1	C/T		2495					
			24	Intron 1	T/C		1460					
			25	Intron 1	C/T		1226					
Cm-ACO1/PS	933	0	-									
	933	1	26	Exon 4	C 728 T	T 243 I	C142	10/6/3	1.00			
Cm-NOR/PS	1272	0										
Cm-DET1/PS	1185	1	27	Exon 2	T 402 A	L 135 L	I76					
	776	2	28	Exon 4	G 760 A	V 255 I	F201	2/11/2	0.37			
			29	Exon 4	C 848 T	S 284 L	F1876	2/4/3	0.25			
	1232	2	30	Intron 6	G 2981 A		F1248					
			31	Intron 7	C 3211 T		C487					
Cm-DHS/PS	578	1	32	Exon 1	G 184 A	V 62 I	F 1728	13/20/0	0.58			
	754	1	33	Exon 3	G 610 A	V 204 I	C348	2/6/0	0.27			

¹Number of mutations corresponds to those represented in Figures 1 and 2²For some mutations, it was difficult to obtain homozygous MM plants as seedlings showed altered development.

repair of DNA damage [43]. Differences in seed structure between genotypes, as well as an increase in the efficiency and accuracy of the repair of the induced lesions may explain the differences between populations. Perhaps,

using a higher EMS dosage and/or an extended length of mutagen application, combined with a more intense wash length, and a more accurate seedling nursery would have increased mutation efficiency. Other authors that



also found variable results in mutation frequency when using EMS within a crop suggest the use of alternative mutagens as a way for getting a higher density of mutations with less toxicity to the treated seeds [20].

Despite this higher mutation rate in the Charentais population, the efficiency of both populations in producing mutations predicted to be damaging to the proteins in *Cm-eIF4E* and *Cm-PDS* was similar. The new mutations found in *Cm-eIF4E* could be of interest, more so when this factor is highly conserved in eukaryotes. The natural diversity of *Cm-eIF4E* has been studied in melon and related species by EcoTILLING [31] and no variation has been reported within *Cucumis melo*, except for the *nsv* mutation (a point mutation of Leu228-His in exon 5), which controls resistance to *Melon necrotic spot virus* [44]. Variants in exons 1 to 3 were identified using EcoTILLING in another *Cucumis* species, *Cucumis zeyheri*, a wild relative of melon, isolated from the cultivated species by strong crossability barriers. Therefore, our new *Cm-eIF4E* alleles represent non-transgenic variants absent in nature. However it remains to be demonstrated if these mutations will lead to a gain of function, thereby producing new functional phenotypes, or to a loss of function. In other crops, recessive resistance to virus results from defective forms of eIF4E, for example in pepper (*pvr2*), pea (*sbm1*), tomato (*pot-1*), lettuce (*mo1*) and barley (*rym*) [45-47]. In most cases, resistance results from a few amino acid changes clustered in two neighboring regions of the

eIF4E structure, located near the structural pocket involved in CAP-binding [48]. Some of the coding changes identified in our TILLING assay (in exons 1 and 2) are located near these two regions, close to the conserved tryptophan residues required for CAP-binding activity (Figure 5). The other coding mutations are found in different regions of the gene, but most affect highly conserved aminoacids. EcoTILLING has also shown that most natural variation in *Cm-ACO1* occurs in exons 1, 2 and 3 [49]. No variants are found in large germplasm collections in exon 4. The effect of the identified mutation in exon 4 on ethylene production needs to be demonstrated.

The mutants found in *Cm-DET1* and *Cm-DHS* are the first reported for these genes in *inodorus* melons. It has been demonstrated that phenotypes of the tomato mutants high pigment-2dg (*hp-2dg*) and *hp-2j* are caused by lesions in the *DET1* gene. Point missense mutations and intron mutations directing alternative splicing have been reported in both the N-terminus and the C-terminus of the protein in *Arabidopsis* and tomato, suggesting that both ends of the protein are important for this function [50,51]. It has been reported that suppression of *DHS* delays loss of tissue integrity in senescing tomato fruit, leading to an extended shelf life. Further analysis will indicate if melon mutants *Cm-DET1* and *Cm-DHS* have alterations in fruit senescence. Segregation studies are now being performed along with phenotype analysis of the mutant lines.

work was supported by projects GEN2003-20237-C06-02/03, funded by the Ministerio de Educación y Ciencia (Spain), and GEN2006-27773-C2-1/2, funded by the ERA-PG programme.

Author details

¹IRTA, Centre de Recerca en Agrigenòmica CSIC-IRTA-UAB, Carretera de Cabrils Km 2, 08348 Cabrils (Barcelona), Spain. ²Institute for the Conservation and Breeding of Agricultural Biodiversity (COMAV-UPV), Universitat Politècnica de València, Camino de Vera s/n, 46022 Valencia, Spain. ³Instituto de Biología Molecular y Celular de Plantas (IBMCP), Universitat Politècnica de València (UPV)-Consejo Superior de Investigaciones Científicas (CSIC), Ciudad Politécnica de la Innovación (CPI), Ed. 8E, C/Ingeniero Fausto Elio s/n, 46022 Valencia, Spain. ⁴Unité de Recherche en Génétique Végétale (INRA-URGV), 2 rue Gaston Crémieux CP 5708, 91057 Evry Cedex, France. ⁵Department of Plant Science, School of Agriculture and Biology, Shanghai Jiaotong University, 200030 Shanghai, People's Republic of China.

Authors' contributions

JG-M, MG and AJM performed the EMS mutagenesis. FN and BP contributed to the development and conservation of the M2 seed. BP, CR and CE extracted the DNA and performed the phenotypic analysis. DNA pooling, TILLING screens and analysis, and PDS mutant phenotypic analysis were done by MG, CT and MX. MX and MP participated in the sequencing of the eIF4E mutants in the M2 population. JG-M, AB, BP, and FN coordinated the study. BP, MG, CE, and JG-M were primarily responsible for drafting and revising the manuscript with contributions from co-authors. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Received: 15 March 2011 Accepted: 11 August 2011
Published: 11 August 2011

References

- The International Cucurbit Genomics Initiative (ICuGI). [http://www.icugi.org].
- González-Ibeas D, Blanca J, Roig C, González-To M, Picó B, Truniger V, Gómez P, Deleu W, Caño-Delgado A, Arús P, Nuez F, García-Mas J, Puigdomènech P, Aranda MA: **MELOGEN: an EST database for melon functional genomics.** *BMC Genomics* 2007, **8**:306.
- Fita A, Picó B, Monforte A, Nuez F: **Genetics of Root System Architecture Using Near-isogenic Lines of Melon.** *J Am Soc Hortic Sci* 2008, **133**:448-458.
- Fernandez-Silva I, Eduardo I, Blanca J, Esteras C, Picó B, Nuez F, Arús P, García-Mas J, Monforte AJ: **Bin mapping of genomic and EST-derived SSRs in melon (*Cucumis melo* L.).** *Theor Appl Genet* 2008, **118**:139-150.
- Deleu W, Esteras C, Roig C, González-To M, Fernández-Silva I, Blanca J, Aranda MA, Arús P, Nuez F, Monforte AJ, Picó MB, García-Mas J: **A set of EST-SNPs for map saturation and cultivar identification in melon.** *BMC Plant Biol* 2009, **9**:90.
- Mascarell-Creus A, Cañizares J, Vilarrasa J, Mora-García S, Blanca J, González-Ibeas D, Saladié M, Roig C, Deleu W, Picó B, López-Bigas N, Aranda MA, García-Mas J, Nuez F, Puigdomènech P, Caño-Delgado A: **An oligo-based microarray offers novel transcriptomic approaches for the analysis of pathogen resistance and fruit quality traits in melon (*Cucumis melo* L.).** *BMC Genomics* 2009, **10**:467.
- Blanca JM, Cañizares J, Ziarolo P, Esteras C, Mir G, Nuez F, García-Mas J, Picó B: **Melon transcriptome characterization. SSRs and SNPs discovery for high throughput genotyping across the species.** *Plant Genome* 2011, **4**(2):118-131.
- González VM, Benjak A, Hénaff EM, Mir G, Casacuberta JM, García-Mas J, Puigdomènech P: **Sequencing of 6.7 Mb of the melon genome using a BAC pooling strategy.** *BMC Plant Biology* 2010, **10**:246.
- Moreno E, Obando JM, Dos-Santos N, Fernández-Trujillo JP, Monforte AJ, García-Mas J: **Candidate genes and QTLs for fruit ripening and softening in melon.** *Theor Appl Genet* 2007, **116**:589-602.
- Essafi A, Diaz-Pendón JA, Moriones E, Monforte AJ, García-Mas J, Martín-Hernández AM: **Dissection of the oligogenic resistance to Cucumber mosaic virus in the melon accession PI 161375.** *Theor Appl Genet* 2009, **118**:275-284.
- Comai L, Henikoff S: **TILLING: practical single-nucleotide mutation discovery.** *Plant J* 2006, **45**:684-94.
- Cooper JL, Till BJ, Laport RG, Darlow MC, Kleffner JM, Jamai A, El-Mellouki T, Liu S, Ritchie R, Nielsen N, et al: **TILLING to detect induced mutations in soybean.** *BMC Plant Biol* 2008, **8**(1):9.
- Dalmais M, Schmidt J, Le Signor C, Moussy F, Burstin J, Savoix V, Aubert G, de Oliveira Y, Guichard C, Thompson R, Bendahmane A: **UTILLDb, a *Psidium sativum* in silico forward and reverse genetics tool.** *Genome Biol* 2008, **9**:R43.
- Dierking EC, Bilyeu KD: **New sources of soybean meal and oil composition traits identified through TILLING.** *BMC Plant Biol* 2009, **9**:89.
- Perry J, Brachmann A, Welham T, Binder A, Charpentier M, Groth M, Haage K, Markmann K, Wang TL, Parniske M: **TILLING in *Lotus japonicus* identified large allelic series for symbiosis genes and revealed a bias in functionally defective ethyl methanesulfonate alleles toward glycine replacements.** *Plant Physiol* 2009, **151**(3):1281-1291.
- Caldwell DG, McCallum N, Shaw P, Muehlbauer GJ, Marshall DF, Waugh R: **A structured mutant population for forward and reverse genetics in Barley (*Hordeum vulgare* L.).** *Plant J* 2004, **40**(1):143-150.
- Henikoff S, Bradley JT, Comai L: **TILLING. Traditional mutagenesis meets functional genomics.** *Plant Physiol* 2004, **135**:630-636.
- Wu JL, Wu C, Lei C, Baraoidan M, Bordeos A, Madamba MR, Ramos-Pamplona M, Mauleon R, Portugal A, Ulat VJ, et al: **Chemical- and irradiation-induced mutants of indica rice IR64 for forward and reverse genetics.** *Plant Mol Biol* 2005, **59**(1):85-97.
- Slade AJ, Fuerstenberg SJ, Loeffler D, Steine MN, Facciotti D: **A reverse genetic, nontransgenic approach to wheat crop improvement by TILLING.** *Nat Biotechnol* 2005, **23**:75-81.
- Till BJ, Cooper J, Tai TH, Colowit P, Greene EA, Henikoff S, Comai L: **Discovery of chemically induced mutations in rice by TILLING.** *BMC Plant Biol* 2007, **7**:19.
- Xin Z, Wang ML, Barkley NA, Burow G, Franks C, Pederson G, Burke J: **Applying genotyping (TILLING) and phenotyping analyses to elucidate gene function in a chemically induced sorghum mutant population.** *BMC Plant Biol* 2008, **8**:103.
- Dong C, Dalton-Morgan J, Vincent K, Sharp P: **A modified TILLING method for wheat breeding.** *Plant Genome* 2009, **2**:39-47.
- Sestili F, Botticella E, Bedo Z, Phillips A, Lafiandra D: **Production of novel allelic variation for genes involved in starch biosynthesis through mutagenesis.** *Mol Breeding* 2010, **25**:145-154.
- Watanabe S, Mizoguchi T, Aoki K, Kubo Y, Mori H, Imanishi S, Yamazaki Y, Shibata D, Ezura H: **Ethylmethanesulfonate (EMS) mutagenesis of *Solanum lycopersicum* cv. Micro-Tom for large-scale mutant screens.** *Plant Biotech* 2007, **24**:33-38.
- Elias R, Till BJ, Miba CH, Al-Safadi B: **Optimizing TILLING and EcoTILLING techniques for potato (*Solanum tuberosum* L.).** *BMC Res Notes* 2009, **2**:141.
- Piron F, Nicolai M, Minoia S, Piednoir E, Moretti A, Salgues A, Zamir D, Caranta C, Bendahmane A: **An induced mutation in tomato eIF4E leads to immunity to two Potyviruses.** *PLoS ONE* 2010, **5**(6):e11313.
- Himelblau E, Gilchrist EJ, Buono K, Bizell C, Mentzer L, Vogelzang R, Osborn T, Amasino RM, Parkin IAP, Haughn J: **Forward and reverse genetics of rapid cycling *Brassica oleracea*.** *Theor Appl Genet* 2009, **118**:953-961.
- Stephenson P, Baker D, Girtin T, Perez A, Amoah S, King GJ, Østergaard L: **A rich TILLING resource for studying gene function in *Brassica rapa*.** *BMC Plant Biol* 2010, **10**:62.
- Pitrat M: **Melon (*Cucumis melo* L.).** In *Handbook of Crop Breeding Vol 1. Vegetables*. Edited by: Prohens J, Nuez F. New York:Springer; 2008:283-315.
- Dahmani-Mardas F, Troadec Ch, Boualem A, Leveque S, Alsdon AA, Aldoss AA, Dogimont C, Bendahmane A: **Engineering Melon Plants with Improved Fruit Shelf Life Using the TILLING Approach.** *PLoS ONE* 2010, **5**:e15776.
- Nieto C, Piron F, Dalmais M, Marco CF, Moriones E, Gómez-Guillamón ML, Truniger V, Gómez P, García-Mas J, Aranda MA, Bendahmane A: **EcoTILLING for the identification of allelic variants of melon eIF4E, a factor that controls virus susceptibility.** *BMC Plant Biol* 2007, **7**:34.
- Qin G, Gu H, Ma L, Peng Y, Deng XW, Chen Z, Qu LJ: **Disruption of phytoene desaturase gene results in albino and dwarf phenotypes in *Arabidopsis* by impairing chlorophyll, carotenoid, and gibberellin biosynthesis.** *Cell Res* 2007, **17**:471-482.
- Codons Optimized to Deliver Deleterious Lesions (CODDLE). [http://www.proweb.org/input].

34. Lasserre E, Bouquin T, Hernández JA, Bull J, Pech JC, Balague C: Structure and expression of three genes encoding ACC oxidase homologs from melon (*Cucumis melo* L.). *Mol Gen Genet* 1996, **251**(1):81-90.
35. Giovannoni JJ: Fruit ripening mutants yield insights into ripening control. *Curr Opin Plant Biol* 2007, **10**:1-7.
36. Davuluri GR, van Tuinen A, Mustilli AC, Manfredonia A, Newman R, Burgess D, Brummell DA, King SR, Pals J, Uhlrig J, Penning H, Bowler C: Manipulation of DET1 expression in tomato results in photomorphogenic phenotypes caused by post-transcriptional gene silencing. *Plant J* 2004, **40**:344-354.
37. Wei S, Li X, Gruber ML, Li R, Zhou R, Zebajadi A, Hannoufa A: RNAi-mediated suppression of DET1 alters the levels of carotenoids and sinapate esters in seeds of *Brassica napus*. *J Agric Food Chem* 2009, **57**(12):5326-5333.
38. Wang TW, Zhang CG, Wu W, Nowack LM, Madey E, Thompson JE: Antisense suppression of deoxyhypusine synthase in tomato delays fruit softening and alters growth and development DHS mediates the first of two sequential enzymatic reactions that activate eukaryotic translation initiation factor-5A. *Plant Physiol* 2005, **138**:1372-1382.
39. Ng PC, Henikoff S: SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Res* 2003, **31**(13):3812-3814.
40. Guzman P, Ecker JR: Exploiting the triple response of Arabidopsis to identify ethylene-related mutants. *The Plant Cell* 1990, **2**:513-523.
41. Henikoff S, Comai L: Single-nucleotide mutations for plant functional genomics. *Ann Rev Plant Biol* 2003, **54**:375-401.
42. Greene EA, Codomo CA, Taylor NE, Henikoff JG, Till BJ, Reynolds SH, Enns LC, Burtner C, Johnson JE, Odden AR, et al: Spectrum of chemically induced mutations from a large-scale reverse genetic screen in Arabidopsis. *Genetics* 2003, **164**(2):731-740.
43. Britt AB: DNA damage and repair in plants. *Annu Rev Plant Physiol Plant Mol Biol* 1996, **47**:75-100.
44. Truniger V, Nieto C, González-Ibeas D, Aranda M: Mechanism of plant eIF4E-mediated resistance against a Carmovirus (Tombusviridae): cap-independent translation of a viral RNA controlled in cis by an (a) virulence determinant. *Plant J* 2008, **56**(5):716-727.
45. Gao Z, Johansen E, Evers S, Thomas CL, Ellis THN, Maule AJ: The potyvirus resistance gene, sbm1, identifies a novel role for translation initiation factor eIF4E in cell-to-cell trafficking. *Plant J* 2004, **40**(3):376-385.
46. Kang BC, Yeam I, Frantz JD, Murphy JF, Jahn MM: The pvr1 locus in Capsicum encodes a translation initiation factor eIF4E that interacts with Tobacco etch virus VPg. *Plant J* 2005, **42**(3):392-405.
47. Ruffel S, Gallois J, Lesage M, Caranta C: The recessive potyvirus resistance gene pot-1 is the tomato orthologue of the pepper pvr2-eIF4 genes. *Mol Genet Genom* 2005, **274**(4):346-353.
48. Nicaise V, German-Retana S, Sanjuán R, Dubrana MP, Mazier M, Maisonneuve B, Candresse T, Caranta C, LeGall O: The Eukaryotic Translation Initiation Factor 4E Controls Lettuce Susceptibility to the Potyvirus Lettuce mosaic virus1. *Plant Physiol* 2003, **132**:1272-1282.
49. Esteras C, Pascual L, Saladie M, Dogimont C, Garcia-Mas J, Nuez F, Picó B: Use of Ecotilling to identify natural allelic variants of melon candidate genes involved in fruit ripening. *Proceedings Plant GEM8 Lisbon* 2009.
50. Levin I, Frankel P, Gilboa N, Tanny S, Lalazar A: The tomato dark green mutation is a novel allele of the tomato homolog of the DEETIOLATED1 gene. *Theor Appl Genet* 2003, **106**:454-460.
51. Kolotilin I, Koltai H, Tadmor Y, Bar-Or C, Reuveni M, Meir A, Nahon S, Shlomo S, Chen L, I Levin: Transcriptional profiling of high pigment-2dg tomato mutant links early fruit plastid biogenesis with its overproduction of phytonutrients. *Plant Physiol* 2007, **145**:389-401.

doi:10.1186/1756-0500-4-289

Cite this article as: González et al.: Towards a TILLING platform for functional genomics in Piel de Sapo melons. *BMC Research Notes* 2011 **4**:289.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

Research article

Open Access

MELOGEN: an EST database for melon functional genomics

Daniel Gonzalez-Ibeas¹, José Blanca², Cristina Roig², Mireia González-To³,
 Belén Picó², Verónica Truniger¹, Pedro Gómez¹, Wim Deleu³,
 Ana Caño-Delgado⁴, Pere Arús³, Fernando Nuez², Jordi Garcia-Mas³,
 Pere Puigdomènech⁴ and Miguel A Aranda*¹

Address: ¹Departamento de Biología del Estrés y Patología Vegetal, Centro de Edafología y Biología Aplicada del Segura (CEBAS)- CSIC, Apdo. correos 164, 30100 Espinardo (Murcia), Spain, ²Departamento de Biotecnología, Instituto de Conservación y Mejora de la Agrodiversidad Valenciana (COMAV-UPV), Camino de Vera s/n, 46022 Valencia, Spain, ³Departament de Genètica Vegetal, Centre de Recerca en Agrigenòmica CSIC-IRTA, Carretera de Cabrils Km2, 08348 Cabrils (Barcelona), Spain and ⁴Departament de Genètica Molecular, Centre de Recerca en Agrigenòmica CSIC-IRTA, Jordi Girona 18-26, 08034 Barcelona, Spain

Email: Daniel Gonzalez-Ibeas - agr030@cebas.csic.es; José Blanca - jblanca@btc.upv.es; Cristina Roig - croig@btc.upv.es; Mireia González-To - tmp2115@irta.es; Belén Picó - mpicosi@btc.upv.es; Verónica Truniger - truniger@cebas.csic.es; Pedro Gómez - pglopez@cebas.csic.es; Wim Deleu - wim.deleu@irta.es; Ana Caño-Delgado - acdgm1@cid.csic.es; Pere Arús - pere.arus@irta.es; Fernando Nuez - fnuez@btc.upv.es; Jordi Garcia-Mas - Jordi.Garcia@IRTA.Es; Pere Puigdomènech - pprgmp@ibmb.csic.es; Miguel A Aranda* - m.aranda@cebas.csic.es

* Corresponding author

Published: 3 September 2007

Received: 8 May 2007

BMC Genomics 2007, 8:306 doi:10.1186/1471-2164-8-306

Accepted: 3 September 2007

This article is available from: <http://www.biomedcentral.com/1471-2164/8/306>

© 2007 Gonzalez-Ibeas et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.**Abstract**

Background: Melon (*Cucumis melo* L.) is one of the most important fleshy fruits for fresh consumption. Despite this, few genomic resources exist for this species. To facilitate the discovery of genes involved in essential traits, such as fruit development, fruit maturation and disease resistance, and to speed up the process of breeding new and better adapted melon varieties, we have produced a large collection of expressed sequence tags (ESTs) from eight normalized cDNA libraries from different tissues in different physiological conditions.

Results: We determined over 30,000 ESTs that were clustered into 16,637 non-redundant sequences or unigenes, comprising 6,023 tentative consensus sequences (contigs) and 10,614 unclustered sequences (singletons). Many potential molecular markers were identified in the melon dataset: 1,052 potential simple sequence repeats (SSRs) and 356 single nucleotide polymorphisms (SNPs) were found. Sixty-nine percent of the melon unigenes showed a significant similarity with proteins in databases. Functional classification of the unigenes was carried out following the Gene Ontology scheme. In total, 9,402 unigenes were mapped to one or more ontology. Remarkably, the distributions of melon and Arabidopsis unigenes followed similar tendencies, suggesting that the melon dataset is representative of the whole melon transcriptome. Bioinformatic analyses primarily focused on potential precursors of melon micro RNAs (miRNAs) in the melon dataset, but many other genes potentially controlling disease resistance and fruit quality traits were also identified. Patterns of transcript accumulation were characterised by Real-Time-qPCR for 20 of these genes.

Conclusion: The collection of ESTs characterised here represents a substantial increase on the genetic information available for melon. A database (MELOGEN) which contains all EST sequences, contig images and several tools for analysis and data mining has been created. This set of sequences constitutes also the basis for an oligo-based microarray for melon that is being used in experiments to further analyse the melon transcriptome.

Background

Melon (*Cucumis melo* L.) is an important horticultural crop grown in temperate, subtropical and tropical regions worldwide. Melon is among the most important fleshy fruits for fresh consumption, its total production in 2004 exceeding 874 million metric tons, of which 72.5% are produced in Asia, 11.7% in Europe, 8.4% in America and 6.1% in Africa, being a significant component of fresh fruit traded internationally [1]. Melon belongs to the *Cucurbitaceae* family, which comprises up to 750 different species distributed in 90 genera. Species in this family include watermelon, cucumber, squash and marrow, all of them cultivated essentially because of their fruits, but this family also includes species of interest for other reasons, as, for example, their contents in potentially therapeutic compounds (e.g. *Momordica charantia*) [2]. Melon is a diploid species, with a basic number of chromosomes $x = 12$ ($2x = 2n = 24$) and an estimated genome size of 450 to 500 Mb [3], similar in size to the rice genome (419 Mb) [4,5] and about three times the size of the Arabidopsis genome (125 Mb) [6]. Melon has been classified into two subspecies, *C. melo* ssp. *agrestis* and *C. melo* ssp. *melo* with India and Africa being their centres of origin, respectively [7,8].

Melon has a great potential for becoming a model for understanding important traits in fruiting crops. Melon fruits have wide morphological, physiological and biochemical diversity [7,9] which can be exploited to dissect biological processes of great technological importance, among them flavour development and textural changes that occur during fruit ripening. The contemporary melon cultivars can be divided into two groups, climacteric and nonclimacteric, according to their ripening patterns [10]. Climacteric fruits are characterized by rapid and profound changes during ripening associated to increased levels of respiration and release of ethylene, whereas the nonclimacteric varieties do not produce ethylene and have long shelf-life. Analyses of climacteric and nonclimacteric melons have illustrated the process of aroma formation [11-14] and the temporal sequence of cell wall disassembly [15-17]. Melon can be also a very useful experimental system to analyse other aspects of fundamental plant biology. For example, melon and other cucurbits have been used to analyse the development of the plant vasculature and the transportation of macromolecules through it [18-20], and different interactions between melon and pests and pathogens have been characterised with varying depths [21-27].

Important genetic tools have been described for melon, as for example linkage genetic maps [28,29] and the development of a genomic library of near isogenic lines (NILs) from an exotic accession [30]; also, biotechnology is feasible in melon [31-33]. However, the great majority of

genes involved in the aforementioned traits are yet to be identified in melon. Partial sequencing of cDNA inserts of expressed sequence tags (ESTs) have been used as an effective method for gene discovery. By sequencing clones derived from RNA from different sources, and/or by normalizing cDNA libraries, the total set of genes sampled can be maximized. Bioinformatic analysis, annotation and clustering of sequences could yield databases which mining can be used to select candidate genes implicated in traits of interest. EST collections can also serve to construct microarrays useful for identifying sets of plant genes expressed during different developmental stages and/or responding to environmental stimuli [34,35]. In addition, EST collections are good sources of simple sequence repeats (SSRs) and single-nucleotide polymorphisms (SNPs) that can be used for creating saturated genetic maps [36,37]. Thus, EST collections have been generated for many plant species, being the most comprehensive those of Arabidopsis [6] and rice [38]. Fruit crops have been less extensively surveyed, but important collections are publicly available for several species, including tomato [39], apple [40], grape [41] and citrus [42].

Despite the importance of the family *Cucurbitaceae*, relatively little EST information is currently available: only 16,039 nucleotide sequences have been annotated from the whole *Cucurbitaceae* family in the publicly accessible GenBank database as of November 2006; out of these, 12,180 correspond to the *Cucumis* genus and 6,061 to melon. These numbers are in sharp contrast with the data available for families composed of other important food crops like *Solanaceae* (1,020,102 sequences), *Fabaceae* (1,466,518 sequences), *Brassicaceae* (1,010,148 sequences excluding Arabidopsis), *Vitaceae* (449,478) and *Rosaceae* (390,066 sequences). Here we describe a public EST sequencing project in melon. We report the determination and analysis of 30,675 high-quality melon ESTs, sequenced from eight normalized cDNA libraries corresponding to different tissues in different physiological conditions. We have classified the sequences into functional categories and described SSRs and SNPs of potential use in genetic maps and marker-assisted breeding programs. A database which contains all EST sequences, contig images and several tools for analysis and data mining has been created. In addition, we have analyzed the EST melon dataset to identify candidate genes potentially coding microRNAs or involved in fruit maturation processes and pathogen defence. The pattern of transcript accumulation in different physiological conditions has been characterised by Real-Time-qPCR for 20 of these candidate genes.

Results

EST Sequencing and Clustering

Eight cDNA libraries were constructed using material from "Piel de Sapo" Spanish cultivars, the C-35 cantaloupe line (both belonging to *Cucumis melo* L. ssp. *melo*) and the accession pat81 of *C. melo* L. ssp. *agrestis* (Naud.) Pangalo. The sources of RNA to construct each library were fruits of 15 and 46 days after pollination (dap), leaves, photosynthetic cotyledons inoculated with *Cucumber mosaic virus* (CMV), healthy roots and *Monosporascus cannonballus* Pollack et Uecker (the causal agent of melon vine decline) infected roots (Table 1). Approximately 3,700 sequences were determined from each library by single-pass 5' sequencing, except for the library prepared from CMV infected cotyledons for which approximately 6,600 sequences were determined, yielding a total of 33,292 raw sequences. Processing to eliminate vector sequences, low quality chromatograms and sequences of less than 100 base pairs (bp) gave rise to 29,604 good quality expressed sequence tags (ESTs) (Table 2) implying a cloning success of approximately 89%. The average edited length was 674 bp, and only a 6.4% of the sequences had less than 350 bp.

Clustering of the sequences using default parameters of the EST analysis pipeline EST2uni [43] yielded 6,023 tentative consensus sequences (also called contigs) and 10,614 unclustered sequences (also called singletons), with a total of 16,637 non-redundant sequences or unigenes (Table 2). All good quality ESTs were used for clustering, independently of the melon genotype of origin, because single nucleotide polymorphisms (SNPs) were expected among genotypes. The number of ESTs per unigene was between 1 and 44 (1 case), with an average of 1.8 ESTs per contig, as a high proportion of contigs (4,886 out of 6,023) contained less than 5 ESTs and contigs with more than 8 ESTs were scarce (Fig. 1A). Therefore, redundancy values were notably low (around 16%). The unigene length varied between 101 bp and 2,664 bp, averaging 751 bp (Fig. 1B). Library specific unigenes were about one third of the total for each library (Table 2). A second round of clustering yielded 14,480 unigene clusters, referred to as superunigenes. A web integrated data-

base that contains all EST sequences, contig images and several tools for analysis and data mining has been created and named MELOGEN [44]. Codon usage was estimated using this EST collection. As expected, the codon usage of melon was very similar to that of Arabidopsis and other dicots. The preferred stop codon was UGA occurring in the 48% of the sequences. Suppression of the CG dinucleotide in the last two codon positions is very frequent in dicots, possibly as a consequence of methylation of C in the CG dinucleotide, resulting in an increased mutation rate [45]; in agreement with these data, the ratio XCG/XCC for melon was 0.52, very similar to the corresponding figure for tomato (0.58), pea (0.51), potato (0.48) and other dicots [45].

Libraries obtained from tissues inoculated with *M. cannonballus* were expected to contain sequences from the fungus. To estimate the proportion of sequences of fungal origin in these libraries, BLAST analyses against a database with plant and fungal sequences were carried out [46]. Only 56 sequences from these libraries were found to have a more significant similarity with fungal sequences than with plant sequences (Table 3). Consequently, these sequences were considered of fungal origin [46].

SSRs and SNPs

We have analysed the nature and frequency of microsatellites or simple sequence repeats (SSRs) in the melon sequence dataset. A search for repeats of two, three or four nucleotides in the dataset yielded 1,052 potential SSRs. Approximately, 6% of the unigenes contained at least one of the considered SSRs motifs, with repeats of three nucleotides being prevalent (Table 4). The maximum and minimum lengths of the repeats were 68 and 17 nucleotides, respectively, and the average length was 26 nucleotides. The most common repeat among dinucleotides was, by far, the AG repeat, constituting the 83% (Table 4). Repeats of AT and AC dinucleotides followed, with approximately 9% and 7%, respectively. Among the trinucleotide repeats, the most frequent was AAG (66%, Table 4), and the least frequent was ACT (0.6%, Table 4). Among tetranucleotide repeats, the most frequent was AAAG (51%, Table 4). A high proportion of SSRs (29.5%)

Table 1: Description of cDNA libraries

Name	Subspecies/cultivar/accession	Tissue/physiological condition
15d	<i>Ssp. melo</i> cv. "Piel de Sapo" T-111	Fruit 15 days after pollination
46d	<i>Ssp. melo</i> cv. "Piel de Sapo" T-111	Fruit 46 days after pollination
A	<i>Ssp. agrestis</i> accession pat81	Roots
AI	<i>Ssp. agrestis</i> accession pat81	Roots infected with <i>M. cannonballus</i>
CI	<i>Ssp. melo</i> var. cantaloupe accession C-35	Photosynthetic cotyledons infected with CMV
HS	<i>Ssp. melo</i> var. cantaloupe accession C-35	Leaves
PS	<i>Ssp. melo</i> cv. "Piel de Sapo" Piñonet torpedo	Roots
PSI	<i>Ssp. melo</i> cv. "Piel de Sapo" Piñonet torpedo	Roots infected with <i>M. cannonballus</i>

Table 2: EST statistics

Library	Raw sequences	Good-Quality ESTs	EST length	Singletons	Contigs	Unigenes	Redundancy (%)	Library-specific unigenes	Novelty (%)
15d	3,936	3,582	608.1 ± 175.2	1,009	1,930	2,939	18	1,100	37
46d	3,840	3,493	583.0 ± 161.1	1,000	1,854	2,854	18	1,063	37
A	3,936	3,666	700.0 ± 185.4	1,289	1,900	3,189	13	1,365	43
AI	3,647	3,255	756.3 ± 137.1	928	1,688	2,616	20	1,005	38
CI	6,605	5,664	651.4 ± 205.7	2,089	2,590	4,679	17	2,264	48
HS	3,648	3,012	669.3 ± 171.1	939	1,609	2,548	15	998	39
PS	3,840	3,377	679.9 ± 198.7	1,179	1,766	2,945	13	1,258	43
PSI	3,840	3,555	749.3 ± 156.2	1,279	1,826	3,105	13	1,363	44
	33,292	29,604		10,614	6,023	16,637			

were found in open reading frames (ORFs), though an analysis of the localization of di-, tri- and tetranucleotides separately showed that di- and tetranucleotides localised preferentially in untranslated regions (UTRs), whereas trinucleotides localised in both, UTRs and ORFs (Table 5).

Single nucleotide polymorphisms (SNPs) are the most abundant variations in genomes and, therefore, constitute a powerful tool for mapping and marker-assisted breeding. We initially identified in the melon sequence dataset 14,074 single nucleotide sequence variations and therefore potential SNPs (pSCH; Table 6) distributed in 4,663 contigs; however, these variations would include high-quality SNPs (pSNP) but also sequencing errors and mutations introduced during the cDNA synthesis step. Using more stringent criteria, these figures were substantially reduced: Putative SNPs were annotated only when the least represented allele was present in at least two EST sequences from the same genotype in a given contig and showing the same base change. Two accessions of the same cultivar (cv. "Piel de sapo") represented 47.3% of the sequences, but more than one half of the sequences were from two other more distant genotypes, the C-35 cantaloupe accession (29.3%) and the pat81 agrestis accession (23.4%). Thus, a total of 356 high-quality SNPs were found in 292 contigs, averaging 1.2 SNPs per contig.

Table 3: ESTs showing significant similarity with fungal sequences

Library	Number of ESTs
15d	0
46d	0
A	2
AI	26
CI	0
HS	1
PS	1
PSI	30
Total	60

Transitions were much more common than transversions. There were 117 AG and 112 CT transitions compared with 28 AC, 37 AT and 33 GT transversions (Table 6). CG transversions were not detected. The MELOGEN database [44] includes a tool for designing oligonucleotide primers to amplify the region containing the polymorphism to generate the corresponding molecular marker.

Functional annotation

In order to identify melon unigenes potentially encoding proteins with known function, we carried out a BLASTX analysis [47] of the sequence dataset against the databases listed in Table 7. Out of the 13,019 unigenes with a hit with proteins in databases, 11,431 (68.7%) unigenes showing an E value of $\leq 1e-10$ were annotated. On the other hand, 31.3% of the unigenes did not show significant similarity to any protein in the databases and, therefore, were not annotated.

Additionally, we performed a functional classification of the unigenes following the Gene Ontology scheme. Gene Ontology provides a structured and controlled vocabulary to describe gene products according to three ontologies: molecular function, biological process and cellular component [48]. To do that, we added GO terms based on the automated annotation of each unigene using the Arabidopsis database [6]. A summary of the results with the percentage of unigenes annotated in representative categories corresponding to the GO slim terms [48] is shown, as well as a comparison of the distribution of melon and Arabidopsis unigenes (Fig. 2). The distributions of melon and Arabidopsis unigenes follow similar tendencies, suggesting that the melon dataset is representative of the whole melon transcriptome. In total, 9,402 unigenes could be mapped to one or more ontologies, with multiple assignments possible for a given protein within a single ontology. A high percentage of unigenes in both species was classified as "unknown function". Out of the 9,791 assignments made to the cellular component category, 25.8% corresponded to membrane proteins and 17.8% to plastidial proteins (Fig. 2A). Under the molecu-

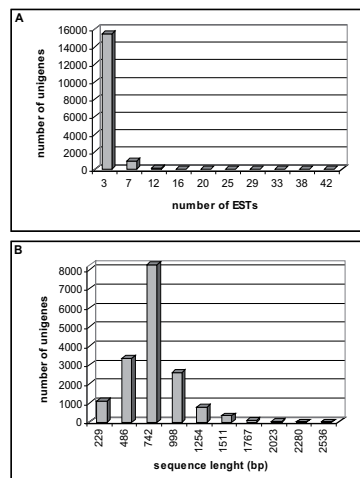


Figure 1
Unigenes statistics. (A) Distribution of melon ESTs among unigenes (contigs and singletons). (B) Size distribution of melon unigenes.

lar function category, assignments were mainly to catalytic activity (23.0%) and to hydrolase activity (14.7%) (Fig. 2B). The distribution of unigenes under the biological process category was more uniform, with 19.9% of assignments to cellular process and 12.7% to biosynthesis (Fig. 2C).

We have also identified 6,673 (40.1%) melon unigenes with an ortholog in the Arabidopsis database, and a HMMER motif has been assigned to 4,655 (28.0%) unigenes by comparisons with the Pfam database [49] (Table 7). All these results are compiled in the MELOGEN database, which also contains direct links to the databases used to carry out analyses.

Genes potentially encoding microRNAs

Central to RNA silencing are small RNA molecules (sRNAs) that can arise from endogenous or exogenous sources from precursors with double-stranded RNA (dsRNA) pairing. One class of such sRNAs are microRNAs (miRNAs), which originate from endogenous long self-complementary precursors that mature in a multi-step process involving many enzymes [50,51]. Recently, a

Table 4: Simple sequence repeats (SSRs) statistics*

Dinucleotide repeat	Number of di-pSSR	%
AG	205	83.3
AT	23	9.4
AC	18	7.3
Total	246	100
Trinucleotide repeat	Number of tri-pSSR	%
AAC	41	5.7
AAG	471	66.0
AAT	22	3.1
ACC	18	2.5
ACG	14	2.0
ACT	4	0.6
AGC	25	3.5
AGG	54	7.6
ATC	48	6.6
CCG	17	2.4
Total	714	100
Tetranucleotide repeat	Number of tetra-pSSR	%
AAAC	8	8.7
AAGG	7	7.6
AATC	4	4.3
AATG	3	3.3
AATT	2	2.2
ATCG	2	2.2
ACTC	1	1.1
AAGC	1	1.1
ACAT	1	1.1
AAAG	47	51.1
AAAT	16	17.3
Total	92	100

*The number of di-, tri- and tetranucleotide repeats identified in the melon database is shown for the complete set of putative SSRs (pSSRs).

comprehensive strategy to identify new miRNA homologs in EST databases has been developed [52,53]. We have followed this strategy to identify potential melon miRNAs. A total of 20 ESTs that contained homologs to miRNAs in the microRNA Registry database [54] were identified and grouped into 12 contigs and, after manual inspection of secondary foldback hairpin structure, 5 unigenes were selected (Table 8). Contig sequences varied between 536 and 840 nucleotides long, and had negative folding free energies of -206.8 to -160.8 kcal mol⁻¹ (Table 8) according to MFOLD [55], which are in the range of the computational values of Arabidopsis miRNA precursors [52]. Their predicted secondary structures showed that there were at least 16 nucleotides paired between the sequence of the potential mature miRNA and its opposite arm (miRNA*) in the corresponding hairpin structure (Fig. 3). The location of the potential miRNAs varied among ESTs, 4 were found in the sense orientation of the EST, 1 was found in

Table 5: Localization of simple sequence repeats (SSRs) with respect to putative initiation and termination codons in the melon sequence dataset*

	Dinucleotide repeats		Trinucleotide repeats		Tetranucleotide repeats		All SSRs analyzed	
	No.	%	No.	%	No.	%	No.	%
5'-UTR	65	82.3%	121	51.7%	18	62.1%	204	59.6%
ORF	0	0.0%	99	42.3%	2	6.9%	101	29.5%
3'-UTR	13	16.5%	13	5.6%	9	31.0%	35	10.2%
Other†	1	1.3%	1	0.4%	0	0.0%	2	0.6%
Total	79	100.0%	234	100.0%	29	100.0%	342	100%

*Only full length melon unigenes were used for this analysis. Full length unigenes were automatically selected by checking the presence of the start codon by comparison of the 5'-terminal region of the unigene with the corresponding Arabidopsis gene used for annotation. †Imprecise localization of the SSR with respect to putative initiation or termination codons.

the antisense orientation. We have also searched for potential targets of the potential miRNAs in the melon EST dataset, identifying 3 of them (Table 8). However, minimal folding free energy indexes (MFEIs) [53] were below the -0.85 cut-off value proposed by Zhang et al. [53] only for m12 (Table 8). Potential melon miRNA m12 has a precursor of 536 nt in length and codes for a melon ortholog of the Arabidopsis miR319. miR319 targets a transcription factor of the TCP family [56,57]; in the melon dataset, an ortholog of this Arabidopsis gene has been found in a unigene annotated as a TCP transcription factor. In this case, the melon miRNA and its potential target have a pattern of paired/non-paired bases between the target and the miRNA identical to the corresponding target-miRNA pattern in Arabidopsis (data not shown).

Genes potentially encoding pathogen resistance and fruit quality traits

Pathogens affect severely the productivity of melon crops. Three of the cDNA libraries sequenced here correspond to pathogen-infected tissues and, thus, should contain transcripts from genes whose expression is induced in

response infection. We have carried out a bioinformatics search for homologs of genes involved in pathogen resistance response (see [58] for a review) and virus susceptibility [59-61], finding among them at least one melon ortholog to the Arabidopsis FLS2 receptor [62], several unigenes potentially encoding disease resistance proteins as well as mitogen-activated protein kinases, homologs to translation initiation factors constituting potential virus susceptibility factors, etc. [see Additional file 1].

Fruit development and ripening are the most important processes determining the fruit quality traits of fleshy fruits like melon. At present most of the molecular and genetic data available about fruit development and ripening come from tomato [63,64] and Arabidopsis [65,66]. In recent years, several genes and quantitative trait loci controlling fruit quality traits have been described in melon [67,68]. As for developmental processes, homologs to genes involved in melon fruit development, ripening and quality have been found in the melon dataset. These include several MADS-box genes, homologs to the *fw2.2* and *ovate* QTLs [69,70], several homologs to

Table 6: Single nucleotide polymorphisms (SNPs) statistics*

Variation	Number	%		
pSNP transversions:	127	35.7		
pSNP transitions:	229	64.3		
pSCH transversions:	4,273	30.4		
pSCH transitions:	9,801	69.6		
Mutation	Number of pSNP	Mutation	Number of pSCH	
A<->C	28	A<->C	859	
A<->G	117	A<->G	5,860	
A<->T	37	A<->T	1,897	
C<->G	29	C<->G	431	
C<->T	112	C<->T	3,941	
G<->T	33	G<->T	1,086	

*Type and number of transition and transversions are shown for putative single nucleotide variations in sequence (pSCH) and for putative high-quality single nucleotide polymorphism (pSNP) identified in the melon database.

Table 7: Functional annotation statistics

A. Number of unigenes with BLAST hits		
Database*	Number of unigenes	%
Arabidopsis	11,724	70.5
Cucurbitaceae	5,340	32.1
Uniref90	11,893	71.5
Any database	13,019	78.3
B. Number of unigenes with HMMER hits		
Database*	Number of unigenes	%
Pfam	4,655	28.0
C. Number of unigenes with orthologue		
Database*	Number of unigenes	%
Arabidopsis	6,673	40.1

* Databases searched were: Arabidopsis [6,109]; Uniref90 [110,115]; Cucurbitaceae: all cucurbitaceae sequences available in the National Center for Biotechnology Information (NCBI); any database: results using Arabidopsis, cucurbitaceae and Uniref90 databases all together; Pfam [49,112].

members of the SBP-box gene family to which the major tomato ripening gene *COLORLESS NON-RIPENING* belongs [71], several ACC synthase and ACC oxidase genes, unigenes from several cell wall-metabolism enzymes, etc [see Additional file 1].

Expression analysis of selected ESTs by Real-Time-qPCR

The accumulation of transcripts for 20 selected genes was analyzed by reverse transcription Real-Time-qPCR. ESTs for this analysis were preferentially chosen among those showing significant similarity with genes related to response to infection and fruit quality characteristics in melon and other species, and included *CTLI*, *EIF4A-2*, *EIF4E*, *EIN4*, *GA2OX1*, *HSP101*, *HSP70*, *IAA9*, *LSM1*, *LUT2*, *NCBP*, *SVP*, *HIR*, *TCH4*, *TIP4*, *TOM1*, *TOM2A*, *TOM3*, *UGE5* and *WRKY70* (Table 9). Preliminary experiments were carried out to choose between *GAPDH* and *CYCLOPHILIN (CYP7)* RNAs as endogenous controls; results showed that the *CYP7* RNA levels varied the least among treatments (data not shown) and, therefore, transcript accumulation levels were expressed relative to *CYP7* RNA levels.

Figure 4A illustrates the alteration of the RNA accumulation levels of selected genes that occurred in photosynthetic cotyledons after CMV infection. A significant increase in the level of transcripts from *HSP101*, *HSP70*, *HIR*, *TOM2A*, *WRKY70* and *EIN4* was observed; for *HSP101*, *HSP70*, *WRKY70* and *EIN4*, transcript accumulation levels in inoculated cotyledons were up to five times greater than in uninoculated controls (Fig. 4A). All of

these genes, except *TOM2A*, have been shown to be responsive to virus infection in other hosts [72-74]. Notably, the expression of *EIF4E*, known to be required for MNSV multiplication [27], remained unaltered. A shutoff of host gene expression also occurs in association with virus infection [75]; for the set of genes analysed here, only *GA2OX1* and *NCBP* responded to CMV infection with a reduction in the accumulation of their transcripts.

The response of selected genes in roots inoculated with *M. cannonballus* was analysed in melon genotypes known to be susceptible (cultivar "Piel de sapo"; Fig. 4B) and partially resistant (accession pat81 of *C. melo* L. ssp. *agrestis*; Fig. 4C) to the infection by this fungus. The patterns of transcript accumulation resulted clearly different for both genotypes. For pat81 (resistant), transcription factors *WRKY70* and *SVP* increased their expression between 2 and 3 times after inoculation; other stress-inducible genes (*HSP101*, *HSP70*) showed only a moderate increase (Fig. 4C). For "Piel de sapo" (susceptible), accumulation of *WRKY70* and *SVP* transcripts only increased about 1.5 times after inoculation whereas the expression of *HSP101* showed a marked increase (Fig. 4B). It is also worth noting the differential response of the *GA2OX1* gene in the two genotypes. Expression of *GA2OX1* increased about 1.5 times in pat81 roots after the *M. cannonballus* attack, whereas it decreased in "Piel de sapo" roots after fungal infection (compare Figs. 4B and 4C).

Comparison of patterns of transcript accumulation at two stages of fruit development showed increased levels of

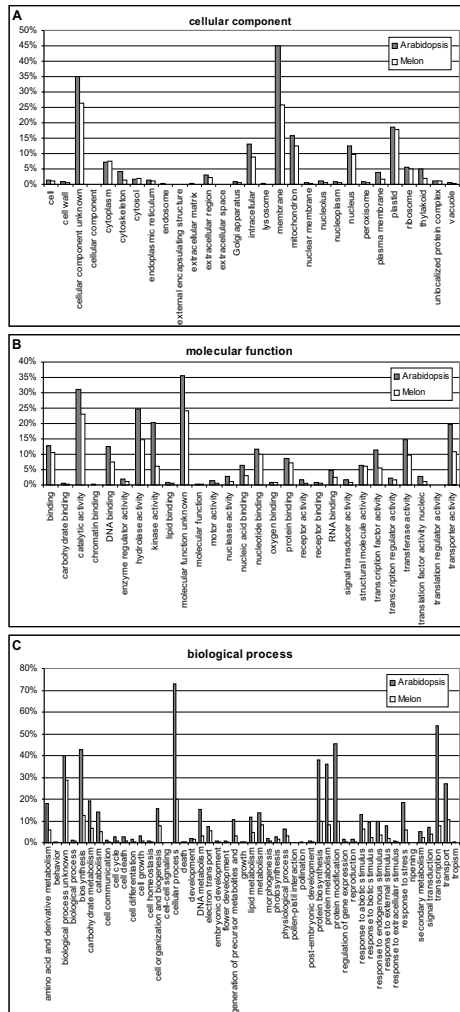


Figure 2
Distribution of melon and Arabidopsis unigenes according to the Gene Ontology scheme for functional classification of gene products.

Table 8: Potential melon miRNAs

Name	Potential mature miRNA sequence (5'->3')*	Melogen unigene [†]	Precursor folding free energy (kcal mol ⁻¹) [‡]	MFEI (kcal mol ⁻¹) [‡]	miRNA family	Potential target in melon
m2	ugaagcugccagcaugaucu	bCL2353Contig1	-206.8	-0.67	miR167	--
m4	ugauugagccgugccaaauac	bPSI_40-F10-M13R_c	-195.0	-0.66	miR171	bHS_39-C12-M13R_c
m7	ucggaccagguucauuucccc	bA_31-D02-M13R_c	-160.8	-0.70	miR166	--
m8	uugacagaagauagagagcac	bCI_04-H02-M13R_c	-188.2	-0.74	miR157	bCI_30-A09-M13R_c
m12	uuggacugaaggagucuccu	b15d_24-H05-M13R_c	-163.6	-0.86	miR319	bCL2243Contig1

*Nucleotide sequences correspond to potential mature miRNA sequences as deduced from BLAST searches using known plant miRNA sequences [54] and analysis of secondary structure predictions [52]. [†]Accession numbers (MELOGEN database) of potential precursors of melon miRNAs are given. [‡]Computational values of folding free energies have been calculated using MFOLD 3.1 [55]. [‡]Minimal folding free energy indexes (MFEIs) have been calculated as described by Zhang et al. [53].

gene expression for 9 of the analysed genes. This was particularly evident for *HSP70*, *TOM2A*, *TOM3*, *EIN4* and *IAA9*. In contrast, decreased levels of transcript accumulation were observed for the other 11 genes.

Discussion

In this paper we provide an initial platform for functional genomics of melon by the identification of more than 16,000 unigenes assembled from almost 30,000 ESTs sequenced from 8 melon cDNA libraries. It is probably premature to estimate the proportion of melon genes represented in this dataset, but based on available data for other plant species (i.e. Arabidopsis and rice), it is likely that the melon unigene set characterised here represents approximately between half and one-third of the number of expressed, protein coding genes of melon. Libraries were constructed from various tissue types, but with a bias towards fruit development and pathogen-infected tissues. Data from these libraries will become a useful resource of

genes for experiments aimed at understanding important processes involved in fruit development and resistance to viral and fungal pathogens. Also, data presented here provide an important tool for generating markers to saturate melon genetic maps.

In contrast to typical EST gene-sampling strategies reported previously, we have found a low degree of redundancy in the sequences determined. The process of clustering reduced the number of sequences to 56%, from 29,604 good quality ESTs to 6,023 contigs and 10,614 singletons. Contigs with more than 8 ESTs were scarce, the majority of them being formed by 3 or 2 ESTs. Redundancy of the sequences derived from each library ranged from 13% to 20%, with singletons constituting approximately one third of the unigenes determined per library. This low redundancy is probably due to the success of the normalization process, responsible for the suppression of superabundant transcripts specific for a given tissue or condition. Normalization precludes *in silico* analysis of gene expression, but greatly increases the number of unigenes that can be determined by reducing redundancy [76]. Here we have used a recently described normalization protocol which is based on the cleavage of DNA or DNA-RNA duplexes by a specific DNase [77]; this process, in our hands, has proven simple, reproducible and efficient. Another factor that has contributed to the low redundancy values obtained has been the sequencing of libraries from very distinct tissues. Thus, the number of library specific unigenes was about one half of the total number of unigenes contributed by each library, suggesting that further sequencing of the libraries still has the potential to provide a good number of new, non-redundant sequences.

cDNA sequences are a useful source of SSRs, which are excellent molecular markers due to their high degree of polymorphism. A common feature of cDNA sequences obtained from plants is the high frequency of SSRs that they contain [36]. We have identified more than 1,000 potential SSRs in the melon dataset, with approximately

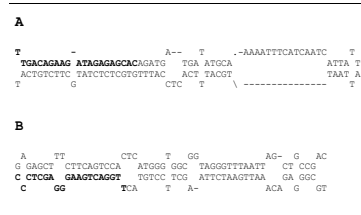


Figure 3
Potential precursors of melon microRNAs. (A) Stem loop sequence of putative precursor miRNA corresponding to unigene bCI_04-H02-M13R_c. (B) Stem loop sequence of putative precursor miRNA corresponding to unigene b15d_24-H05-M13R_c. The mature miRNA sequences are shown in bold.

Table 9: Transcripts selected from the database for gene expression analysis by Real Time qPCR

Gene	Melogen unigene	Sequence length (bp)	Arabidopsis locus	Aminoacid similarity (%)	Annotation (HMHR domain)
<i>CTL1</i>	bCL1465Contig1	1,387	AT1G05850	66.8	Chitinase-like protein 1, similar to class I chitinase (Glyco_hydro_19)
<i>CYP 7</i>	bCL3337Contig1	801	AT5G58710	78.4	Peptidyl-prolyl cis-trans isomerase, cyclophilin
<i>EIF4A-2</i>	bCL2906Contig1	1,146	AT1G54270	92.6	Eukaryotic translation initiation factor 4A-2 (DEAD)
<i>EIF4E</i>	bCL4710Contig1	815	AT4G18040	77.5	Eukaryotic translation initiation factor 4E 1 (IF4E)
<i>EIN4</i>	bCL1742Contig1	829	AT3G04580	55.7	Ethylene receptor (Response_reg)
<i>GA2OX1</i>	bCL1313Contig1	1,454	AT1G78440	59.9	Gibberellin 2-oxidase (2OG-Fell_Oxy)
<i>HSP101</i>	bCL138-F11-M13R_c	726	AT1G74310	73.4	Heat shock protein 101 (AAA_2)
<i>HSP70</i>	bPSL_41-D06-M13R_c	819	AT5G09590	86.8	Heat shock protein 70/HSC70-5 (HSP70)
<i>IAA9</i>	bCL1341Contig1	1,672	AT5G65670	51.2	Auxin-responsive protein/indoleacetic acid- induced protein 9 (AUX_IAA)
<i>LSM1</i>	bHS_37-F11-M13R_c	766	AT3G14080	72.4	Small nuclear ribonucleoprotein/snRNP, putative/Sm protein, putative, similar to U6 snRNA-associated Sm-like protein
<i>LUT2</i>	bCL3563Contig1	1,082	AT5G57030	75.5	Lycopene epsilon cyclase (Lycopene_cycl)
<i>NCBP</i>	bCL183Contig1	1,081	AT5G18110	85.3	Novel cap-binding protein (IF4E)
<i>SVP</i>	bCL2852Contig1	826	AT2G22540	54.9	Short vegetative phase protein, MAD5 box transcription factor related cluster (SRF-TF)
<i>HIR</i>	bCL144Contig1	1,266	AT1G69840	89.7	Band 7 family protein, strong similarity to hypersensitive-induced response protein (Band_7)
<i>TCH4</i>	bCL1212Contig1	1,228	AT5G57560	65.9	Xyloglucan:xyloglucosyl transferase/xyloglucan endotransglycosylase/endo- xyloglucan transferase/ TCH4
<i>TIP4</i>	bA_23-B12-M13R_c	751	AT2G25810	50.5	Tonoplast intrinsic protein (MIP)
<i>TOM1</i>	bCL4416Contig1	780	AT4G21790	75.4	Transmembrane protein-related/TOM1 (DUF1084)
<i>TOM2A</i>	bCL3115Contig1	1,113	AT1G32400	60.1	Senescence-associated family protein
<i>TOM3</i>	bPSL_36-E03-M13R_c	857	AT1G14530		Tobamovirus multiplication protein 3/THH1 (DUF1084)
<i>UGE5</i>	bCL1153Contig1	1,313	AT4G10960	70.9	UDP-glucose 4-epimerase/UDP-galactose 4-epimerase/Galactowaldenase (Epimerase)
<i>WRKY70</i>	bA_25-B01-M13R_c	889	AT3G56400	42.1	WRKY family transcription factor, DNA-binding protein (WRKY)

6% of the melon unigenes containing di-, tri- or tetranucleotide repeats. A clear bias toward AG and AAG repeats existed, that account for 67% of the SSRs. In contrast, the GC repeat was not found in the melon dataset. A similar bias toward AG and against CG repeats has been identified in Arabidopsis and other plant species [40,78]. As proposed at least in one other instance [40], this may be due to the tendency of CpG sequences to be methylated [79], which potentially might inhibit transcription. Another interesting feature of melon SSRs relates to their pattern of localization with respect to putative initiation and termination codons. It is known that the UTRs of transcribed sequences are richer in SSRs than coding regions, particularly at the 5'-UTRs [36,40]. However, in the melon dataset, a high proportion of SSRs (29.5%) were found in ORFs. An analysis of the localization of di-, tri- and tetranucleotide repeats separately showed that di- and tetranucleotides were preferentially located in UTRs, whereas trinucleotides localised in both, UTRs and ORFs, consistently with maintenance of the ORFs coding capacity. Thus, the prevalence of trinucleotide repeats in the melon dataset (71%) explains this result.

We identified in the melon sequence dataset 356 high-quality SNPs. Since non-redundant sequences analysed here encompassed 4.5 Mb, one SNP was found every 12,000 pb of sequence. This small figure is probably due to the limited number of melon genotypes used and the low redundancy found among libraries. In fact, when the frequency of SNPs is computed in relation to the length and number of contigs containing SNPs, the corresponding value (one SNP in every 616 bp of sequence) is of the same order of magnitude as values previously calculated for melon (441 bp; [80]) and other plant species [40]. With the advent of high-throughput detection systems, the SSRs and SNPs identified here will constitute an important resource for mapping and marker-assisted breeding in melon and closely related crops.

As an approach to the function of melon unigenes, we carried out a bioinformatics analysis based on BLASTX and matches with the Pfam database [49]. The proportion of melon unigenes with no similar sequences in databases was quite high, suggesting that the melon dataset may encompass an important number of melon-specific

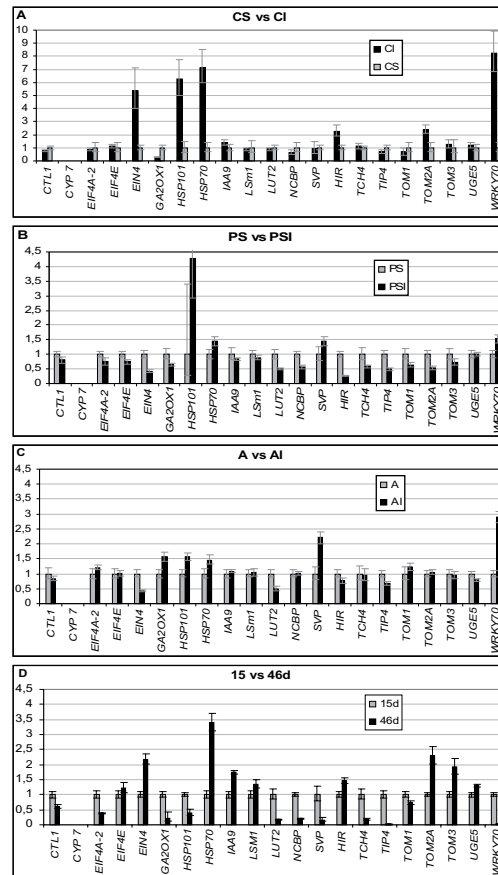


Figure 4
Transcripts analyzed by Real Time qPCR. (A) Pattern of transcripts accumulation in CMV-infected melon cotyledons (CI) relative to that of healthy cotyledons (CS). (B) Pattern of transcripts accumulation in *M. cannonballus* infected roots of *C. melo* L. cv. "Piel de sapo" (PSI) relative to that of healthy roots (PS). (C) Pattern of transcripts accumulation in *M. cannonballus* infected roots of *C. melo* L. ssp. *agrestis* (AI) relative to that of healthy roots (A). (D) Pattern of transcripts accumulation in fruits of 15 days after pollination of *C. melo* L. cv. "Piel de sapo" (15d) relative to that of fruits of 46 days after pollination (46d). cy: cyclophilin endogenous control; see Table 9 for the rest of genes.

sequences. However, the proportion specific sequences might be overestimated because blasting has been made with unigene sequences, which in many cases do not cover the complete length of the transcript. We performed a functional classification of the unigenes following the Gene Ontology scheme, which is one of the more versatile and complete systems for functional classification [48]. A comparison of the distributions of melon and Arabidopsis unigenes in GO categories showed that both followed similar tendencies, suggesting that the melon dataset is representative of the whole melon transcriptome. This is remarkable, as the number of different libraries sequenced has been relatively small; again, this is probably due to the success of the normalization process. We have also carried out specific searches for genes involved in pathways of particular relevance in melon, such as resistance response and fruit development, identifying a remarkable number of melon candidates. For example, an ortholog of the flagellin receptor FLS2 from Arabidopsis [62] has been identified, together with 163 candidate RLKs that may have critical roles in pathogen recognition or diverse signalling processes. Similarly, up to 8 MADS-box gene homologs with potential roles in development have been found in the melon dataset. Moreover, a bioinformatics approach [52,53] allowed the identification of potential precursors of melon miRNAs together with several potential targets in the melon dataset. This finding opens the door to biotechnology approaches based on the use of artificial miRNAs to specifically silence melon genes [81,82].

The transcript accumulation analysis for the 20 selected genes revealed important changes in gene expression associated with pathogen infection and fruit development. For virus infection, the accumulation of transcripts remained unaltered for 12 genes, but showed a significant increase for 6 genes and a decrease for 2. Among the set of genes analysed, *TOM2A* and *EIF4E* were known to code for virus susceptibility factors [27,83]; the expression of *TOM2A* was increased after CMV infection, consistently with its requirement by the virus, but this was not the case for *EIF4E*. Different hypotheses can explain this result: since *EIF4E* is an abundant, housekeeping protein, increased expression may not be essential for virus multiplication; alternatively, CMV may not use *EIF4E* in melon or may use a factor coded by a different member of the 4E family; it may also be that timing of the sampling for this experiment was not appropriate to detect such an effect, as requirement of *EIF4E* might occur very early during virus multiplication. In the cases of infection of susceptible ("Piel de sapo") and resistant (pat81) melons by *M. cannonballus*, more extensive alterations in gene expression seemed to occur in the susceptible than in the resistant accession. Significantly, for the susceptible accession, stress responsive genes (e.g. *HSP101*) appeared to be max-

imally induced, whereas for the resistant accession, a gene encoding a WRKY70 transcription factor, potentially involved in resistance response, was induced to high levels. Significantly, expression of *GA2OX1* increased about 1.5 times in pat81 after the *M. cannonballus* attack, whereas it decreased in "Piel de sapo". *GA2ox* is a major gibberellin (GA) catabolic enzyme, with an important role in controlling GA levels in plants. Hormones control many plant developmental processes, and strong evidence indicates that hormone signalling is involved in the regulation of root growth and architecture [84,85]. The differential response of the *GA2OX1* gene in the two melon genotypes is consistent with an enhanced root growth in pat81 after infection [86]. Notably, other genes involved in hormone-mediated signalling pathways, such as the *IAA9* gene, did not show such differential response to *M. cannonballus* infection in both genotypes. In the case of fruit development, differences in the expression of selected genes between immature and ripening fruits appeared to be even sharper than in the cases of healthy and pathogen-infected tissues. Specific roles during fruit development for *HSP70*, *TOM2A* and *TOM3* have not been identified, though an increased expression has been shown at least in the case of *TOM2A* in tomato [87]. The ethylene receptor gene *EIN4* showed a two-fold increase in expression. *EIN4* is the ortholog of Arabidopsis *EIN4* and tomato *LeETR4* [88,89]. In tomato, *LeETR4* is also highly expressed in ripening fruit, suggesting that it responds by modulating ethylene signalling during ripening [63]. The MADS-box gene (*SVP*) showed about a four-fold decrease in expression. This gene is the ortholog of tomato *JOINTLESS*, which specifies the abscission zone in tomato. In tomato fruit microarray hybridizations, the expression of *JOINTLESS* also decreased from 7 to 57 DAP [87], in agreement with our data for melon. The lycopene epsilon cyclase (*LUT2*) and xyloglucan endotransglycosylase (*TCH4*) genes showed an approximately four-fold decrease in expression during melon fruit development. These findings fit with the patterns of expression of these genes in tomato, where their transcript levels decrease to a non-detectable level in the ripe fruits [90,91].

Conclusion

In summary, this collection of ESTs represents a substantial increase on the information available for melon. The dataset contains SSR and SNP markers that can be used for breeding, as well as a significant number of candidate genes that can be experimentally tested for their roles in various important processes. This set of genes constitutes also the basis for a microarray for melon that is being used in experiments to further analyse fruit development and maturation and responses to pathogen infections.

Methods

Plant material

The cDNA libraries were prepared using material from four different melon genotypes: the line T-111 (Semillas Fitó, Barcelona, Spain), which corresponds to a Piel de Sapo breeding line, the Piel de Sapo cultivar "Piñonet torpedo" (Semillas Batlle, Barcelona, Spain), the accession C-35 of the germplasm collection of La Mayora-CSIC (EELM-CSIC, Málaga, Spain), which corresponds to a cantaloupe-type of melon, and the accession pat81 of *C. melo* L. ssp. *agrestis* (Naud.) Pangalo maintained at the germplasm bank of COMAV (COMAV-UIPV, Valencia, Spain) (Table 1). Seeds of line T-111 were germinated at 30°C for two days and plants were grown in a greenhouse in peat bags, drip irrigated, with 0.25-m spacing between plants. Fruits of 15 and 46 days after pollination were collected and mesocarp tissues were recovered and used for RNA extractions. Root samples were from Piel de Sapo and pat81 plants, both healthy and inoculated with *M. cannonballus*. Piel de Sapo is fully susceptible to the infection by this fungus whereas pat81 has been shown to be partially resistant [92,93]. Seeds were pre-germinated in Petri dishes. After 4 days, seedlings were transplanted to 0.5-l pots filled with sterile soil substrate and grown in a greenhouse (20–35°C, 60–85% relative humidity). Inoculations were carried out by adding 50 colony-forming units (CFU) of *M. cannonballus* per gram of sterile soil as described by Iglesias et al. [94]. Fourteen days after inoculation, healthy and inoculated roots were collected for RNA extraction. The presence of the fungus and the infection levels were assessed by real-time quantitative PCR as described by Picó et al. [95]. CMV infected cotyledons were collected from plants of the C-35 accession. In this case, seeds were pregerminated in Petri dishes for 24 h at 28°C in the dark, planted in 0.5-l pots and maintained in an insect-proof green house (20–28°C, 45 to 85% relative humidity) for 6 to 7 days, until the first true leaf started emerging. At this stage, cotyledons were mechanically inoculated with CMV following standard procedures [96]. Inoculated cotyledons were harvested 4 days after inoculation and used for RNA extractions. Dot-blot hybridisation [97] was used to check infection by CMV. Plants of the C-35 accession were also used for collecting healthy leaves. Plants were maintained in the greenhouse for 21 days, and second and third leaves above cotyledons were harvested for RNA extractions.

Construction of cDNA libraries and EST sequencing

Total RNA was prepared as described by Aranda et al. [98]. Poly(A⁺) RNA from total RNA was purified using Micro-Poly(A⁺) Purist (Ambion, Austin, TX, USA), a cellulose-oligo(dT)-based method. Integrity and quality of both total and poly(A⁺) RNA were tested by gel electrophoresis. cDNA libraries were constructed with the SMART cDNA Library Construction kit (Clontech, Mountain View, CA,

USA), using a modified primer to include a *Sfi* I enzyme restriction site. A normalization step was carried out with TRIMMER kit (Evrogen, Moscow, Russia). After normalization, a cDNA fractionation step was performed with SizeSep 400 Spun Columns (Amersham Biosciences, Buckinghamshire, England). cDNA was digested with *Sfi* I, generating *Sfi* IA-*Sfi* IB cohesive ends for directional cloning into a modified version of BlueScript SK plasmid vector (Stratagene, La Jolla, CA, USA). Ligation products were transformed into *E. coli* electrocompetent cells DH10B (Invitrogen, Carlsbad, CA, USA) by electroporation. The titer of the libraries was evaluated by plating an aliquot on LB agar plates with ampicillin at 100 µg ml⁻¹. Only libraries of 10⁵ cfu ml⁻¹ or more were considered as acceptable. Prior to large scale sequencing, the average insert size was estimated by restriction analyses of 24 plasmid DNA minipreps per library from randomly picked colonies.

Sequencing was carried out from the 5'-end of the inserts without library amplification using the universal M13 reverse primer. An external custom service was contracted for this task (Macrogen Inc., Seoul, Korea). Approximately 6,000 clones were sequenced from the CI library, and 3,500 clones were sequenced from each of the other libraries (Table 2). Sequences obtained in this work can be found in GenBank [accession numbers [AM713476](#) to [AM743079](#)] and MELOGEN [44].

Bioinformatics

EST sequences were automatically trimmed, clustered and annotated using the EST2uni analysis pipeline [43]. EST2uni comprises the analysis pipeline written in PERL [99], a database (MySQL) [100] and a web site to browse the results coded in PHP [101]. Thus, for the EST pre-processing step, base calling was performed with Phred [102], low quality regions and vector sequences were trimmed with Lucy [103], and repeats and low complexity regions were masked with RepeatMasker [104] and Seqclean [105]. Further vector contamination was also eliminated with Seqclean using NCBI's UniVec [106]. High-quality EST sequences were then assembled to obtain the unigenes set using Tgicl [105].

Detection of SSRs was performed using Sputnik [107]. Putative SNPs were annotated when the least represented allele was present in two EST sequences or more. ORFs were predicted in the ESTs with the aid of the ESTScan software [108].

For functional annotation, comparisons against the Arabidopsis (TAIR) [109] and Uniref [110] databases were carried out using BLASTN or BLASTX for nucleotide or protein sequences, respectively. Functional domains were searched with HMMPFAM [111] using the Pfam database

[112]. The Gene Ontology (GO) classification [48] was derived from the BLASTN results against the Arabidopsis proteome. Also, a bi-directional BLASTN comparison was performed in order to obtain a set of putative orthologs with Arabidopsis. Finally, a set of superunigenes was obtained grouping different unigenes with the same expected mRNA target, as judged by extensive sequence overlapping.

To assess codon usage, we generated a set of melon sequences predicted to contain full-length coding regions. These sequences were subjected to BLASTX and, after manual inspection, sequences showing a high similarity to Arabidopsis proteins were selected to ensure that no sequences containing frame-shift errors were included in the analysis. From this smaller dataset, which included 588 sequences, ORFs were defined and a codon usage table was created. Codon usage was calculated from sequences using the GCUA program [113]. All codons were found in the dataset, with the least frequent codon represented 134 times.

To identify potential melon miRNAs, the 33,292 melon ESTs were subjected to a BLAST search against mature sequences of known miRNAs from the miRNA Registry Database (released January 2007) [54] using BLASTN [47]. ESTs with only 0-1 mismatched nucleotides with known miRNAs were considered. Selected ESTs were subjected to a BLAST search against protein databases in order to remove potential protein-coding sequences. ESTs pertaining to the same melon unigenes of the MELOGEN

database were grouped. The secondary structures of the unigenes encoding potential miRNA precursors were predicted with the web-based tool MFOLD [55], using default parameters. In each case, only the lower energy structure was selected for visual inspection, as previously described [52,114]. In order to select unigenes with perfect or near-perfect secondary foldback hairpin structures, only sequences with a maximum size of 3 nucleotides for a bulge in the miRNA sequence and with at least 16 paired nucleotides between the mature sequence and the opposite arm were considered as potential miRNA candidates. In addition, the minimal folding free energy index (MFEI) for each sequence was calculated following Zhang et al. [53].

Gene expression analyses

Real time quantitative PCR was performed with an AB 7500 System (Applied Biosystems, Foster City, CA, U.S.A) to quantify mRNA corresponding to some transcripts of interest, in the tissues and physiological conditions used for library construction. Twenty ESTs representing these transcripts were chosen from the database and used to generate gene-specific primers (Table 10) with Primer Express Software (Applied Biosystems). The chemistry used for PCR product detection was the Power SYBR green dye (Applied Biosystems) and ROX as passive reference. *CYCLOPHILIN* served as endogenous control (sequence extracted from the database). $\Delta\Delta C_t$ was the method of calculation to perform relative quantification, and three technical replicates were carried out and considered for statistical analysis. Melting curves analyses at the end of

Table 10: Primer sequences for Real Time-qPCR analysis of transcript accumulation

Gene	Primer sequence (5'→3') Forward	Reverse
<i>CTL1</i>	tggccatgttggctctaa	ctcccgtgacaactccatca
<i>CYP 7</i>	cgatgtggaatgacggaa	cggtgcataatgctcggaa
<i>EIF4A-2</i>	ttcccagagtttcaaatgca	caaatgcttctggggcact
<i>EIF4E</i>	ttcggcttcttccctccat	cgccgatgtagtttcac
<i>EIN4</i>	tgcaacgtgactgtttct	ctggcatgtgaatccaaga
<i>GA2OX1</i>	tagggcaaatcggttagcga	caaatgcaaacgattgaa
<i>HSP101</i>	aacgatggtcggattgaca	tcaccctcatggtatccaact
<i>HSP70</i>	gctgaggcgtaccctggaaa	atctgcagcactcctttgt
<i>IAA9</i>	gacgaaagccaggttcaag	ccccctcatactcatttcaaa
<i>LSM1</i>	ctacttcgagatgggcggaa	tcaccaacgatcaccctttca
<i>LUT2</i>	gctggcgtggaacactcttt	ggaatgcttcaatgctcagt
<i>NCBP</i>	cgctggctgtcttaattgca	gcctacgaactcattgaca
<i>SVP</i>	cgaggcaggtcacgttctct	agagaagacgagagcgcaa
<i>HIR</i>	tgacgggctcagagacagt	gtccaggagcgttttcagc
<i>TCH4</i>	ggaggstaggccttgaggaaat	ctggacattgctcgaacaa
<i>TIP4</i>	tccttctgggtcggatc	cgttgccaaaacgattg
<i>TOM1</i>	ggggagaaggaaacttcatgag	gctgcaaaagcggataaac
<i>TOM2A</i>	ctcctcagcagcgaagaaa	tgaccctgctaaacaccac
<i>TOM3</i>	aatggatctgggctgtgt	aaggcaaggcttggcatgt
<i>UGES</i>	gcgaaagtgtccaaaagcca	acacaagctttttgcatcgt
<i>WRKY70</i>	ggattgctctggcctgac	tcggctcttttttcgac

the process and No Template Controls (NTC) were carried out to ensure product-specific amplification and no primer-dimer quantification. A control reaction as for reverse transcription but without the enzyme was performed to evaluate genomic DNA contamination.

Authors' contributions

Daniel Gonzalez-Ibeas prepared RNAs for two libraries, constructed the eight libraries, carried out the gene expression analysis by Real-Time-qPCR and participated in the bioinformatics analyses and in the drafting of the manuscript. José Blanca carried out the bioinformatics analyses, EST database and web page, and participated in the drafting of the manuscript. Cristina Roig and Belén Picó prepared RNAs for the root libraries and participated in the drafting of the manuscript. Mireia González-To, Wim Deleu and Jordi Garcia-Mas prepared RNAs from melon fruits and participated in the drafting of the manuscript. Pere Puigdomènech is the main coordinator of The MEL-OGEN Project and participated in the conception of the study together with Pere Arús, Fernando Nuez, Jordi Garcia-Mas and Miguel A. Aranda. Miguel A. Aranda is the principal investigator of this work, supervised it and wrote the manuscript. All authors read and approved the final manuscript.

Additional material

Additional file 1

Genes potentially encoding pathogen resistance and fruit quality traits. Genes were identified in the melon data set by comparison with the Arabidopsis database [6, 109]. A brief description, the corresponding Arabidopsis locus and the HMMR domain identified are given for each unigene.

Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-8-306-S1.pdf]

Acknowledgements

This work was supported by grants from Ministerio de Educación y Ciencia (Spain) (GEN2003-20237-C06) and Consejería de Educación y Cultura (Región de Murcia, Spain) (BIO2005/04-6436). Wim Deleu, Cristina Roig and Daniel Gonzalez-Ibeas are recipient of a postdoctoral fellowship from the Centre de Recerca en Agrigenòmica CSIC-IRTA (Spain), a Juan de la Cierva grant from Ministerio de Educación y Ciencia (Spain) and a predoctoral fellowship from Ministerio de Educación y Ciencia (Spain), respectively.

References

1. FAOSTAT Agriculture data [http://faostat.fao.org/default.aspx]
2. Jayasooriya AP, Sakono M, Yukizaki C, Kawano M, Yamamoto K, Fukuda N: Effects of *Momordica charantia* powder on serum glucose levels and various lipid parameters in rats fed with cholesterol-free and cholesterol-enriched diets. *J Ethnopharmacol* 2000, **72**:331-336.
3. Arumuganathan K, Earle ED: Nuclear DNA content of some important plant species. *Plant Mol Biol Rep* 1991, **9**:208-218.
4. Goff SA, Ricke D, Lan TH, Presting G, Wang R, Dunn M, Glazebrook J, Sessions A, Oeller P, Varma H, others: A Draft Sequence of the Rice Genome (*Oryza sativa* L. ssp. japonica). *Science* 2002, **296**:92-100.
5. Yu J, Hu S, Wang J, Wong GK-S, Li S, Liu B, Deng Y, Dai L, Zhou Y, Zhang X, others: A Draft Sequence of the Rice Genome (*Oryza sativa* L. ssp. indica). *Science* 2002, **296**:79-92.
6. Arabidopsis Genome Initiative: The Arabidopsis Information Resource (TAIR): a comprehensive database and web-based information retrieval, analysis, and visualization system for a model plant. *Nucleic Acids Res* 2001, **29**:102-105.
7. Kirkbride JH: *Biosystematic monograph of the genus Cucumis (Cucurbitaceae)* Boone, North Carolina: Parkway Publishers; 1993.
8. Garcia-Mas J, Monforte AJ, Arus P: Phylogenetic relationships among *Cucumis* species based on the ribosomal internal transcribed spacer sequence and microsatellite markers. *Plant Syst Evol* 2004, **248**:191-203.
9. Liu L, Kakihara F, Kato M: Characterization of six varieties of *Cucumis melo* L. based on morphological and physiological characters, including shelf-life of fruit. *Euphytica* 2004, **135**:305-313.
10. Miccolis V, Saltveit ME: Morphological and physiological changes during fruit growth and maturation of seven melon cultivars. *J Am Soc Hort Sci* 1991, **116**:1025-1029.
11. Shalit M, Katzir N, Tadmor Y, Larkov O, Burger Y, Shalekhet F, Las-tochkin E, Ravid U, Amar O, Edelstein M, others: Acetyl-CoA: Alcohol acetyltransferase activity and aroma formation in ripening melon fruits. *J Agric Food Chem* 2001, **49**:794-799.
12. Baudouin AD, Mottram DS, Dodson AT, John P: Effect of Aminocyclopropane-1-carboxylic Acid Oxidase Antisense Gene on the Formation of Volatile Esters in Cantaloupe Charentais Melon (Cv. Védrandais). *J Agric Food Chem* 1998, **46**:4787-4792.
13. Flores F, El Yahyaoui F, de Billebeck G, Romojaro F, Latche A, Bouzayen M, Pech JC, Ambid C: Role of ethylene in the biosynthetic pathway of aliphatic ester aroma volatiles in Charentais Cantaloupe melons. *J Exp Bot* 2002, **53**:201-206.
14. Yahyaoui FEL, Wongs-Aree C, Latche A, Hackett R, Grierson D, Pech JC: Molecular and biochemical characteristics of a gene encoding an alcohol acyl-transferase involved in the generation of aroma volatile esters during melon ripening. *FEBS J* 2002, **269**:2359-2366.
15. Hadfield KA, Bennett AB: Polygalacturonases: Many Genes in Search of a Function. *Plant Physiol* 1998, **117**:337-343.
16. Rose JK, Hadfield KA, Labavitch JM, Bennett AB: Temporal sequence of cell wall disassembly in rapidly ripening melon fruit. *Plant Physiol* 1998, **117**:345-361.
17. Bennett AB: Biochemical and genetic determinants of cell wall disassembly in ripening fruit: A general model. *Hortscience* 2002, **37**:447-450.
18. Haritatos E, Keller F, Turgeon R: Raffinose oligosaccharide concentrations measured in individual cell and tissue types in *Cucumis melo* L. leaves: Implications for phloem loading. *Planta* 1996, **198**:614-622.
19. Volk GM, Turgeon R, Beebe DU: Secondary plasmodesmata formation in the minor-vein phloem of *Cucumis melo* L. and *Cucurbita pepo* L. *Planta* 1996, **199**:425-432.
20. Gomez G, Torres H, Pallas V: Identification of translocatable RNA-binding phloem proteins from melon, potential components of the long-distance RNA transport system. *Plant J* 2005, **41**:1107-1116.
21. Chen JQ, Rahbé Y, Delobel B, Sauvion N, Guillaud J, Febvay G: resistance to the aphid *Aphis gossypii*: behavioural analysis and chemical correlations with nitrogenous compounds. *Entomol Exp Appl* 1997, **78**:5:33-44.
22. Luo MZ, Wang YH, Frisch D, Joobeur T, Wing RA, Dean RA: Melon bacterial artificial chromosome (BAC) library construction using improved methods and identification of clones linked to the locus conferring resistance to melon Fusarium wilt (Fom-2). *Genome* 2001, **44**:154-162.
23. Klingler J, Powell G, Thompson GA, Isaacs R: Phloem specific aphid resistance in *Cucumis melo* line AR 5: effects on feeding behaviour and performance of *Aphis gossypii*. *Entomol Exp Appl* 1998, **86**:79-88.
24. Marco CF, Aguilar JM, Abad J, Gomez-Guillamon ML, Aranda MA: Melon resistance to Cucurbit yellow stunting disorder virus

- is characterized by reduced virus accumulation. *Phytopathology* 2003, **93**:844-852.
25. Diaz JA, Nieto C, Moriones E, Truniger V, Aranda MA: **Molecular characterization of a Melon necrotic spot virus strain that overcomes the resistance in melon and nonhost plants.** *Mol Plant-Microbe Interact* 2004, **17**:668-675.
 26. Klingler J, Greasy R, Gao LL, Nair RM, Calix AS, Jacob HS, Edwards OR, Singh KB: **Aphid resistance in *Medicago truncatula* involves antixenosis and phloem-specific, inducible antibiosis, and maps to a single locus flanked by NBS-LRR resistance gene analogs.** *Plant Physiol* 2005, **137**:1445-1455.
 27. Nieto C, Morales M, Orjeda G, Clepet C, Monfort A, Sturbois B, Puigdemonech P, Pitrat M, Caboche M, Dogimont C, others: **An eIF4E allele confers resistance to an uncapped and non-polyadenylated RNA virus in melon.** *Plant J* 2006, **48**:452-462.
 28. Perin C, Hagen L, Conto VD, Katzir N, Danin-Poleg Y, Portnoy V, Baudracco-Arnas S, Chadouef J, Dogimont C, Pitrat M: **A reference map of *Cucumis melo* based on two recombinant inbred line populations.** *Theor Appl Genet* 2002, **104**:1017-1034.
 29. Gonzalo MJ, Oliver M, Garcia-Mas J, Monfort A, Dolcet-Sanjuan R, Katzir N, Arus P, Monforte A: **Simple-sequence repeat markers used in merging linkage maps of melon (*Cucumis melo* L.).** *Theor Appl Genet* 2005, **110**:802-811.
 30. Eduardo I, Arus P, Monforte AJ: **Development of a genomic library of near isogenic lines (NILs) in melon (*Cucumis melo* L.) from the exotic accession P1161375.** *Theor Appl Genet* 2005, **112**:139-148.
 31. Ayub R, Guis M, BenAmor M, Gillot L, Roustan JP, Latche A, Bouzayen M, Pech JC: **Expression of ACC oxidase antisense gene inhibits ripening of cantaloupe melon fruits.** *Nat Biotechnol* 1996, **14**:862-866.
 32. Guis M, Roustan JP, Dogimont C, Pitrat M, Pech JC: **Melon biotechnology.** *Biotechnol Genet Engng Rev* 1998, **15**:289-311.
 33. Gaba V, Zelcer A, Gal-On A: **Cucurbit biotechnology – The importance of virus resistance.** *In Vitro Cell Dev Biol Plant* 2004, **40**:346-358.
 34. Rudd S: **Expressed sequence tags: alternative or complement to whole genome sequences?** *Trends Plant Sci* 2003, **8**:321-329.
 35. Alba R, Fei Z, Payton P, Liu Y, Moore S, Debbie P, Cahn J, D'Ascenzo M, Gordon JS, Rose JKC, others: **ESTs, cDNA microarrays, and gene expression profiling: tools for dissecting plant physiology and development.** *Plant J* 2004, **39**:697-714.
 36. Morgante M, Hanafey M, Powell W: **Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes.** *Nat Genet* 2002, **30**:194-200.
 37. Rafalski JA: **Novel genetic mapping tools in plants: SNPs and LD-based approaches.** *Plant Sci* 2002, **162**:329-333.
 38. Qiyang S, Zhu W, Hamilton J, Lin H, Campbell M, Childs K, Thibaud-Nissen F, Malek RL, Lee Y, Zheng L, others: **The TIGR Rice Genome Annotation Resource: improvements and new features.** *Nucleic Acids Res* 2007, **35**:D883-D887.
 39. Fei ZJ, Tang XM, Alba R, Giovannoni J: **Tomato Expression Database (TED): a suite of data presentation and analysis tools.** *Nucleic Acids Res* 2006, **34**:D766-D770.
 40. Newcomb RD, Crowhurst RN, Gleave AP, Rikkerink EHA, Allan AC, Beuning LL, Bowen JH, Gera E, Jamieson KR, Janssen BJ, others: **Analyses of expressed sequence tags from apple.** *Plant Physiol* 2006, **141**:147-166.
 41. Goes da Silva F, Iandolino A, Al Kayal F, Bohlmann MC, Cushman MA, Lima R, Ergul A, Figueroa R, Kabiloglu EK, Osborne C, others: **Characterizing the Grape Transcriptome. Analysis of Expressed Sequence Tags from Multiple Vitis Species and Development of a Compendium of Gene Expression during Berry Development.** *Plant Physiol* 2005, **139**:574-597.
 42. Forment J, Gadea J, Huerta L, Abizanda L, Agusti J, Alamar S, Alos E, Andres F, Arribas R, Beltran JP, others: **Development of a citrus genome-wide EST collection and cDNA microarray as resources for genomic studies.** *Plant Mol Biol* 2005, **57**:375-391.
 43. ESTzuni [http://www.melogen.upv.es/genomical/web_estpipe/index.php]
 44. **MELOGEN database** [<http://www.melogen.upv.es>]
 45. Sterky F, Bhalerao RR, Unneberg P, Segerman B, Nilsson P, Brunner AM, Charbonnel-Campaa L, Lindvall JJ, Tandri K, Strauss SH, others: **A Populus EST resource for plant functional genomics.** *Proc Natl Acad Sci USA* 2004, **101**:13951-13956.
 46. Hsiang T, Goodwin PH: **Distinguishing plant and fungal sequences in ESTs from infected plant tissues.** *J Microbiol Methods* 2003, **54**:339-351.
 47. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ: **Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.** *Nucleic Acids Res* 1997, **25**:3389-3402.
 48. The Gene Ontology Consortium: **Gene Ontology: tool for the unification of biology.** *Nat Genet* 2000, **25**:25-29.
 49. Finn RD, Mistry J, Schuster-Bockler B, Griffiths-Jones S, Hollich V, Lassmann T, Moxon S, Marshall M, Khanna A, Durbin R, others: **Pfam: clans, web tools and services.** *Nucleic Acids Res* 2006, **34**:D247-D251.
 50. Carrington JC, Ambros V: **Role of microRNAs in plant and animal development.** *Science* 2003, **301**:336-338.
 51. Bartel DP: **MicroRNAs: Genomics, biogenesis, mechanism, and function.** *Cell* 2004, **116**:281-297.
 52. Zhang BH, Pan XP, Cannon CH, Cobb GP, Anderson TA: **Conservation and divergence of plant microRNA genes.** *Plant J* 2006, **46**:243-259.
 53. Zhang BH, Pan XP, Cox SB, Cobb GP, Anderson TA: **Evidence that miRNAs are different from other RNAs.** *Cell Mol Life Sci* 2006, **63**:246-254.
 54. Griffiths-Jones S: **The microRNA Registry.** *Nucleic Acids Res* 2004, **32**:D109-D111.
 55. Zuker M: **Mfold web server for nucleic acid folding and hybridization prediction.** *Nucleic Acids Res* 2003, **31**:3406-3415.
 56. Palamou JF, Allen E, Wu XL, Schommer C, Schwab R, Carrington JC, Weigel D: **Control of leaf morphogenesis by microRNAs.** *Nature* 2003, **425**:257-263.
 57. Gustafson AM, Allen E, Givan S, Smith D, Carrington JC, Kasschau KD: **ASRP: The Arabidopsis Small RNA Project Database.** *Nucleic Acids Res* 2005, **33**:D637-D640.
 58. Chisholm ST, Coaker G, Day B, Staskawicz BJ: **Host-microbe interactions: Shaping the evolution of the plant immune response.** *Cell* 2006, **124**:803-814.
 59. Kushner DB, Lindenbach BD, Grdzelskivili VZ, Noueiry AO, Paul SM, Ahlquist P: **Systematic, genome-wide identification of host genes affecting replication of a positive-strand RNA virus.** *Proc Natl Acad Sci USA* 2003, **100**:15764-15769.
 60. Diaz-Pendon JA, Truniger V, Nieto C, Garcia-Mas J, Bendahmane A, Aranda MA: **Advances in understanding recessive resistance to plant viruses.** *Mol Plant Pathol* 2004, **5**:223-233.
 61. Robaglia C, Caranta C: **Translation initiation factors: a weak link in plant RNA virus infection.** *Trends Plant Sci* 2006, **11**:40-45.
 62. Gomez-Gomez L, Boller T: **FLS2: An LRR receptor-like kinase involved in the perception of the bacterial elicitor flagellin in Arabidopsis.** *Mol Cell* 2000, **5**:1003-1011.
 63. Giovannoni JJ: **Genetic regulation of fruit development and ripening.** *Plant Cell* 2004, **16**:S170-S180.
 64. Tanksley SD: **The genetic, developmental, and molecular bases of fruit size and shape variation in tomato.** *Plant Cell* 2004, **16**:S181-S189.
 65. Pinyopich A, Ditta GS, Savidge B, Liljegren SJ, Baumann E, Wisman E, Yanofsky MF: **Assessing the redundancy of MADS-box genes during carpel and ovule development.** *Nature* 2003, **424**:85-88.
 66. Dinneny JR, Weigel D, Yanofsky MF: **A genetic framework for fruit patterning in Arabidopsis thaliana.** *Development* 2005, **132**:4687-4696.
 67. Pitrat M: **2002 gene list for melon.** *Cucurbit Genet Coop Rep* 2002, **25**:76-93.
 68. Monforte AJ, Oliver M, Gonzalo MJ, Alvarez JM, Dolcet-Sanjuan R, Arus P: **Identification of quantitative trait loci involved in fruit quality traits in melon (*Cucumis melo* L.).** *Theor Appl Genet* 2004, **108**:750-758.
 69. Frary A, Nesbitt TC, Frary A, Grandillo S, van der Knaap E, Cong B, Liu JP, Meller J, Elber R, Alpert KB, others: **fw2.2: A quantitative trait locus key to the evolution of tomato fruit size.** *Science* 2000, **289**:85-88.
 70. Liu JP, Van Eck J, Cong B, Tanksley SD: **A new class of regulatory genes underlying the cause of pear-shaped tomato fruit.** *Proc Natl Acad Sci USA* 2002, **99**:13302-13306.
 71. Manning K, Tor M, Poole M, Hong Y, Thompson AJ, King GJ, Giovannoni JJ, Seymour GB: **A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening.** *Nat Genet* 2006, **38**:948-952.

72. Aranda MA, Escaler M, Wang D, Maule AJ: **Induction of HSP70 and polyubiquitin expression associated with plant virus replication.** *Proc Natl Acad Sci USA* 1996, **93**:15289-15293.
73. Escaler M, Aranda MA, Roberts IM, Thomas CL, Maule AJ: **A comparison between virus replication and abiotic stress (heat) as modifiers of host gene expression in pea.** *Mol Plant Pathol* 2000, **1**:159-167.
74. Whitlam SA, Yang CL, Goodin MM: **Global impact: Elucidating plant responses to viral infection.** *Mol Plant-Microbe Interact* 2006, **19**:1207-1215.
75. Aranda M, Maule A: **Virus-induced host gene shutoff in animals and plants.** *Virology* 1998, **243**:261-267.
76. Bonaldo MDF, Lennon G, Soares MB: **Normalization and subtraction: Two approaches to facilitate gene discovery.** *Genome Res* 1996, **6**:791-806.
77. Zhulidov PA, Bogdanova EA, Shcheglov AS, Vagner LL, Khaspekov GL, Kozhemyako VB, Mize MV, Meleshkevitch E, Moroz LL, Lukyanov SA, others: **Simple cDNA normalization using kamchatka crab duplex-specific nuclease.** *Nucleic Acids Res* 2004, **32**:e37.
78. Zhang LD, Yuan DJ, Yu SW, Li ZG, Cao YF, Miao ZQ, Qian HM, Tang KK: **Preference of simple sequence repeats in coding and non-coding regions of Arabidopsis thaliana.** *Bioinformatics* 2004, **20**:1081-1086.
79. Finnegan EJ, Genger RK, Peacock VJ, Dennis ES: **DNA methylation in plants.** *Annu Rev Plant Physiol Plant Mol Biol* 1998, **49**:223-247.
80. Morales M, Roig E, Monforte AJ, Arus P, Garcia-Mas J: **Single-nucleotide polymorphisms detected in expressed sequence tags of melon (Cucumis melo L.).** *Genome* 2004, **47**:352-360.
81. Niu QW, Lin SS, Reyes JL, Chen KC, Wu HW, Yeh SD, Chua NH: **Expression of artificial microRNAs in transgenic Arabidopsis thaliana confers virus resistance.** *Nat Biotechnol* 2006, **24**:1420-1428.
82. Schwab R, Ossowski S, Riester M, Warthmann N, Weigel D: **Highly specific gene silencing by artificial microRNAs in Arabidopsis.** *Plant Cell* 2006, **18**:1121-1133.
83. Yamanaka T, Imai T, Satoh R, Kawashima A, Takahashi M, Tomita K, Kubota K, Meshi T, Naito S, Ishikawa M: **Complete inhibition of tobamovirus multiplication by simultaneous mutations in two homologous host genes.** *J Virol* 2002, **76**:2491-2497.
84. Malamy JE: **Intrinsic and environmental response pathways that regulate root system architecture.** *Plant Cell Environ* 2005, **28**:67-77.
85. Radi A, Dina P, Guy A: **Expression of sarcotoxin IA gene via a root-specific tob promoter enhanced host resistance against parasitic weeds in tomato plants.** *Plant Cell Rep* 2006, **25**:297-303.
86. Dias RDS, Pico B, Espinos A, Nuez F: **Resistance to melon vine decline derived from Cucumis melo ssp agrestis: genetic analysis of root structure and root response.** *Plant Breeding* 2004, **123**:66-72.
87. **Tomato expression database** [<http://ted.bti.cornell.edu>]
88. Sakai H, Hua J, Chen QHG, Chang CR, Medrano LJ, Bleecker AB, Meyerowitz EM: **ETR2 is an ETR1-like gene involved in ethylene signaling in Arabidopsis.** *Proc Natl Acad Sci USA* 1998, **95**:5812-5817.
89. Tieman DM, Klee HJ: **Differential expression of two novel members of the tomato ethylene-receptor family.** *Plant Physiol* 1999, **120**:165-172.
90. Ronen G, Cohen M, Zamir D, Hirschberg J: **Regulation of carotenoid biosynthesis during tomato fruit development: Expression of the gene for lycopene epsilon-cyclase is down-regulated during ripening and is elevated in the mutant Delta.** *Plant J* 1999, **17**:341-351.
91. Catala C, Rose JKC, Bennett AB: **Auxin-regulated genes encoding cell wall-modifying proteins are expressed during early tomato fruit growth.** *Plant Physiol* 2000, **122**:527-534.
92. Esteve J, Nuez F: **Field resistance to melon dieback in Cucumis melo L.** *Cucurbit Genet. Coop* 1994, **17**:76-77.
93. Iglesias A, Pico B, Nuez F: **A temporal genetic analysis of disease resistance genes: resistance to melon vine decline derived from Cucumis melo var. agrestis.** *Plant Breeding* 2000, **119**:329-334.
94. Iglesias A, Pico B, Nuez F: **Artificial inoculation methods and selection criteria for breeding melons against vine decline.** *Acta Hortic* 2000, **510**:155-162.
95. Pico B, Roig C, Fita A, Nuez F: **Detección de Monosporascus cannonballus en raíces de melón mediante PCR cuantitativa en tiempo real.** *Acta Hort* 2005, **7**:169-176.
96. Hull S: *Matthews's plant Virology* San Diego: Academic Press; 2002.
97. Sambrook J, Russell DW: *Molecular cloning A laboratory manual.* New York: CSHL PRESS; 2001.
98. Aranda MA, Escaler M, Thomas CL, Maule AJ: **A heat shock transcription factor in pea is differentially controlled by heat and virus replication.** *Plant J* 1999, **20**:153-161.
99. **PERL** [<http://www.perl.org>]
100. **MySQL** [<http://www.mysql.com>]
101. **PHP** [<http://www.php.net>]
102. Ewing B, Hillier L, Wendl MC, Green P: **Base-calling of automated sequencer traces using Phred. I. Accuracy assessment.** *Genome Res* 1998, **8**:175-185.
103. Chou HH, Holmes MH: **DNA sequence quality trimming and vector removal.** *Bioinformatics* 2001, **17**:1093-1104.
104. **RepeatMasker** [<http://www.repeatmasker.org>]
105. **DFCI Gene Indices Software Tools** [<http://compbio.dfci.harvard.edu/cgi/software/>]
106. **NCBI's UniVec** [<http://www.ncbi.nlm.nih.gov/VecScreen/UniVec.html>]
107. **Sputnik** [<http://capb.dbi.udel.edu/main/ssr-proj/ssr-file/description.html>]
108. **ESTScan software** [<http://www.isrec.isb.sib.ch/ftp-server/ESTScan/>]
109. **TAIR: the Arabidopsis information resource** [<http://www.arabidopsis.org/>]
110. **Uniref** [<http://www.ebi.ac.uk/uniref/>]
111. Eddy SR: **Profile hidden Markov models.** *Bioinformatics* 1998, **14**:755-763.
112. **Pfam database** [<http://www.sanger.ac.uk/Software/Pfam/>]
113. McInerney JO: **GCUA (General Codon Usage Analysis).** *Bioinformatics* 1998, **14**:372-373.
114. Reinhardt BJ, Weinstein EG, Rhoades MW, Bartel B, Bartel DP: **MicroRNAs in plants.** *Genes Dev* 2002, **16**:1616-1626.
115. Apweiler R, Bairoch A, Wu CH, Barker WVC, Boeckmann B, Ferro S, Gasteiger E, Huang H, Lopez R, Magrane M, others: **UniProt: the Universal Protein knowledgebase.** *Nucleic Acids Res* 2004, **32**:D115-D119.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp



Research article

Open Access

A set of EST-SNPs for map saturation and cultivar identification in melon

Wim Deleu¹, Cristina Esteras^{1,2}, Cristina Roig², Mireia González-To¹, Iria Fernández-Silva¹, Daniel Gonzalez-Ibeas³, José Blanca², Miguel A Aranda³, Pere Arús¹, Fernando Nuez², Antonio J Monforte^{1,4}, Maria Belén Picó² and Jordi Garcia-Mas*¹

Address: ¹IRTA, Centre de Recerca en Agrigenòmica CSIC-IRTA-LIAB, Carretera de Cabrils Km 2, 08348 Cabrils (Barcelona), Spain, ²COMAV-UPV, Institute for the Conservation and Breeding of Agricultural Biodiversity, Universidad Politécnica de Valencia, Camino de Vera s/n, 46022 Valencia, Spain, ³Departamento de Biología del Estrés y Patología Vegetal, Centro de Edafología y Biología Aplicada del Segura (CEBAS)- CSIC, Apdo. correos 164, 30100 Espinardo (Murcia), Spain and ⁴Instituto de Biología Molecular y Celular de Plantas (IBMCP) UPV-CSIC, Ciudad Politécnica de la Innovación Edificio 8E, Ingeniero Fausto Elio s/n, 46022 Valencia, Spain

Email: Wim Deleu - wd@ramiroarnedo.com; Cristina Esteras - crieras@upvnet.upv.es; Cristina Roig - croig@btc.upv.es; Mireia González-To - mireia.gonzalez@irta.cat; Iria Fernández-Silva - iriafernandezsilva@gmail.com; Daniel Gonzalez-Ibeas - agr030@cebas.csic.es; José Blanca - jblanca@btc.upv.es; Miguel A Aranda - m.aranda@cebas.csic.es; Pere Arús - pere.arus@irta.cat; Fernando Nuez - fnuez@btc.upv.es; Antonio J Monforte - amonforte@ibmcp.upv.es; Maria Belén Picó - mpicosi@btc.upv.es; Jordi Garcia-Mas* - jordi.garcia@irta.cat

* Corresponding author ¹Equal contributors

Published: 15 July 2009

Received: 31 March 2009

BMC Plant Biology 2009, 9:90 doi:10.1186/1471-2229-9-90

Accepted: 15 July 2009

This article is available from: <http://www.biomedcentral.com/1471-2229/9/90>

© 2009 Deleu et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: There are few genomic tools available in melon (*Cucumis melo* L.), a member of the *Cucurbitaceae*, despite its importance as a crop. Among these tools, genetic maps have been constructed mainly using marker types such as simple sequence repeats (SSR), restriction fragment length polymorphisms (RFLP) and amplified fragment length polymorphisms (AFLP) in different mapping populations. There is a growing need for saturating the genetic map with single nucleotide polymorphisms (SNP), more amenable for high throughput analysis, especially if these markers are located in gene coding regions, to provide functional markers. Expressed sequence tags (ESTs) from melon are available in public databases, and resequencing ESTs or validating SNPs detected *in silico* are excellent ways to discover SNPs.

Results: EST-based SNPs were discovered after resequencing ESTs between the parental lines of the PI 161375 (SC) × 'Piel de sapo' (PS) genetic map or using *in silico* SNP information from EST databases. In total 200 EST-based SNPs were mapped in the melon genetic map using a bin-mapping strategy, increasing the map density to 2.35 cM/marker. A subset of 45 SNPs was used to study variation in a panel of 48 melon accessions covering a wide range of the genetic diversity of the species. SNP analysis correctly reflected the genetic relationships compared with other marker systems, being able to distinguish all the accessions and cultivars.

Conclusion: This is the first example of a genetic map in a cucurbit species that includes a major set of SNP markers discovered using ESTs. The PI 161375 × 'Piel de sapo' melon genetic map has around 700 markers, of which more than 500 are gene-based markers (SNP, RFLP and SSR). This genetic map will be a central tool for the construction of the melon physical map, the step prior to sequencing the complete genome. Using the set of SNP markers, it was possible to define the genetic relationships within a collection of forty-eight melon accessions as efficiently as with SSR markers, and these markers may also be useful for cultivar identification in Occidental melon varieties.

Background

Single-nucleotide polymorphisms (SNPs) are the most frequent type of variation found in DNA [1] and are valuable markers for high-throughput genetic mapping, genetic variation studies and association mapping in crop plants. Several methods have been described for SNP discovery [2]: SNP mining from expressed sequence tag (EST) databases [3]; based on array hybridization [4] or amplicon resequencing [5]; from the complete sequence of a genome [6] and more recently, using high-throughput sequencing technologies [7]. The discovery of SNP markers based on transcribed regions has become a common application in plants because of the large number of ESTs available in databases, and EST-SNPs have been successfully mined from EST databases in non-model species such as Atlantic salmon [8], catfish [9], tomato [10] and white spruce [11].

Melon (*Cucumis melo* L.) is an important crop worldwide. It belongs to the *Cucurbitaceae* family, which also includes cucumber, watermelon, pumpkin and squash. The melon genome has an estimated size of 450 Mb [12] and is a diploid with a basic chromosome number of $x = 12$. In recent years research has been carried out to increase the genetic and genomic resources for this species, such as the sequencing of ESTs [13], the construction of a BAC library [14], the development of an oligo-based microarray [15] and the development of a collection of near isogenic lines (NILs) [16]. Genetic maps have also been reported for melon, but they have been constructed with different types of molecular markers and genetic backgrounds [17-21], making it difficult to transfer markers from one map to another. The aim of the International Cucurbit Genomics Initiative (ICuGI) [22], currently in progress, is to obtain a consensus genetic map by merging genetic maps available using a common set of SSRs as anchor markers.

A double haploid line (DHL) population from the cross between the Korean accession PI 161375 (SC) and the *inodorus* type 'Piel de sapo' T111 (PS) was the basis for the construction of a genetic map with 221 co-dominant, transferable RFLP and SSR markers [21]. New EST-derived SSR markers, added to this map using a bin-mapping strategy with only 14 mapping individuals, gave a new map with 296 markers distributed in 122 bins and a density of 4.2 cM/marker [21]. There is a need for saturating the SC \times PS genetic map with more markers that are amenable for large-scale genotyping, as are SNPs. In a preliminary experiment with melon, amplicon resequencing of 34 ESTs in SC and PS was used for SNP discovery, obtaining a frequency of one SNP every 441 bp and one indel every 1,666 bp [23]. The availability of more than 34,000 melon ESTs from normalized cDNA libraries from different melon genotypes and tissues [13] is a valuable resource for the identification of SNPs to be added to the current genetic map.

Genetic markers can also be used for variability analysis studies. In melon, there have been several attempts to elucidate intraspecific relationships among melon germplasm, using isozyme [24], RFLP [25], RAPD [26], AFLP [27] and SSR [28] markers, with SSRs the preferred marker for fingerprinting and genetic variability analysis in melon [28]. Due to the absence of a known set of SNPs in the species, this marker has not been compared with other types for variability analysis. It would be of special interest to have a set of these markers for a high-throughput system to identify the germplasm used in breeding programs, mainly from *inodorus* and the *cantalupensis* melon types.

The objectives of this work were to increase the marker resolution in the melon genetic map, discovering EST-SNPs in a melon EST database, and to study the performance of a subset of EST-SNPs for variability analysis in a collection of melon accessions.

Results and discussion

SNP discovery

Two strategies were used to discover SNPs in melon. The first was based on producing amplicons from randomly selected melon ESTs and resequencing the parental lines of the melon genetic map PI 161375 (SC) \times 'Piel de sapo' T111 (PS). Primers were designed from 223 melon ESTs (Table 1). After discarding primers that did not amplify a PCR product, amplicons that did not produce high quality sequences and monomorphic amplicons, 93 ESTs (56.3%) showed at least one polymorphism between SC and PS.

The second strategy was the validation of *in silico* SNPs from the ICuGI database [22]. Three hundred and sixty-six *in silico* SNPs found in the database were selected, belonging to two types of SNPs: pSNP and pSCH (Table 1; see methods). Primers were designed from 269 ESTs containing pSNP and 97 containing pSCHs. Putative *in silico* SNPs were validated in 51.8% and 21.3% of the amplicons for pSNPs and pSCHs, respectively. In some instances additional SNPs were detected in the sequenced regions, giving a slightly higher percentage of polymorphic amplicons (69.7% and 31.3% for pSNP and pSCH amplicons, respectively). From the ESTs reported by Gonzalez-Ibeas et al. [13], 47.3% were obtained from two accessions of the 'Piel de sapo' cultivar type (Pinyonet and PS), and the remainder from two genotypes, the C-35 cantaloupe accession (29.3%) and the pat81 *agrestis* accession (23.4%). The pSNPs and pSCHs were deduced from this set of EST sequences, with a high proportion found between pat81 and 'Piel de sapo', and SNPs experimentally validated after resequencing amplicons from PS and SC. SC belongs to the *agrestis* melon type as the accession pat81 but has a different origin, so, as expected not all the SNPs were conserved between SC and PS, giving a pSNP validation of 51.8%. On the other hand, only 21.3% of

Table 1: Amplicons designed from ESTs for SNP discovery

	Amplicons	Failed	Monomorphic	Polymorphic	Polymorphic amplicons*	<i>In silico</i> SNP validation
Random ESTs	223	58	72	93	56.3%	
<i>in silico</i> pSNPs	269	41	69	159	69.7%	51.8%
<i>in silico</i> pSCHs	97	14	57	26	31.3%	21.3%
TOTAL	589	113	198	278	58.4%	

ESTs were selected at random or chosen because they contained pSNPs or pSCHs in the MELOGEN database. Columns show the number of amplicons that failed to amplify or gave bad quality sequences, and monomorphic and polymorphic amplicons between SC and PS. The percentages of polymorphic amplicons and *in silico* SNPs that were validated are shown in the last two columns. (*) Polymorphic amplicons rate was calculated without considering failed amplicons.

the pSCHs were validated, indicating that many may represent sequencing errors or mutations introduced during the cDNA synthesis procedure. The SNPs in a subset of amplicons containing *in silico* SNPs between 'Piel de Sapo' and pat81 were validated using different genotyping methods (see below) rather than resequencing in PS and SC.

A total of 368 amplicons (random and containing *in silico* SNPs) were resequenced in PS and SC and produced 177.5 kb of melon DNA, with 431 SNPs and 59 short indels, at an average of one SNP every 412 bp and one indel every 3.0 kb, (Table 2). This is in agreement with the values obtained in a previous small-scale experiment using the same two melon accessions, which gave one SNP every 441 bp and one indel every 1.6 kb [23]. SC and PS belong to the *agrestis* (*C. melo* ssp. *agrestis*) and *inodorus* (*C. melo* ssp. *melo*) melon groups, respectively, which are two of the more distant groups in the species [28]. This may explain the relatively high frequency of SNPs between the cultivars.

SNP detection

Various detection methods were used for genotyping the SNPs in each EST. A restriction site around the SNP position, different in the parental sequences, was used to develop a CAPS marker for 103 EST-SNPs. When more than one SNP was discovered in one amplicon, we selected the most suitable SNP for detection using CAPS. When no restriction enzyme was available to produce a CAPS marker, we used the SNaPshot SNP detection system. Seventy-seven EST-SNPs were genotyped with SNaP-

shot. For 14 ESTs, PS and SC gave a different amplicon size, so they could be genotyped as SCAR markers. Four EST-SNPs were genotyped using DNA sequencing and two were converted into dCAPS. The SNP detection method used for each mapped EST-SNP is shown in Additional file 1.

SNP variability

Forty-five SNPs (see Additional file 2) were randomly chosen to study their variability in a set of melon accessions of worldwide cultivar and botanical types (see Additional File 3). The *inodorus* cultivars were overrepresented in order to assess whether SNPs between distant melon accessions (SC and PS) were also variable among more closely related genotypes.

All SNPs were polymorphic and the mean major allele frequency was 0.69 (Table 3). Only one SNP (AL_24-H05) had a rare allele (frequency = 0.08), whereas the frequencies of the two alleles were similar in 28 SNPs (major allele frequency < 0.65). Average gene diversity (He) was 0.4 (ranging from 0.14 to 0.5). Forty-three SNPs yielded He > 0.20, demonstrating that most of the chosen SNPs were highly informative, as found for SNPs in rye [29] but contrasting with crops such as soybean [30] and wheat [31] where SNPs yielding rare alleles are more frequent.

The mean gene diversity index for SNPs was considerably lower than the values reported for SSRs in melon (e.g. PIC = 0.58 [21], He = 0.66 [28]). To ensure the difference was not due to sampling, gene diversity indexes were estimated using a subset of genotypes that had been included

Table 2: Frequency of SNPs and indels found after resequencing EST-derived amplicons

Amplicons sequenced in SC and PS	Length sequenced (bp)	SNPs	bp per SNP	indels	bp per indel	Reference
368	177,518	431	411.9	59	3,008.8	this report
34	15,000	34	441.2	9	1,666.6	[23]

Data from a previous report using the same two melon parental lines is shown as a comparison.

Table 3: Gene diversity indexes for SNP and SSR alleles using all, *inodorus* or genotypes described in a previous study [28]

Genotypes	Marker type	Major allele frequency	Ho	He	He range
all	SNP	0.69	0.10	0.40	0.14–0.50
<i>inodorus</i>	SNP	0.85*	0.07	0.15	0–0.50
group used in [28]	SNP	0.63	0.09	0.47	0.16–0.50
group used in [28]	SSR	0.47	0.14	0.64	0.51–0.83

Ho, observed heterozygosity; He expected heterozygosity. * Major allele frequency was only calculated for polymorphic SNPs.

in a previous study with SSRs [28] (see Additional file 3). The differences in gene diversity were confirmed, demonstrating that they were intrinsic to the different marker type. SNPs are biallelic, implying that the He value can not exceed 0.5, whereas SSRs are multiallelic and so it can be higher. Haplotypes may yield higher gene diversity values than individual SNPs and provide more efficient application of SNP markers [29].

All *inodorus* genotypes could be distinguished with the set of SNPs, although polymorphism was notably reduced (Table 3). Fourteen SNPs were monomorphic and 18 were informative (minor allele frequency > 0.1). As most of the SNPs were discovered between the *agrestis* and *inodorus* cultivar and not within *inodorus*, we expected the SNP polymorphism within *inodorus* to be lower. Nevertheless, these results demonstrate that SNPs discovered using a germplasm sample can be successfully transferred to different germplasm samples in melon.

The genetic relationships among accessions based on SNP polymorphism were investigated by cluster analysis. The NJ dendrogram (Figure 1) fits very well with previous classifications using different markers [26,28,32]. Comparing the common genotype set in [28], the average pair-wise distances based on SNPs and SSR were 0.47 and 0.64, respectively. The correlation between the two distance matrices was 0.73 ($P < 0.00001$) according to Mantel's test, confirming that the current SNP set is as effective as SSRs in establishing genetic relationships among melon accessions, as shown in species such as rye [29] and soybean [30].

The population structure was estimated using the STRUCTURE software [33]. The *a posteriori* probability of the data increased rapidly from $K = 1$ to 4 and begun to reach a plateau for $K = 5$, inferring that our collection can be divided in five populations. Genetic variability among melon germplasm seems to be highly structured. The subdivision of the accessions in 5 populations agrees with the botanical classification and the cluster analysis (Figure 1): group 1 included all the *inodorus* cultivars from Spain; group 2, a diverse group of traditional *inodorus* landraces and similar

ones from the Near-East region such as elongated (*chate* and *flexuosus*) and Asiatic *ananas* and *chandalak* types; group 3, modern *cantalupensis* cultivars; group 4, mainly traditional varieties and wild melons from India and Africa and group 5 included *conomon* accessions from the Far East. The population structure should be taken into account when establishing a collection of genotypes for association mapping studies in melon and models including population structure should be used [34]. Alternatively, melon collections without structure, as we found with the *inodorus* melon accessions included in our studies, could be used.

These results demonstrate that SNPs discovered using a small germplasm sample can be transferred to different cultivar groups, being useful for depicting genetic relationships as well as for cultivar identification.

SNP mapping using a bin-mapping strategy

Two hundred and seventy-eight SNP-containing ESTs (Table 1) plus twelve additional SNP-containing ESTs previously discovered between the two parental lines [22] were used for mapping in the SC × PS genetic map using 14 DHLs of the melon bin-mapping population [21]. In total, 199 EST-derived SNPs were mapped, yielding 200 new markers (Figures 2 and 3). F112 produced two SCAR markers (F112a and F112b) that mapped to groups I and V, respectively. Our previous melon bin-map contained 296 markers distributed in 122 bins, with a density of 4.2 cM/marker and 2.4 markers per bin [21]. With the addition of 35 candidate genes previously reported for resistance to virus and fruit ripening [23,35,36] and the SNPs now described, the new bin-map contains 528 markers, distributed in 145 bins, with an increased density of 2.35 cM/marker and 3.64 markers per bin. The SNP-based markers defined 23 new bins with an average bin length of 8.55 cM. Some of the new bins were located in regions with poor marker density in the previous SC × PS melon map [21], such as HS_30-B08 in group XI, AI_12-B08 in group VII, A_38-F04 in group VI or P06.05 in group III.

Essentially the new version of the melon bin-map is a gene-based map, with 412 markers (78%) obtained from

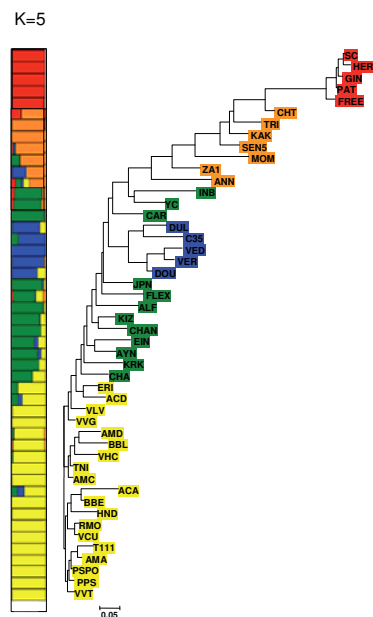


Figure 1
Dendrogram and population structure based on the variability of 45 SNPs in 48 melon accessions. The neighbor-joining (NJ) tree based on Nei genetic distances [44] for the selected melon accessions is shown on the right. The subdivision based on STRUcTURE is shown on the left; each accession on the NJ is colored according to its group assignment defined from STRUcTURE analysis.

gene sequences. Additionally, 114 RFLPs derived from ESTs were previously mapped in an F2 population from the cross SC × PS [37], and their approximate position can also be plotted in the corresponding bin-map. As a large proportion of the markers are codominant and based on gene sequences, this makes this map a very useful tool for melon breeding and comparative analysis in cucurbit species.

With the advent of next generation sequencing technologies, SNP discovery has become more feasible in non-model crop species, allowing the discovery of thousands of SNPs in a single experiment [7]. In *Eucalyptus grandis* more than 23,000 SNPs were discovered using 454 sequencing technology, with a validation rate of 83%

[38]. In melon, a preliminary analysis of 100,000 reads obtained after 454 sequencing of leaf cDNAs from SC and PS produced more than 1,000 SNPs (Garcia-Mas, unpublished). This indicates that the use of next generation sequencing technologies is the next step towards saturation of the melon genetic map.

Conclusion

The set of 200 SNP markers discovered and mapped have increased the marker resolution of the melon genetic map by defining new bins. The genetic map contains more than 500 gene-based codominant markers (SNPs, RFLPs and SSRs), which can be used as anchor points with other genetic maps in this species. This genetic map is also a useful resource for comparative mapping in the *Cucurbitaceae*, for the construction of the melon physical map and for sequencing the melon genome. Additionally, the set of SNPs has proven to be as useful as microsatellites for studying genetic relationships in melon and for varietal identification.

Methods

Plant material and DNA extraction

The parent lines of the melon double haploid line (DHL) mapping population, PI 161375 'Songwan Charmi' (SC) and 'Piel de sapo' line T111 (PS), were used for SNP discovery [20]. Fourteen DHLs from the SC × PS segregating population were used to bin-map the SNP set [21]. The 48 melon genotypes selected for analysis with a subset of SNPs (see Additional file 3) were obtained from the germplasm collection maintained at COMAV (Valencia, Spain) and from a previous study of germplasm variability using SSRs [28]. DNA from all genotypes was extracted using a modified CTAB method [27]. DNA of the forty-eight melon accessions was extracted from leaves of five individuals per accession to take into account the genetic variability within heterogeneous accessions.

SNP discovery and detection

SNPs were discovered using two different strategies. Firstly, random ESTs were selected from the International Cucurbit Genomics Initiative (ICuGI) webpage [22]. Primer pairs were designed from each EST using the Primer3 software [39] with an average length of 20 nucleotides, a melting temperature around 60°C and an expected PCR product of 500–700 bp. Genomic DNA from the parental lines of the melon mapping population was amplified with each primer pair as previously described [23]. Amplified fragments were purified with Sepharose columns and sequenced using the ABI Prism BigDye Terminator Cycle Sequencing kit (Applied Biosystems, Foster City, CA, USA) in an ABI Prism 3130 sequencer (Applied Biosystems, Foster City, CA, USA). Sequences were aligned and screened for polymorphism with the Bioedit software [40]. Putative SNP positions were visually verified on the sequence chromatogram, and

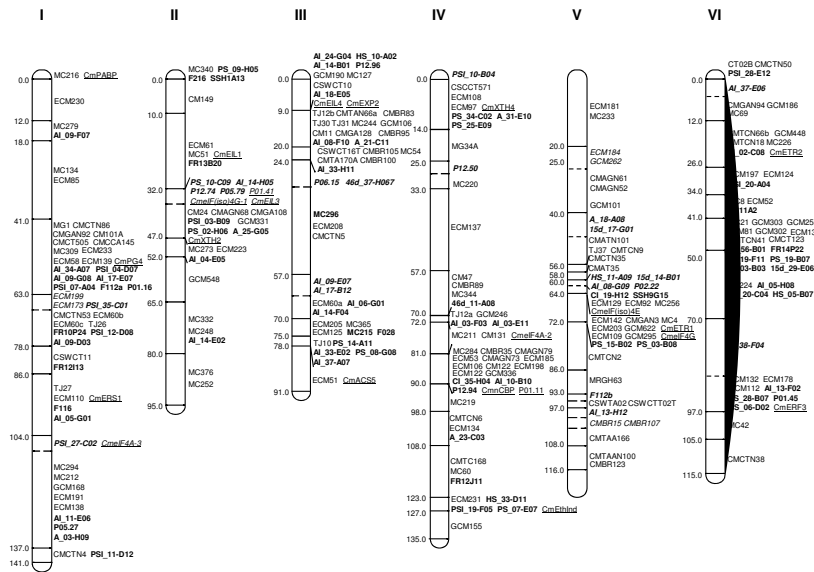


Figure 2
EST-SNP bin map of *Cucumis melo* obtained by selective genotyping of fourteen DHLs. Linkage groups are represented by vertical bars, divided in bins defined by the joint genotype of the selected DHLs. The mapped SNPs in this report are shown in bold. Underlined markers are candidate genes previously reported [23,35,36]. The other markers have been described in [21]. Genetic distances are shown on the left, indicating the position of the last marker included in the bin according to the framework map in [21]. Markers defining new bins are shown in italics. The hypothetical position of the last marker of these bins is indicated by a dashed horizontal line within the linkage group bar, without the genetic distance.

the genomic sequences compared with the original EST sequence to identify any introns. In the second strategy, *in silico* SNPs previously identified [13] using EST2uni [41] were classified as i) pSNPs, corresponding to SNPs present in at least two EST sequences from the same genotype in a given contig and with the same base change and ii) pSCHs, corresponding to single nucleotide variations in sequence that did not follow the above criteria for pSNPs. Selected pSNPs and pSCHs were verified in most cases after resequencing the parental lines of the melon mapping population. For a small subset, the SNP was verified with an appropriate SNP detection method.

Bioedit software was used to generate restriction maps from sequences obtained from SC and PS. SNPs (or indels) showing differential restriction maps were used to

develop cleaved amplified polymorphic sequence (CAPS) markers. When no differential restriction maps were available, the ABI Prism SNaPshot ddNTP Primer Extension Kit (Applied Biosystems) was used for SNP genotyping [23]. Markers F112, 46d_11-A08, FR12J11, 15d_17-G01, P01.45, PSI_26-B12, F012, PS_18-F05, PS_16-C09, F088, A_02-H11, AI_13-G03 and FR15D10 produced amplicons of different sizes in the parental lines, which were not sequenced and were genotyped as sequence characterized amplified region markers (SCARs) after electrophoresis in agarose gels or using a LI-COR IR2 sequencer (Li-Cor Inc, Lincoln, Nebraska, USA). Markers PSI_12-D08 and PSI_35-F11 were converted into dCAPS markers [42]. Markers F028, F149, F080 and PSI_25-B05 were genotyped using direct sequencing.

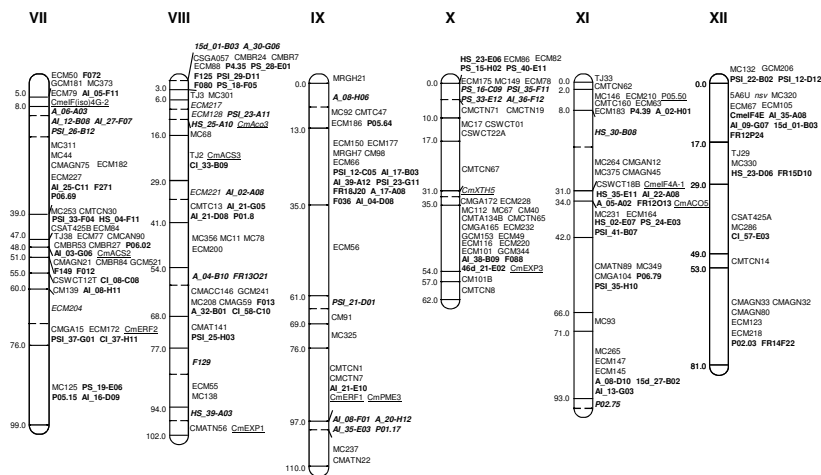


Figure 3
EST-SNP bin map of *Cucumis melo* obtained by selective genotyping of fourteen DHLs. Linkage groups are represented by vertical bars, divided in bins defined by the joint genotype of the selected DHLs. The mapped SNPs in this report are shown in bold. Underlined markers are candidate genes previously reported [23,35,36]. The other markers have been described in [21]. Genetic distances are shown on the left, indicating the position of the last marker included in the bin according to the framework map in [21]. Markers defining new bins are shown in italics. The hypothetical position of the last marker of these bins is indicated by a dashed horizontal line within the linkage group bar, without the genetic distance.

SNP mapping

SNPs and indels were mapped by selective genotyping using the bin-mapping strategy [43], adapted for the melon mapping population [21]. Fourteen out of 72 DHLs from the melon mapping population were selected to obtain the maximum resolution with a minimum number of genotypes. SNPs and indels were placed in the bin map by visual inspection of the genotypes predicted by the markers and genotypes in the bin set.

Genetic variability analysis

Forty-five SNPs from 44 amplicons (two SNPs were selected from F241) were chosen for genetic variability analysis. SNPs were genotyped as CAPS or by pyrosequencing as shown in Additional file 2. Thirty SNPs, described in Additional file 1, were used. Twelve SNPs that were not polymorphic between SC and PS were also included in the variability analysis, and the primers for each amplicon are provided in Additional file 2. The SNPs CmERF1, CmPm3 and CmXTH5 have been previously described [36].

Eight SNPs were genotyped by minisequencing the region surrounding the polymorphism (two SNPs were detected for F241 in the same reaction). Pyrosequencing was performed using a PSQ™ HS 96 system (Pyrosequencing AB, Uppsala, Sweden) following the manufacturers' instructions. Primers were designed with the Pyrosequencing™ Assay Design Software (Biotage AB, Uppsala, Sweden). One of the amplifying primers was 5' end labeled with biotin, allowing the immobilization of the fragment onto M-280 streptavidin coated Sepharose™ dynabeads (DynaL AS, Oslo, Norway). The genotyping primer was hence designed to anneal several nucleotides upstream of the SNP. After denaturation of the streptavidin-captured PCR fragments, the single stranded DNA fragments were released into the wells of the PSQ HS 96 plate. Pyrosequencing was performed using the PSQ HS SNP Reagent kit (Pyrosequencing AB, Uppsala, Sweden), and bioluminescent quantification of pyrophosphate (P_{pp}) released as a result of nucleotide incorporation during DNA synthesis was measured with the PSQ™ HS 96 system.

Allele frequencies, major allele frequency, gene diversity (measured as expected heterozygosity, H_e [44]), genetic distances and neighbor-joining (NJ) tree were calculated using Powermarker 3.25 [45]. The NJ tree was plotted with MEGA 3.0 [46]. Distance matrices were compared by the Mantel test [47].

The number of populations in our collection was deduced with the STRUCTURE software [33]. This package uses a Bayesian clustering approach to identify subpopulations and to assign individuals to these populations on the basis of their genotypes. Given a sample of individuals, K populations are assumed (where K may be unknown) and individuals are assigned to these populations. A posteriori probability for each K (Pr(K)) can be calculated, which is very small for K values lower than the appropriate value. Usually, the researcher fixes a minimum K (for example K = 1), recording Pr(K) after the analysis, and tests increasing Ks, plotting K against Pr(K). The final K is defined when Pr(K) reaches a plateau for higher K values. Consequently, in the current report, several number of populations (from K = 1 to 8) were tested with the software and the total number of populations was set when the probability reached a plateau for higher K.

Authors' contributions

WD discovered and mapped the SNPs and performed the genotyping for the variability analysis. CE and MGT discovered and mapped SNPs. CR discovered SNPs. IFS mapped SNPs. DGI identified and selected *in silico* SNPs. JB carried out the bioinformatics analyses for *in silico* SNPs. AJM performed the variability analysis, coordinated the SNP mapping and participated in the drafting of the manuscript. MBP prepared DNAs for the melon accessions and participated in the genotyping for the variability analysis and in the drafting of the manuscript. JGM, PA, FN, MBP and MAA were involved in the conception of the study. JGM is the principal researcher of this work, supervised it and wrote the manuscript. All authors read and approved the final manuscript.

Additional material

Additional file 1

SNPs markers mapped in the SC × PS genetic map. Shown here, for each SNP marker: the EST and accession number from where it was obtained; best BlastX hit and E-value for each EST; amplicon primer sequences; SNP/indel position; SNP detection method; linkage group and BIN where the marker maps to. *sequence available in [22] or <http://www.melogen.upv.es> without accession number. †SNPs published by Morales et al (2004). ‡SNP position is provided when located in exons and referred to EST in first column. ‡third primer was used for SNaPshot genotyping.

Click here for file
[<http://www.biomedcentral.com/content/supplementary/1471-2229-9-90-S1.xls>]

Additional file 2

SNP markers used for genotyping the melon accessions. The EST from where the SNPs were discovered, the genotyping method (CAPS or pyrosequencing), linkage group where the marker maps to, and source of the marker are given. For unmapped SNP markers, the amplicon primer sequences are given. For SNP markers genotyped using pyrosequencing, forward, reverse and internal primers were used for genotyping. 5'bio: Forward or reverse primer was 5' labeled with biotine. ST1: Additional file 1. Click here for file
[<http://www.biomedcentral.com/content/supplementary/1471-2229-9-90-S2.xls>]

Additional file 3

Forty-eight melon accessions that were examined in this study. Plant assignment (or common name), code used in the current study, accession number from the respective gene banks, cultivar group, origin and seed bank donor (COMAV, Instituto de Conservación y Mejora de la Agrodiversidad Valenciana, Valencia, Spain; USDA/ARS/NCRPIS, North Central Regional Plant Introduction Station, Ames, IA, USA; IPK, Institute of Plant Genetics and Crop Plant Research, Gatersleben, Germany; INRA, Institut National de la Recherche Agronomique, Montfavet, Avignon, France; Semillas Fitó SA, Barcelona, Spain; ARO, Agricultural Research Organization, Ramat Yishay, Israel) are specified for each genotype. Accessions marked with (*) were previously used by Monforte et al. (2003) for an SSR study. Click here for file
[<http://www.biomedcentral.com/content/supplementary/1471-2229-9-90-S3.xls>]

Acknowledgements

This work was supported by a grant from the Ministerio de Educación y Ciencia (Spain) (GEN2003-20237-C06). WD is recipient of a postdoctoral fellowship from the Centre de Recerca en Agrigenòmica CSIC-IRTA-UAB (Spain). CR is recipient of a Juan de la Cierva grant from the Ministerio de Educación y Ciencia (MEC) (Spain). DGI and CS are recipients of pre-doctoral fellowships from MEC (Spain). IFS is recipient of a pre-doctoral fellowship from INIA (Spain). We are grateful to Armand Sanchez and Anna Mercader (UAB) for their help with the pyrosequencing analysis.

References

- Brookes AJ: **The essence of SNPs.** *Gene* 1999, **234(2)**:177-186.
- Ganal MW, Altmann T, Roder MS: **SNP identification in crop plants.** *Curr Opin Plant Biol* 2009, **12(2)**:211-217.
- Batley J, Barker G, O'Sullivan H, Edwards KJ, Edwards D: **Mining for single nucleotide polymorphisms and insertions/deletions in maize expressed sequence tag data.** *Plant Physiol* 2003, **132(1)**:84-91.
- Borevitz JO, Liang D, Plouffe D, Chang HS, Zhu T, Weigel D, Berry CC, Winzeler E, Chory J: **Large-scale identification of single-feature polymorphisms in complex genomes.** *Genome Res* 2003, **13(3)**:513-523.
- Choi YI, Hyten DL, Matukumalli LK, Song Q, Chaky JM, Quigley CV, Chase K, Lark KG, Reiter RS, Yoon MS, et al: **A soybean transcript map: gene distribution, haplotype and single-nucleotide polymorphism analysis.** *Genetics* 2007, **176(1)**:685-696.
- Velasco R, Zharikh A, Troglio M, Cartwright DA, Cestaro A, Pruss D, Pindo M, Fitzgerald LM, Vezzulli S, Reid J, et al: **A high quality draft consensus sequence of the genome of a heterozygous grapevine variety.** *PLoS ONE* 2007, **2(12)**:e1326.
- Barbazuk WB, Emrich SJ, Chen HD, Li L, Schnable PS: **SNP discovery via 454 transcriptome sequencing.** *Plant J* 2007, **51(5)**:910-918.
- Hayes B, Laerdahl JK, Lien S, Moen T, Berg P, Hindar K, Davidson WS, Koop BF, Adzhubei A, Heyheim B: **An extensive resource of sin-**

- gle nucleotide polymorphism markers associated with Atlantic salmon (*Salmo salar*) expressed sequences. *Aquaculture* 2007, **265**:82-90.
9. Wang S, Sha Z, Sonstegard TS, Liu H, Xu P, Somridhivej B, Peatman E, Kucuktas H, Liu Z: **Quality assessment parameters for EST-derived SNPs from catfish.** *BMC Genomics* 2008, **9**:450.
 10. Yamamoto N, Tsugane T, Watanabe M, Yano K, Maeda F, Kuwata C, Toriki M, Ban Y, Nishimura S, Shibata D: **Expressed sequence tags from the laboratory-grown miniature tomato (*Lycopersicon esculentum*) cultivar Micro-Tom and mining for single nucleotide polymorphisms and insertions/deletions in tomato cultivars.** *Gene* 2005, **356**:127-134.
 11. Pavy N, Parsons LS, Paule C, MacKay J, Bousquet J: **Automated SNP detection from a large collection of white spruce expressed sequences: contributing factors and approaches for the categorization of SNPs.** *BMC Genomics* 2006, **7**:174.
 12. Arumuganathan K, Earle ED: **Nuclear DNA content of some important plant species.** *Plant Mol Biol Rep* 1991, **9**:208-218.
 13. Gonzalez-Ibeas D, Blanca J, Roig C, Gonzalez-To M, Pico B, Truniger Y, Gomez P, Deleu W, Cano-Delgado A, Arus P, et al.: **MELOGEN: an EST database for melon functional genomics.** *BMC Genomics* 2007, **8**:306.
 14. van Leeuwen H, Monfort A, Zhang HB, Puigdomenech P: **Identification and characterisation of a melon genomic region containing a resistance gene cluster from a constructed BAC library. Microcolinearity between *Cucumis melo* and *Arabidopsis thaliana*.** *Plant Mol Biol* 2003, **51**(5):703-718.
 15. Mascarelli-Cireus A, Carrizosa J, Vilarasa J, Mora-Garcia S, Blanca J, Gonzalez-Ibeas D, Saladie M, Roig C, Deleu W, Pico B, et al.: **An oligo-based microarray offers novel transcriptomic approaches for the analysis of pathogen resistance and fruit quality traits in melon.** *BMC Genomics* in press.
 16. Eduardo I, Arus P, Monforte AJ: **Development of a genomic library of near isogenic lines (NILs) in melon (*Cucumis melo* L.) from the exotic accession PI161375.** *Theor Appl Genet* 2005, **112**(1):139-148.
 17. Wang YH, Thomas CE, Dean RA: **A genetic map of melon (*Cucumis melo* L.) based on amplified fragment length polymorphism (AFLP) markers.** *Theor Appl Genet* 1997, **95**:791-798.
 18. Danin-Poleg Y, Reis N, Baudracco-Arnas S, Pitrat M, Staub JE, Oliver M, Arus P, deVicente CM, Katzir N: **Simple sequence repeats in *Cucumis* mapping and map merging.** *Genome* 2000, **43**(6):963-974.
 19. Perin C, Hagen S, De Conto V, Katzir N, Danin-Poleg Y, Portnoy V, Baudracco-Arnas S, Chadoeuf J, Dogimont C, Pitrat M: **A reference map of *Cucumis melo* based on two recombinant inbred line populations.** *Theor Appl Genet* 2002, **104**(6-7):1017-1034.
 20. Gonzalo MJ, Oliver M, Garcia-Mas J, Monforte A, Dolcet-Sanjuan R, Katzir N, Arus P, Monforte AJ: **Simple-sequence repeat markers used in merging linkage maps of melon (*Cucumis melo* L.).** *Theor Appl Genet* 2005, **110**(5):802-811.
 21. Fernandez-Silva I, Eduardo I, Blanca J, Esteras C, Pico B, Nuez F, Arus P, Garcia-Mas J, Monforte AJ: **Bin mapping of genomic and EST-derived SSRs in melon (*Cucumis melo* L.).** *Theor Appl Genet* 2008, **118**(1):139-150.
 22. **The International Cucurbit Genomics Initiative (ICuGI)** [<http://www.icugi.org>]
 23. Morales M, Roig E, Monforte AJ, Arus P, Garcia-Mas J: **Single-nucleotide polymorphisms detected in expressed sequence tags of melon (*Cucumis melo* L.).** *Genome* 2004, **47**(2):352-360.
 24. Staub JE, Box J, Meglic V, Horejsi TF, McCreight JD: **Comparison of isozyme and random amplified polymorphic DNA data for determining intraspecific variation in *Cucumis*.** *Genet Res Crop Evol* 1999, **44**:257-269.
 25. Neuhausen SL: **Evaluation of restriction fragment length polymorphisms in *Cucumis melo*.** *Theor Appl Genet* 1992, **83**:379-384.
 26. Stepansky A, Kovalski I, Perl-Treves R: **Intraspecific classification of melons (*Cucumis melo* L.) in view of their phenotypic and molecular variation.** *Plant Syst Evol* 1999, **217**:313-332.
 27. Garcia-Mas J, Oliver M, Gomez H, de Vicente MC: **Comparing AFLP, RAPD and RFLP markers to measure genetic diversity in melon.** *Theor Appl Genet* 2000, **101**:860-864.
 28. Monforte AJ, Garcia-Mas J, Arus P: **Genetic variability in melon based on microsatellite variation.** *Plant Breeding* 2003, **122**:153-157.
 29. Varshney RK, Beier U, Khlestkina EK, Kota R, Korzun V, Graner A, Börner A: **Single nucleotide polymorphisms in rye (*Secale cereale* L.): discovery, frequency, and applications for genome mapping and diversity studies.** *Theor Appl Genet* 2007, **114**(6):1105-1116.
 30. Yoon MS, Song QJ, Choi IY, Specht JE, Hyten DL, Cregan PB: **BARCSoySNP23: a panel of 23 selected SNPs for soybean cultivar identification.** *Theor Appl Genet* 2007, **114**(5):885-899.
 31. Ravel C, Praud S, Murigneux A, Canaguier A, Sapet F, Samson D, Bal-fourier F, Dufour P, Chalhoub B, Brunel D, et al.: **Single-nucleotide polymorphism frequency in a set of selected lines of bread wheat (*Triticum aestivum* L.).** *Genome* 2006, **49**(9):1131-1139.
 32. Monforte AJ, Eduardo I, Abad S, Arus P: **Inheritance mode of fruit traits in melon-heterosis for fruit shape and its correlation with genetic distance.** *Euphytica* 2005, **144**:31-38.
 33. Pritchard JK, Stephens M, Donnelly P: **Inference of population structure using multilocus genotype data.** *Genetics* 2000, **152**(2):945-959.
 34. Pritchard JK, Stephens M, Rosenberg NA, Donnelly P: **Association mapping in structured populations.** *Am J Hum Genet* 2000, **67**(1):170-181.
 35. Essafi A, Diaz-Pendon JA, Moriones E, Monforte AJ, Garcia-Mas J, Martin-Hernandez AM: **Dissection of the oligogenic resistance to Cucumber mosaic virus in the melon accession PI 161375.** *Theor Appl Genet* 2009, **118**(2):275-284.
 36. Moreno E, Obando JM, Dos-Santos N, Fernandez-Trujillo JP, Monforte AJ, Garcia-Mas J: **Candidate genes and QTLs for fruit ripening and softening in melon.** *Theor Appl Genet* 2008, **116**(4):589-602.
 37. Oliver M, Garcia-Mas J, Cardus M, Pueyo N, Lopez-Sese AL, Arroyo M, Gomez-Panigua H, Arus P, de Vicente MC: **Construction of a reference linkage map for melon.** *Genome* 2001, **44**(5):836-845.
 38. Novaes E, Drost DR, Farmerie WG, Pappas GJ Jr, Grattapaglia D, Sederoff RR, Kirst M: **High-throughput gene and SNP discovery in *Eucalyptus grandis*, an uncharacterized genome.** *BMC Genomics* 2008, **9**:312.
 39. Rozen S, Skaletsky H: **Primer3 on the WWW for general users and for biologist programmers.** *Methods Mol Biol* 2000, **132**:365-386.
 40. Hall TA: **BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT.** *Nucleic Acids Symp Ser* 1998, **41**:95-98.
 41. Forment J, Gilibert F, Robles A, Conejero V, Nuez F, Blanca JM: **EST2uni: an open, parallel tool for automated EST analysis and database creation, with a data mining web interface and microarray expression data integration.** *BMC Bioinformatics* 2008, **9**:5.
 42. Neff MM, Neff JD, Chory J, Pepper AE: **dCAPS, a simple technique for the genetic analysis of single nucleotide polymorphisms: experimental applications in *Arabidopsis thaliana* genetics.** *Plant J* 1998, **14**(3):387-392.
 43. Howard W, Yamamoto T, Dirlwanger E, Testolin R, Cosson P, Cipriani G, Monforte AJ, Georgi L, Abbott AG, Arus P: **Mapping with a few plants: using selective mapping for microsatellite saturation of the *Prunus* reference map.** *Genetics* 2005, **171**(3):1305-1309.
 44. Nei M, Tajima F, Tateno Y: **Accuracy of estimated phylogenetic trees from molecular data. II. Gene frequency data.** *J Mol Evol* 1983, **19**(2):153-170.
 45. Liu K, Muse SV: **PowerMarker: an integrated analysis environment for genetic marker analysis.** *Bioinformatics* 2005, **21**(9):2128-2129.
 46. Tamura K, Dudley J, Nei M, Kumar S: **MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0.** *Mol Biol Evol* 2007, **24**(8):1596-1599.
 47. Mantel N: **The detection of disease clustering and a generalized regression approach.** *Cancer Res* 1967, **27**(2):209-220.

ANEXO II

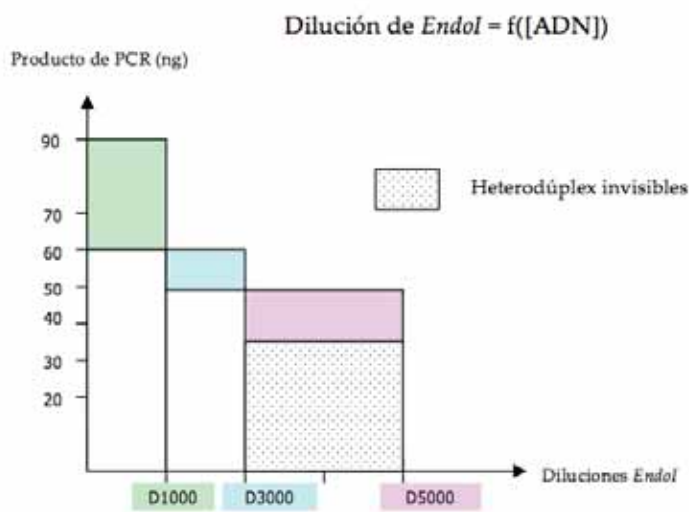
II.a Tabla con los cebadores utilizados para las amplificaciones de cada gen para búsqueda de mutantes TILLING. En la columna de la izquierda el nombre de los cebadores y a la derecha las secuencias correspondientes.

Nombre	Cebador
eIE4E 1er amplicón	
Intro1-R	CTATCTACCAAGTTTCTAAATT
cDNA5'F	GACTCAAACGCCTAACAGAAAATC
Intron1Rev-IRD800	TCCCTAAATCGAACCAAGAAACGCC
FullcDNA5'F-IRD700	GAGGGCGGTGCCATTCTTCTCGG
eIF4E 2o amplicón	
Intron1-F	TTGTTTCGTGTGACATGTCCATC
cDNA3'R	AGCATGCATTTACCTCATTGGCT
Intron1For-IRD700	TGCTTGCGTGTTAATTTATCTCTGC
FullcDNA3'R-IRD800	GTCAAGTACAGAACAAGAATCTGAG
eIF(iso)4E	
iso4EEF	GATCGATAATTTCCCTTTTC
iso4EER	TGACAGGCTTAAGCACTATG
iso4EIF-IRD700	ACCCAATTCATTCTAGGG
iso4EIR-IRD800	TGTGCAAGCGCAACAAGGTAC
ETR1	
ETR1-1F	GCAGAAGAACCCAGATAAAC
ETR1-1R	AGAGCAACGTTCTGCTCCAT
ETR1-8F-IRD700	CATGTGCAACTGCCCTTATG
ETR1-8R-IRD800	CCTGGTCAGCAACAATTCA
PDS	
PDS-6F	CAATGGGATGCTTAGATCAAC
PDSUTR3'-2	TCCTTCCAAAATTACCCTAG
PDS4F-IRD700	GGATTTCTCGGAGTGACTCGG
PDSR-IRD800	GCTCTCGAGCAACTAGCACT

II.b Tampón de PCR

MgCl₂ 15 mM
 KCl 500 mM
 Tris 100 mM
 Gelatina 0.01%

II.c Gráfico de la dilución necesaria de enzima *EndoI* en función de la concentración de ADN obtenida en la *Nested PCR*. Dependiendo de los ng de producto obtenidos en la PCR debe usarse una concentración adecuada de enzima para que la digestión sea óptima. (Datos proporcionados por el grupo del INRA, Dr. Abdel Bendahmane).

**II.d** Tampón de dilución de la enzima *EndoI*

50 mM Tris HCl pH8
 100 mM KCl
 10 μ M ZnCl₂
 0.01% Triton X-100
 50% Glicerol
 20 μ g/mL BSA Acetilado

II.e Tampón de reacción para la digestión con *EndoI*

1M Hepes pH7,5
0.1M MgSO₄
0.02% Triton (X-100)
0.1M KCl
QSP H₂O

II.f Tampón de extracción de ADN Doyle

2% CTAB
1,4M NaCl
20mM EDTA
100mM Tris-HCl (pH 8.0)
Ajustar PH a 8.00, autoclavar y dejar enfriar. Finalmente añadir 0.2% de β-mercaptoetanol

II.g Tampón de lavado Doyle

Etanol 76%
Acetato amónico 10mM

II.h Azul de carga para geles de agarosa

Azul de bromofenol 0.6% (p/v)
Glicerol 70% (v/v)
SDS 0.2% (v/v)
NEB 0.5x
EDTA 20 mM

II.i Tampón NEBX10

Tris 1M
EDTA 10mM
NaAc-3H₂O 0.12M

II.j Tabla con las secuencias de los cebadores utilizados durante la secuenciación de cada gen:

II.j.1 PDS:

Nombre cebador	Secuencia
PDSF	CGATGGGAATGTAATTGAAGG
W207	GCTTCTTTGCCTGATGACT
PDS2F	GAGACGAAGCATGAGGAGTCC
PDS3F	AAGGTTCTTGGCGAACTATG
W208	CTGACATGTCCGCATAAACAC
PDSI3F	TGGCACATCTGGTAGCAATTC
PDSI3R	GTCGAAAAGCAAGGACAGAAC
PDS6F	CAATGGGATGCTTAGATCAAC
PDS4F	GGATTCTCGGAGTGACTCGG
PDS4R	CCGAGTCACTCCGAGAAATCC
PDSR	GCTCTCGAGCAACTAGCACT
PDSUTR3'-2	TCCTTCCAAAATTACCCTAG

II.j,2 *eIF(iso)4E*

Nombre cebador	Secuencia
Iso4EER	TGACAGGCTTAAGCACTATG
Iso4EIR	TGTGCAAGCGCAACAAGGTAC
Iso4EIF	ACCCCAATTCATCTAGGG
Iso4EEF	GATCGATAATTTCCTTTTIC

II.j,3 *ETR1*

Nombre cebador	Secuencia
ETR1-1F	GCAGAAGAACCCAGATAAAC
ETR1-8F	CATGTGCAACTGCCCTTATG
ETR1-8R	CCTGGTCAGCAACAACCTCA
ETR1-2F	GGCCAGAGCTTTCGACTAAG
ETR1-1R	AGAGCAACGTTCTGCTCCAT
ETR1-4F	GGCTAGAGATCTCTTAATGG
ETR1-5F	GCTGTTGGTGACGAGAAACG
ETR1-4R	GACTATCACTGGCACAGGG
ETR1-6F	GGAAGGACATATATGGCTT
ETR1-5R	CGTCCATAACAAGGACTTTG
ETR1-3F	GGAGTTAGTCGGTCAGTGAC
ETR1-6R	CTGTGTTCTGGGAGACTAC
ETR1-3R	CCTATCTGGGTATCTAGTAC

II.k 113 variedades usadas para el EcoTILLING pertenecientes al banco de germoplasma del COMAV.

Código de accesión	Grupo	Nombre local	País de origen
CUM 254	dudaim	DUA	Afganistán
PI 125951	melo-chandalack	AFG	Afganistán
CUM424	melo-cantalup??	pjeper	Albania
Bol-84	melo-adana	melón	Bolivia
CUM535	melo-cantalup??		Bulgaria
CUM206	conomon	COC	China
PI618854	conomon	Longtian	China
CUM510	melo	dinja-obicna	Croacia
ECU1476	melo		Ecuador
PI288233	melo	Kafr hakim	Egipto
CUM198	cantalupensis	Cantaloup de Bellegard	Eslovaquia
ASI-C-11	melo	Melon nanatri	Georgia
PI-314427	melo-cantalup??	Koljonitza	Georgia
PI 185111	melo-tibish??		Ghana
CUM53	melo-cantalup??	BAK (South Balkan)	Grecia
SUD-C-2	melo-cantalup??	Melon	Guatemala
CUM 133	inodorus	Muscatello Cukordinnye	Hungría
CUM 305	cantalupensis-ameri	Ezüst Ananasz	Hungría
CUM 438	momordica	Kharbuja	India
CUM225	flexuosus		India
PI124112	momordica		India
PI 164320	chito-agrestis	CHT	India
CUM 344	flexuosus	Khiar taaruzy	Irak
CUM 349	flexuosus		Irak
CUM 355	melo, inod-asiatico		Irak
PI 140632	cantalupensis		Irán
CUM 363	adzhur -chate	carosello	Italia
CUM372	melo	cucumarazzo	Italia
PI 266947	catalupensis		Japón
PI 420176	conomon-makuwa	ginsen-makuwa	Japón

ASI-C-10	melo	mestnaia	Kazajstán
Pat81	agrestis	Agrestis pat-81	Corea
CUM294	cantalupensis	Khiar	Libia
PI 490388	cantalup	Knakani	Mali
CUM338	chandalak		Mongolia
Afr-c-1	melo-reticulatus		Marruecos
CUM-287	agrestis		Nigeria
ACC04844	melo	melao	Portugal
M-56197	melo	Casca de carvalho	Portugal
CUM 168	melo	Korça	Rusia
CUM259	cantalup	Apelsinaja	Rusia
PI 276660	chandalak		Rusia
PI 506460	cantalupensis	kuvisnka	Rusia
PI 436534	agrestis		Senegal
PI 282448	melo		Sur África
8207	inodorus	Caña Dulce	España
A-049	inodorus	Honey-dew	España
AN-C-109	inodorus	Melón coca	España
AN-C-133	cantalup	Melón escrito oloroso	España
AN-C-7	inodorus	Verde gordo	España
AN-C-70	inodorus	Blanco escriturado	España
AN-C-8	inodorus	Tendral negro de invierno	España
AN-C-90	inodorus	Tempranillo	España
C-198	inodorus	Hilo carrete	España
C-199	inodorus	Blanco liso	España
C-241	inodorus	SE-2811-1-C	España
C-550	cantalup??	Melon azucarado	España
C-62	inodorus	Madura amarilla	España
C-C-19	inodorus	Mollerusa	España
C-C-26	cantalupensis	Meló CCA	España
CC-33	inodorus	Tendral	España
C-C-39	inodorus	Cuarenteno	España

Ch 17	inodorus	Blanco	España
Ch13	inodorus	Melon negro de invierno	España
Ch5	inodorus	Pipa de oro	España
CL-C-1	cantalupensis	Verdejo-casero rayado	España
CM-17	inodorus	Melon piñoncillo	España
CM-25	inodorus	Amarillo	España
CM-C-23	inodorus	Blanco	España
CM-C-50	inodorus	Mochuelo	España
CM-C-53	inodorus	Largo de Villaconejos	España
E -C-29	inodorus	Melona amarilla	España
E-C-1	inodorus	Calamonte	España
E-C-3	inodorus	Hilo Carrete	España
EC-39	inodorus	Verde pinto	España
<u>E-C-45</u>	inodorus	Amarillo manchado	España
Erico	inodorus	Eriçó mallorquin	España
Franceset			España
J.A 8208	inodorus	Invernizo	España
J.A 8214	inodorus	Común	España
J.A 8215	inodorus	Invern	España
J.A 8217	inodorus	Bolas	España
J.A 8221	inodorus	Mochuelo	España
MU-C-17	inodorus	Amarillo oro	España
Muc-21	inodorus	rochet	España
MU-C-35	inodorus	Amarillo canario	España
MU-C-36	inodorus	Tendral	España
MU-C-44	inodorus	Tendral tardío	España
V -C-136	inodorus	Melón amarillo	España
V-C-131	flexuosus	Alficos combro	España
ASI-C-9	melo	Melon tokash	Tadshikistán
CUM416	battik		Tunicia
PI 167057	flexuosus	Acuk	Turquía
PI 169305	inodorus	kirkagac	Turquía

PI 169329	inodorus	winter type	Turquía
PI 169331	melo, inodorus	al timbas	Turquía
PI 169368	melo , inodorus	Hassambey	Turquía
PI 506459	cantalupensis		Ucrania
cum468	melo	Stutz Supreme	USA
cum474	melo, catalupensis	Golden Champlain	USA
cum500	cantalupensis	Golden Honey	USA
nsl30032	cantalupensis	topmark	USA
asic-5	chandalack	kotcha	Uzbekistán
asic-6	ameri	kizil-uruk	Uzbekistán
y285	melo		Yugoslavia
PI 505602	agrestis		Zambia
PI 482429	agrestis		Zimbaue

ANEXO III

III.a Posición de las familias mutantes en las 25 placas que constituyen la población de TILLING de *Cucumis melo* elaborada en este trabajo. Las siglas identifican el lugar de crecimiento de la familia M2 (C:COMAV, F: Semillas Fitó S.A). *Familias con poco ADN o prácticamente nulo.

P1

	1	2	3	4	5	6	7	8	9	10	11	12
A	C1	C10	C21	C33	C42	C51	C60	C73	C83	C91	C104	C117
B	C2	C11	C23	C34	C44	C53	C64	C74	C84	C92	C108	C118
C	C3	C12	C26	C35	C45	C54	C65	C76	C85	C94	C109	C119
D	C4	C14	C27	C36	C46	C55	C66	C77	C86	C96	C111	C120
E	C5	C15	C29	C37	C47	C56	C68	C79	C87	C97	C113	C121
F	C6	C16	C30	C38	C48	C57	C70	C80	C88	C99	C114	C122
G	C7	C17	C31	C39	C49	C58	C71	C81	C89	C100	C115	C123
H	C8	C18	C32	C41	C50	C59	C72	C82	C90	C102	C116	C125

P2

	1	2	3	4	5	6	7	8	9	10	11	12
A	C126	C136	C149	C181	C207	C231	C241	C250	C259	C272	C290	C304
B	C127	C138	C150	C183	C210	C232	C242	C252	C260	C282	C291	C307
C	C128	C139	C157	C186	C220	C234	C243	C253	C261	C283	C293	C308
D	C130	C142	C158	C190	C221	C235	C244	C254	C263	C284	C296	C310
E	C131	C144	C159	C198	C222	C237	C245	C255	C264	C285	C298	C312
F	C132	C146	C163	C199	C223	C238	C246	C256	C265	C287	C299	C313
G	C134	C147	C165	C201	C224	C239	C247	C257	C266	C288	C300	C315
H	C135	C148	C168	C204	C227	C240	C248	C258	C267	C289	C301	C317

P3

	1	2	3	4	5	6	7	8	9	10	11	12
A	C319	C332	C344	C364	C384	C397	C411	C424	C437	C447	C460	C478
B	C320	C333	C350	C366	C385	C398	C414	C425	C438	C449	C462	C479
C	C321	C334	C353	C372	C388	C399	C415	C427	C439	C450	C463	C481
D	C325	C335	C354	C373	C390	C403	C417	C429	C440	C451	C466	C482
E	C326	C336	C355	C377	C391	C406	C418	C430	C441	C453	C467	C483
F	C327	C338	C356	C378	C392	C407	C419	C433	C442	C454	C473	C485
G	C330	C341	C359	C379	C393	C408	C420	C435	C443	C458	C474	C486
H	C331	C343	C360	C380	C395	C410	C423	C436	C444	C459	C477	C487

P4

	1	2	3	4	5	6	7	8	9	10	11	12
A	C489	F13	F21	F29	F37	F47	F56	F65	F74	F83	F93	F102
B	F1	F14	F22	F30	F39	F48	F58	F66	F75	F84	F94	F103
C	F5	F15	F23	F31	F40	F50	F59	F68	F76	F85	F96	F104
D	F6	F16	F24	F32	F41	F51	F60	F69	F77	F86	F97	F105
E	F8	F17	F25	F33	F42	F52	F61	F70	F78	F87	F98	F106
F	F9	F18	F26	F34	F43	F53	F62	F71	F79	F88	F99	F107
G	F11	F19	F27	F35	F44	F54	F63	F72	F80	F89	F100	F108
H	F12	F20	F28	F36	F45	F55	F64	F73	F81	F90	F101	F109

P5

	1	2	3	4	5	6	7	8	9	10	11	12
A	F110	F120	F129	F139	F149	F160	F169	F180	F188	F196	F205	F215
B	F111	F121	F131	F140	F150	F161	F170	F181	F189	F197	F206	F216
C	F112	F122	F133	F141	F151	F162	F171	F182	F190	F198	F207	F217
D	F114	F123	F134	F142	F153	F163	F173	F183	F191	F199	F208	F218
E	F116	F124	F135	F143	F154	F164	F175	F184	F192	F200	F210	F219
F	F117	F125	F136	F144	F157	F165	F176	F185	F193	F201	F211	F220
G	F118	F126	F137	F145	F158	F167	F178	F186	F194	F203	F213	F221
H	F119	F127	F138	F148	F159	F168	F179	F187	F195	F204	F214	F222

P6

	1	2	F	4	5	6	7	8	9	10	11	12
A	F223	F232	F242	F251	F259	F268	F280	F288	F296	F307	F317	F328
B	F225	F235	F243	F252	F260	F271	F281	F289	F297	F309	F318	F329
C	F226	F236	F244	F253	F261	F272	F282	F290	F298	F310	F319	F330
D	F227	F237	F245	F254	F262	F273	F283	F291	F300	F311	F320	F331
E	F228	F238	F247	F255	F263	F274	F284	F292	F302	F312	F321	F332
F	F229	F239	F248	F256	F265	F275	F285	F293	F303	F313	F322	F333
G	F230	F240	F249	F257	F266	F278	F286	F294	F305	F315	F324	F334
H	F231	F241	F250	F258	F267	F279	F287	F295	F306	F316	F326	F335

P7

	1	2	3	4	5	6	7	8	9	10	11	12
A	F336	F345	F354	F362	F370	F381	F390	F399	F408	F418	F429	F437
B	F337	F346	F355	F363	F373	F382	F391	F400	F409	F419	F430	F438
C	F338	F347	F356	F364	F374	F383	F392	F402	F410	F421	F431	F439
D	F339	F348	F357	F365	F375	F384	F394	F403	F411	F422	F432	F440
E	F340	F349	F358	F366	F376	F385	F395	F404	F414	F423	F433	F441
F	F341	F350	F359	F367	F377	F386	F396	F405	F415	F424	F434	F442
G	F343	F352	F360	F368	F378	F387	F397	F406	F416	F425	F435	F443
H	F344	F353	F361	F369	F380	F389	F398	F407	F417	F427	F436	F444

P8

	1	2	3	4	5	6	7	8	9	10	11	12
A	F445	F454	F465	F476	F485	F496	F506	F517	F526	F534	F542	F555
B	F446	F456	F468	F477	F486	F497	F507	F518	F527	F535	F543	F556
C	F447	F458	F469	F478	F488	F498	F508	F519	F528	F536	F544	F558
D	F448	F459	F470	F479	F489	F499	F510	F520	F529	F537	F545	F559
E	F449	F460	F472	F481	F490	F500	F512	F521	F530	F538	F547	F560
F	F450	F461	F473	F482	F491	F501	F513	F522	F531	F539	F550	F563
G	F451	F462	F474	F483	F494	F502	F514	F523	F532	F540	F551	F564
H	F453	F463	F475	F484	F495	F504	F516	F524	F533	F541	F553	F565

P9

	1	2	3	4	5	6	7	8	9	10	11	12
A	F566	F575	F584	F594	F602	F610	F620	F630	F639	F647	F656	F664
B	F567	F576	F585	F595	F603	F611	F623	F631	F640	F648	F657	F665
C	F568	F578	F586	F596	F604	F612	F624	F633	F641	F649	F658	F666
D	F569	F579	F587	F597	F605	F613	F625	F634	F642	F650	F659	F667
E	F570	F580	F588	F598	F606	F614	F626	F635	F643	F651	F660	F668
F	F572	F581	F590	F599	F607	F616	F627	F636	F644	F652	F661	F669
G	F573	F582	F591	F600	F608	F617	F628	F637	F645	F653	F662	F672
H	F574	F583	F593	F601	F609	F618	F629	F638	F646	F654	F663	F674

P10

	1	2	3	4	5	6	7	8	9	10	11	12
A	F675	F683	F691	F703	F712	F722	F732	F743	I3	I11	I19	I29
B	F676	F684	F692	F704	F714	F723	F734	F744	14	I12	I20	I30
C	F677	F685	F693	F705	F715	F724	F735	F745	15	I13	I21	I31
D	F678	F686	F695	F706	F716	F727	F736	F746	16	I14	I22	I32
E	F679	F687	F697	F707	F718	F728	F737	F748	17	I15	I24	I33
F	F680	F688	F699	F708	F719	F729	F738	F749	18	I16	I25	I34
G	F681	F689	F700	F710	F720	F730	F741	I1	19	I17	I26	I35
H	F682	F690	F702	F711	F721	F731	F742	I2	I10	I18	I28	I36

P11

	1	2	3	4	5	6	7	8	9	10	11	12
A	I37	I45	I54	I63	I72	F777	F785	F793	F803	F811	F821	F835
B	I38	I46	I55	I64	I75	F778	F786	F794	F804	F812	F822	F836
C	I39	I47	I56	I66	I76	F779	F787	F795	F805	F813	F825	F837
D	I40	I48	I57	I67	I77	F780	F788	F796	F806	F814	F826	F838
E	I41	I50	I58	I68	I78	F781	F789	F797	F807	F817	F827	F839
F	I42	I51	I59	I69	C324	F782	F790	F798	F808	F818	F829	F840
G	I43	I52	I60	I70		F783	F791	F801	F809	F819	F832	F841
H	I44	I53	I62	I71		F784	F792	F802	F810	F820	F834	F843

P12

	1	2	3	4	5	6	7	8	9	10	11	12
A	F844	F853	F866	F883	F901	F912	F925	F935	F945	F959	F974	F987
B	F845	F855	F869	F885	F903	F913	F927	F936	F947	F961*	F975*	F988
C	F846	F860	F871	F886	F904	F915	F929	F937	F949	F963	F976*	F990
D	F848	F861	F874	F887	F905	F917	F930	F938	F950	F965*	F977	F992
E	F849	F862	F877	F891	F906	F918	F931	F939	F952	F966	F979	F993
F	F850	F863	F879	F892	F907	F920	F932	F941	F953	F968	F981	F994
G	F851	F864	F880	F893	F908	F921	F933	F943	F957	F971	F985	F995
H	F852	F865	F881	F898	F910	F923	F934	F944	F958	F972	F986	F996

P13

	1	2	3	4	5	6	7	8	9	10	11	12
A	F1000	F1009	F1022	F1033	F1041	F1051	F1060	F1070	F1083	F1095	F1105	F1123
B	F1001	F1010	F1023	F1034	F1042	F1052	F1061	F1071	F1084	F1096	F1111	F1125
C	F1002*	F1011	F1024	F1035	F1044	F1054	F1062	F1072	F1085	F1097	F1112	F1126
D	F1003	F1012	F1026	F1036	F1045	F1055	F1063	F1073	F1086	F1098	F1113	F1127
E	F1004	F1014	F1027	F1037	F1046	F1056	F1065	F1074	F1087	F1099	F1114	F1128
F	F1005	F1017	F1028	F1038	F1047	F1057	F1067	F1075	F1088	F1100	F1115	F1130
G	F1006	F1018	F1029	F1039	F1049	F1058	F1068	F1077	F1089	F1101	F1116	F1131
H	F1007	F1021	F1030	F1040	F1050	F1059	F1069	F1079	F1093	F1103	F1121	F1132

P14

	1	2	3	4	5	6	7	8	9	10	11	12
A	F1133	F1146	F1158	F1172	F1184	F1197	F1208	F1221	F1232	F1249	F1269	F1288
B	F1134	F1148	F1159	F1173	F1185	F1201	F1209	F1224	F1234	F1250	F1270	F1289
C	F1135	F1149	F1160	F1174	F1187	F1202	F1212	F1226	F1236	F1251	F1271	F1293
D	F1137	F1150	F1161	F1175	F1190	F1203	F1213	F1227	F1239	F1254	F1272	F1296
E	F1138	F1151	F1163	F1177	F1191	F1204	F1214	F1228	F1242	F1256	F1275	F1300
F	F1141	F1152	F1167	F1179	F1193	F1205	F1215	F1229	F1244	F1259	F1281	F1301
G	F1143	F1154	F1168	F1181	F1194	F1206	F1216	F1230	F1247	F1261	F1282	F1305
H	F1145	F1155	F1170	F1183	F1195	F1207	F1220	F1231	F1248	F1263	F1287	F1309

P15

	1	2	3	4	5	6	7	8	9	10	11	12
A	F1312	F1327	F1344	F1363	F1379	F1391	F1407	F1420	F1430	F1447	F1456	F1480
B	F1313	F1331	F1346	F1365	F1382	F1392	F1409	F1422	F1431	F1448	F1460	F1483
C	F1314	F1332	F1347	F1368	F1383	F1396	F1410	F1423	F1438	F1449	F1472	F1485
D	F1316	F1334	F1352	F1369	F1384	F1402	F1411	F1424	F1439	F1450	F1473	F1486
E	F1319	F1335	F1354	F1375	F1386	F1403	F1412	F1425	F1440	F1451	F1474	F1487
F	F1323	F1338	F1355	F1376	F1387	F1404	F1414	F1426	F1441	F1452	F1477	F1488
G	F1324	F1341	F1357	F1377	F1388	F1405	F1416	F1428	F1442	F1453	F1478	F1491
H	F1325	F1342	F1358	F1378	F1390	F1406	F1419	F1429	F1446	F1455	F1479	F1492

P16

	1	2	3	4	5	6	7	8	9	10	11	12
A	F1493	F1505	F1515	F1527	F1540	F1552	F1562	F1581	F1593	F1606	F1621	F1633
B	F1495	F1506	F1518	F1528	F1541	F1553	F1563	F1582	F1594	F1609	F1624	F1634
C	F1496	F1508	F1521	F1530	F1545	F1555	F1564	F1583	F1595	F1610	F1625	F1635
D	F1497	F1509	F1522	F1532	F1546	F1556	F1565	F1586	F1599	F1611	F1627	F1636
E	F1498	F1510	F1523	F1533	F1547	F1558	F1566	F1588	F1600	F1612	F1629	F1637
F	F1500	F1511	F1524	F1534	F1548	F1559	F1567	F1590	F1603	F1613	F1630	F1639
G	F1501	F1512	F1525	F1535	F1550	F1560	F1569	F1591	F1604	F1615	F1631	F1640
H	F1503	F1514	F1526	F1538	F1551	F1561	F1574	F1592	F1605	F1618	F1632	F1641

P17

	1	2	3	4	5	6	7	8	9	10	11	12
A	F1642	F1652	F1666	F1680	F1689	F1701	F1712	F1727	F1737	F1749	F1758	F1770
B	F1644	F1653	F1669	F1681	F1690	F1702	F1713	F1728	F1738	F1750	F1760	F1771
C	F1645	F1654	F1671	F1683	F1691	F1703	F1715	F1729	F1739	F1752	F1763	F1772
D	F1646	F1655	F1672	F1684	F1695	F1704	F1718	F1731	F1741	F1753	F1764	F1773
E	F1647	F1660	F1673	F1685	F1696	F1707	F1719	F1732	F1743	F1754	F1765	F1775
F	F1648	F1661	F1674	F1686	F1697	F1708	F1723	F1733	F1745	F1755	F1767	F1776
G	F1650	F1662	F1676	F1687	F1698	F1709	F1725	F1734	F1746	F1756	F1768	F1779
H	F1651	F1664	F1677	F1688	F1699	F1711	F1726	F1735	F1748	F1757	F1769	F1780

P18

	1	2	3	4	5	6	7	8	9	10	11	12
A	F1781	F1792	F1804	F1817	F1834	F1849	F1870	F1880	F1892	F1901	F1912	F1924
B	F1783	F1793	F1806	F1819	F1837	F1854	F1871	F1881	F1893	F1902	F1913	F1925
C	F1784	F1796	F1808	F1820	F1839	F1857	F1872	F1882	F1894	F1903	F1914	F1926
D	F1787	F1797	F1810	F1822	F1840	F1861	F1873	F1883	F1896	F1905	F1916	F1927
E	F1788	F1798	F1812	F1823	F1843	F1864	F1874	F1884	F1897	F1906	F1919	F1928
F	F1789	F1800	F1813	F1825	F1845	F1866	F1875	F1886	F1898	F1908	F1920	F1930
G	F1790	F1801	F1814	F1827	F1846	F1867	F1876	F1887	F1899	F1909	F1921	F1931
H	F1791	F1802	F1816	F1829	F1848	F1869	F1878	F1890	F1900	F1911	F1923	F1932

P19

	1	2	3	4	5	6	7	8	9	10	11	12
A	F1934	F1944	F1955	F1970	F1982	F1995	F2009	F2018	F2032	F2041	F2052	F2071
B	F1936	F1946	F1956	F1973	F1983	F1996	F2010	F2022	F2033	F2042	F2056	F2072
C	F1937	F1947	F1957	F1975	F1984	F1997	F2011	F2024	F2034	F2043	F2057	F2073
D	F1938	F1949	F1959	F1976	F1985	F2002	F2013	F2026	F2036	F2044	F2058	F2075
E	F1939	F1950	F1963	F1977	F1987	F2004	F2014	F2027	F2037	F2045	F2061	F2078
F	F1940	F1951	F1964	F1979	F1988	F2005	F2015	F2028	F2038	F2048	F2063	F2079
G	F1941	F1952	F1967	F1980	F1991	F2006	F2016	F2029	F2039	F2049	F2066	F2080
H	F1943	F1954	F1969	F1981	F1994	F2008	F2017	F2031	F2040	F2051	F2067	F2082

P20

	1	2	3	4	5	6	7	8	9	10	11	12
A	F2083	F2100	F2122	F2134	F2149	F2159	F2177	F2189	F2202	F2213	F2224	F2237
B	F2084	F2103	F2123	F2136	F2150	F2160	F2178	F2191	F2203	F2214	F2225	F2238
C	F2087	F2109	F2124	F2137	F2151	F2161	F2179	F2193	F2204	F2215	F2227	F2240
D	F2090	F2110	F2125	F2139	F2152	F2163	F2180	F2195	F2205	F2216	F2228	F2244
E	F2093	F2112	F2126	F2143	F2154	F2164	F2181	F2196	F2207	F2219	F2229	F2245
F	F2095	F2114	F2130	F2144	F2156	F2165	F2184	F2198	F2209	F2220	F2231	F2247
G	F2096	F2116	F2131	F2146	F2157	F2166	F2186	F2199	F2210	F2221	F2233	F2248
H	F2099	F2118	F2133	F2148	F2158	F2167	F2187	F2200	F2212	F2222	F2235	F2249

P21

	1	2	3	4	5	6	7	8	9	10	11	12
A	F2250	F2264	F2277	F2288	F2302	F2315	F2326	C52	C106	C193	C246	C303
B	F2253	F2265	F2279	F2289	F2303	F2317	F2328	C63	C112	C201	C243	C305
C	F2254	F2268	F2280	F2291	F2305	F2319	F2332	C67	C155	C203	C248	C306
D	F2257	F2271	F2281	F2293	F2306	F2321	F2333	C75	C160	C206	C249*	C309
E	F2258	F2273	F2282	F2294	F2308	F2322	F2335	C93	C167	C214	C251	C311
F	F2260	F2274	F2284	F2296	F2310	F2323	C9	C95	C175	C215	C269	C314
G	F2261	F2275	F2286	F2298	F2312	F2324	C20	C103	C177	C226	C281	C320
H	F2263	F2276	F2287	F2301	F2313	F2325	C28	C105	C186	C236	C295	C323

P22

	1	2	3	4	5	6	7	8	9	10	11	12
A	C339	C389	F57	F174	F380	F441	F459	F491	F501	F701	F869	F894
B	C342	C422	F67	F304	F382	F450	F460	F492	F546	F709	F870	F895
C	C347	C428	F88	F368	F393	F451	F464	F495	F577	F712	F873	F911
D	C348	C452	F112	F369	F412	F452	F467	F496	F590	F747	F875	F914
E	C352	C469	F121	F370	F413	F453	F479	F497	F591	F775	F876	F916
F	C362	C471	F128*	F371	F419	F454	F482	F498	F619	F854	F878	F924
G	C370	F10	F146	F377	F438	F455	F483	F499*	F670	F867	F882	F928
H	C388	F49*	F156	F378	F440	F456	F489	F500	F677	F868	F890	F940

P23

		2	3	4	5	6	7	8	9	10	11	12
A	F942	F983	F1189	F1255	F1285	F1337	F1370	F1357	F1570	F2115	C472	C507
B	F946	F1015	F1192	F1257	F1286	F1345	F1373	F1359	F1659	F2141	C484	C508
C	F948	F1048	F1199	F1258	F1294	F1348	F1399	F1465	F1692	C13	C493	C517
D	F951	F1064	F1210	F1260	F1298	F1351	F1401	F1466	F1735	C25	C494	C519
E	F960	F1091	F1235	F1264	F1303	F1353	F1427	F1469	F1751	C61	C498	C520
F	F967	F1107	F1237	F1266	F1318	F1356	F1433	F1502	F1958	C62	C499	C523
G	F969	F1117	F1241	F1271	F1322	F1362	F1434	F1520	F2047	C98	C503	C524
H	F982	F1171	F1252	F1272	F1326	F1367	F1437	F1568	F2069	C439	C506	C525

P24

	1	2	3	4	5	6	7	8	9	10	11	12
A	C526	C538	C554	C573	C583	C596	F2337	F2345	F2354	F2362	F2370	F2378
B	C528	C539	C558	C574	C584	C598	F2338	F2346	F2355	F2363	F2371	F2379
C	C529	C545	C559	C575	C585	C599	F2339	F2347	F2356	F2364	F2372	F2380
D	C532	C547	C562	C576	C588	C601	F2340	F2348	F2357	F2365	F2373	F2381
E	C533	C549	C565	C578	C589	C603	F2341	F2350	F2358	F2366	F2374	F2382
F	C534	C551	C569	C580	C590	C604	F2342	F2351	F2359	F2367	F2375	F2383
G	C535	C552	C570	C581	C591	I73	F2343	F2352	F2360	F2368	F2376	F2384
H	C536	C553	C572	C582	C595	F2336	F2344	F2353	F2361	F2369	F2377	F2385

P25

	1	2	3	4	5	6	7	8	9	10	11	12
A	F2386	F2396	F2404	F2412	F2421	F2431	F2443	F2456	F2467	F2483	F2492	
B	F2387	F2397	F2405	F2413	F2423	F2432	F2444	F2458	F2469	F2484	F2493	
C	F2388	F2398	F2406	F2414	F2424	F2433	F2446	F2461	F2471	F2485		
D	F2390	F2399	F2407	F2415	F2425	F2434	F2449	F2462	F2475	F2486		
E	F2391	F2400	F2408	F2416	F2427	F2437	F2450	F2463	F2477	F2487		
F	F2392	F2401	F2409	F2417	F2428	F2439	F2453	F2464	F2479	F2488		
G	F2394	F2402	F2410	F2418	F2429	F2440	F2454	F2465	F2480	F2490		
H	F2395	F2403	F2411	F2419	F2430	F2441	F2455	F2466	F2482	F2491		

