

# ÍNDICE

## RESUMEN

<b>1. INTRODUCCIÓN</b>	<b>4</b>
1.1. Origen e importancia del arroz	4
1.2. Características del arroz	4
1.3. Control de la floración en <i>Oryza sativa</i>	5
1.4. Variación natural de los genes responsables de la floración	7
1.4.1. Genes principales	7
1.4.2. Otros genes de la ruta	8
1.5. El EcoTILLING como método de estudio	9
<b>2. OBJETIVOS</b>	<b>11</b>
<b>3. MATERIALES Y MÉTODOS</b>	<b>12</b>
3.1. Extracción de ADN genómico	12
3.2. Cuantificación de la expresión	12
3.3. El EcoTILLING:	13
3.3.1. Diseño de cebadores	13
3.3.2. Reacciones de PCR	13
3.3.3. Visualización de las amplificaciones por electroforesis	13
3.3.4. Formación de heterodúplex	14
3.3.5. Digestión con endonucleasa CEL1 de hinojo	14
3.3.6. Visualización de la digestión en gel de agarosa	14
3.4. Secuenciación de ADN	15
3.5. Análisis de las secuencias	15
3.6. Toma de datos de longitud del ciclo	15
<b>4. RESULTADOS</b>	<b>16</b>
4.1. Colección de variedades	16
4.2. Caracterización del tiempo de floración	16
4.3. Análisis genético:	17
4.3.1. <i>Hd3a</i> : Nivel de expresión de <i>Hd3a</i>	17
EcoTILLING de <i>Hd3a</i>	18
Relación alelo <i>Hd3a</i> -floración	20
4.3.2. <i>RFT1</i> : Nivel de expresión de <i>RFT1</i>	20
EcoTILLING de <i>RFT1</i>	21
Relación alelo <i>RFT1</i> -floración	22
4.4. Relación <i>Hd3a</i> - <i>RFT1</i>	23
<b>5. DISCUSIÓN</b>	<b>24</b>
<b>6. CONCLUSIONES</b>	<b>28</b>
<b>7. BIBLIOGRAFÍA</b>	<b>29</b>
<b>ANEXOS</b>	
Anexo I: tabla resumen	31
Anexo II: <i>Hd3a</i>	32
Anexo III: <i>RFT1</i>	37

# ABREVIACIONES

<b>ADN</b>	Ácido Desoxirribonucleico
<b>ARN</b>	Ácido Ribonucleico
<b>ARNi</b>	Ácido Ribonucleico de interferencia
<b>ARNm</b>	Ácido Ribonucleico mensajero
<b>DDS</b>	Días Después de la Siembra
<b>dNTPs</b>	desoxirribonucleótidos
<b>INDEL</b>	Inserción y/o Delección
<b>kb</b>	kilobase
<b>LD</b>	Long Day (día largo)
<b>nm</b>	nanometro
<b>pb</b>	pares de bases
<b>PCR</b>	Polymerase Chain Reaction
<b>rpm</b>	revoluciones por minuto
<b>SD</b>	Short Day (día corto)
<b>SNP</b>	Single Nucleotide Polymorphism
<b>Tm</b>	melting temperature
<b>UV</b>	Ultra-violeta

# RESUMEN

El arroz (*Oryza sativa*) es una planta de origen tropical que depende de fotoperiodo para florecer. En su región nativa, la evocación floral se produce de forma natural en condiciones de día corto, es decir, florece cuando la longitud de los días es inferior a 12 horas. Durante su adaptación a otros climas, principalmente forzado por el hombre, el arroz ha conseguido modular su sensibilidad al fotoperiodo, consiguiendo crecer y reproducirse en latitudes dónde la floración sobreviene en verano (fotoperiodo de día largo).

Este trabajo se centra en la comprensión de los mecanismos genéticos que controlan dicha evocación, particularmente en los genes del final de la ruta de señalización de la floración en arroz. El objetivo es conocer los elementos genéticos que modifican el tiempo de floración en cultivares de tipo *japónica* adaptados a la cuenca mediterránea, dónde florece en condiciones de día largo. Se estudiarán las regiones promotoras de *RFT1* y *Hd3a*, que codifican señales de floración móviles únicas, responsables del cambio de fase. Se estudiará igualmente los niveles de expresión de los dos genes.

El análisis genético se realizará mediante EcoTILLING. El EcoTILLING aprovecha la capacidad de la endonucleasa de hinojo CEL1 para cortar la doble hebra de ADN cuando detecta un desapareamiento de bases. Las variedades se compararán de dos en dos, y se analizará la presencia de SNPs por medio de electroforesis con geles de agarosa, a tenor de los patrones de bandas que presenten, asociables a alelos.

Las variaciones genéticas encontradas se contrastarán con los tiempos de floración medidos en un ensayo de campo para caracterizar el fenotipo de floración de todas las variedades del estudio.

Con todo ello se obtendrá información acerca de la implicación de *Hd3a* y *RFT1* en la inducción de la floración en fotoperiodo largo a través de sus niveles de expresión.

# INTRODUCCIÓN

## 1.1. ORIGEN E IMPORTANCIA DEL ARROZ:

El arroz es uno de los cultivos para consumo humano más importante del planeta. Hay pruebas arqueológicas que demuestran que ya se consumía hacia el 11.000 aC (Mannion, 1999), aunque su domesticación no se produjo hasta el 6.000 aC (Fuller *et al.*, 2009) en la cuenca del río Yangtze, actual China.

Es la base de la alimentación para un tercio de la población mundial (Khush, 1997). Se consume principalmente en Asia, dónde puede representar hasta el 75% del aporte calórico en la dieta de ciertos países. Sin embargo se cultiva y consume en todo el planeta, lo que supone que *Oryza sativa* se ha adaptado a diversos climas, desde los trópicos hasta una latitud límite de 45°N (Izawa *et al.*, 2007). Esta distribución del cultivo se debe a la fuerte presión de selección que sufrió el arroz durante su expansión, acompañando al hombre en sus migraciones. Los datos más recientes apuntan a que la domesticación de *O. sativa* se produjo en dos eventos separados en China, de los que proceden las dos subespecies más cultivadas: *japónica* e *índica* (Xu *et al.*, 2011). Las variedades de tipo *índica* parecen provenir fundamentalmente de *O. nirvana* y están adaptadas a climas tropicales; siendo cultivadas en Asia y Sudamérica. En el tipo *japónica* hay una mayor contribución de *O. rufipogon* y se expandió en zonas templadas; cultivándose en Japón, el mediterráneo y Norteamérica.

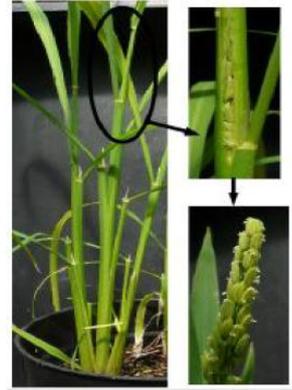
En la cuenca mediterránea el arroz se introdujo en el siglo IV aC, al establecerse las rutas comerciales con Asia. Aquí se ha ido adaptando a las características del clima mediterráneo, con veranos cálidos de días largos. La adaptación al clima templado ha originado un tipo de arroz *japónica* particular, denominado *japónica* templado.

## 1.2. CARACTERÍSTICAS DEL ARROZ:

*Oryza sativa* es una especie de día corto, es decir, es dependiente del fotoperiodo y sus condiciones de floración óptimas se dan en días cortos. Las variedades de tipo *índica* están adaptadas a climas tropicales, dónde florecen bajo fotoperiodo de día corto. Para evitar el frío del invierno, el arroz de tipo *japónica*, ha tenido que adaptar su floración a fotoperiodos de día largo típicos del verano.

El ciclo biológico de *Oryza sativa* se divide en tres fases: vegetativa, reproductiva y de maduración del grano. Las variedades cultivadas en la región mediterránea presentan un ciclo que oscila entre los 120 y los 145 días, hasta la floración. La germinación no ocurre por debajo de los 12 °C, por lo que en la península ibérica no se siembra hasta la primavera. En la fase vegetativa se establece la arquitectura de la planta, ramificándose hasta presentar un aspecto cespitoso.

La fase reproductiva se inicia con la transición de los meristemos apicales de estado vegetativo a reproductivo. El cambio de estado sólo es visible cuando la inflorescencia comienza a aumentar su tamaño y ocasiona la apertura de la vaina de la hoja en la que se halla envuelta, llamada “hoja bandera”. En el arroz cultivado, la cleistogamia es habitual: el polen fecunda el ovario antes de la antesis. Este fenómeno favorece que las variedades tradicionales se mantengan como líneas puras y tiendan a ser homocigotas en todos sus loci; aspecto que resulta muy útil para el uso de la técnica EcoTILLING planteado en este trabajo.



**Figura 1: Visualización del cambio de fase de vegetativa a reproductiva en arroz:**

En zona resaltada se muestra la hoja bandera, dónde está creciendo la panícula. El detalle superior muestra la apertura de la hoja bandera forzada por el crecimiento de la panícula.

La fase de maduración del grano comienza con la fecundación y finaliza en la cosecha, cuando el grano ha completado su engorde y se ha secado. Esta fase depende mucho de las condiciones ambientales, pudiendo alargarse con lluvias y frío. En nuestra latitud (39° 28'N), la fecha de recogida coincide con el final del verano. Las condiciones climatológicas de este periodo se caracterizan por lluvias torrenciales que pueden provocar el encamado del arroz. La reducción de la producción debida a este fenómeno puede llegar a ser considerable. Para evitar cosechar a finales del verano se pueden plantear varias estrategias, una de ellas consistiría en adelantar la siembra, pero las temperaturas mínimas previas a marzo, por debajo de los 12°C, imposibilitan la germinación y, por consiguiente, un adelanto de la siembra. Otra alternativa consistiría en el acortamiento del ciclo vegetativo, lo que implica un adelanto de la floración.

El entendimiento de la regulación de la floración ha sido objeto de numerosos estudios y ya se conocen gran parte de los mecanismos fisiológicos y moleculares que la controlan (Izawa, 2007). Se han caracterizado varios genes implicados en este proceso, que establecen las herramientas con las que trabajar para modificar la longitud del ciclo.

### 1.3. CONTROL DE LA FLORACIÓN EN *ORYZA SATIVA*:

En las últimas décadas, estudios sobre la planta modelo *Arabidopsis thaliana*, que florece en día largo, han permitido esclarecer la ruta que controla la floración en plantas dependientes de fotoperiodo (Baurle y Dean, 2006; Imaizumi y Kay, 2006).

Se han identificado tres genes clave que controlan este proceso: *GIGANTEA (GI)* (Mizoguchi *et al.* 2005), *CONSTANS (CO)* y *FLOWERING LOCUS T (FT)*. El primero de ellos está íntimamente relacionado con el reloj circadiano y la señalización mediada por luz azul y roja. A través de este gen la planta integra la información sobre la longitud del día para activar *CO* llegado el momento. *CO* codifica un factor de transcripción de tipo dedo de zinc (Samach *et al.*, 2000), que induce la transcripción de *FT*. *FT* codifica una proteína móvil que migra desde las hojas a través del floema e induce el cambio de estado del meristemo (Corbesier *et al.*, 2007).

Esta ruta de señalización está conservada en arroz (*OsGI-Hd1-Hd3a*), aunque participan elementos reguladores adicionales, propios de la especie. *OsGI* actúa sobre *Hd1* dependiendo de señales luminosas y del reloj circadiano. *Hd1* es el ortólogo de *CO* (Yano *et al.*, 2000) y su acción depende del fotoperiodo: en día corto induce la expresión de *Hd3a*, ortólogo de *FT*. En día largo, *HD1* coincide en el tiempo con la acumulación de los fitocromos, lo que lleva a un cambio en su conformación, haciendo de él un represor de *Hd3a* (Hayama *et al.*, 2003; Andrés *et al.*, 2009) y retrasando la floración.

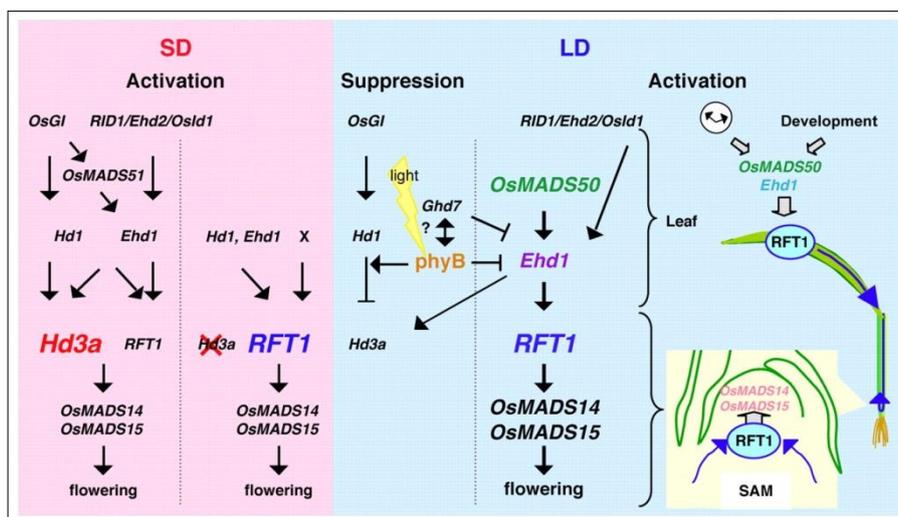
*Hd3a* está también bajo control de *Ehd1*, un regulador de respuesta de tipo B independiente de *Hd1* y sin ortólogo conocido en *Arabidopsis* (Doi *et al.*, 2004). *Ehd1* actúa siempre como un activador, aunque su expresión aumenta en fotoperiodos cortos. Por tanto, *Hd1* y *Ehd1* actúan de forma

antagónica en día largo, impidiendo la floración; mientras que en fotoperiodo de día corto actúan conjuntamente para promover la expresión de *Hd3a* y la floración. Además el arroz posee un represor específico de día largo, *Ghd7*, que actúa modulando a *Ehd1* (Xue *et al.*, 2008). La actuación conjunta de estos tres genes modula la respuesta del arroz a la longitud del día, siendo inhibida la floración en día largo.

Sin embargo, las variedades adaptadas a climas templados en latitudes norte florecen en fotoperiodo largo. Recientemente se han realizado diversos estudios para comprender en qué consistió esta adaptación. Una de las principales pistas se obtuvo al observar que plantas en las que el gen *Hd3a* se encontraba silenciado no veían comprometida su floración (Komiya *et al.*, 2008). En estas plantas, un miembro de la familia *FT-like*, *RFT1* (*Rice Flowering locus T1*), cumplimentaba la función perdida por la falta de *Hd3a*, activando los mismos genes aguas abajo y desencadenando la floración, aunque ligeramente retrasada. *RFT1* es el homólogo más cercano a *Hd3a* de los 13 miembros de la familia *FT-like* de arroz, con un 91% de similitud (Kojima *et al.*, 2002; Hagiwara *et al.*, 2009). Se encuentra a 11,5 kb aguas arriba de *Hd3a*, sugiriendo que su origen sea una duplicación en tándem de *Hd3a* y su posterior diversificación.

Este gen complementa la acción de *Hd3a*, siendo poco relevante en la promoción de la floración bajo fotoperiodos de día corto. Sin embargo, en fotoperiodos largos donde *Hd1*, fuertemente activado por *OsGI*, reduce drásticamente la expresión de *Hd3a* (Hayama *et al.*, 2003), *RFT1* toma relevancia. En estas condiciones *Ehd1* pasa a ser el principal activador de la ruta, inducido aguas arriba por *OsMADS50*. *OsMADS50* no está activo en día corto y establece una ruta de activación por fotoperiodo independiente de *OsGI* (Lee *et al.*, 2004). El papel de *RFT1* se completa con la selección de alelos débiles o nulos de *Ghd7* (Xue *et al.*, 2008) y *Hd1* (Naranjo, TFC, 2010) durante los procesos adaptativos a fotoperiodos largos.

A pesar de tener un esbozo de los cambios genéticos que permiten al arroz florecer en día largo, quedan aún muchos puntos por aclarar. Lo que parece estar claro es que de los 13 miembros de la familia *FT-like*, el papel de “florigeno” sólo lo representan *Hd3a* y *RFT1* en arroz, puesto que un doble mutante de ARNi es incapaz de florecer tanto en día largo como corto (Komiya *et al.*, 2008).



**Figura 2: Modelo de la ruta de control de la floración en arroz propuesto por Komiya *et al.* :**

Bajo condiciones de día corto *Hd1* y *Ehd1* regulan positivamente a *Hd3a* para iniciar la floración. Si *Hd3a* está silenciado (Knockout por ARNi), *RFT1* es inducido más tardíamente para desencadenar la floración.

En condiciones de día largo *RFT1* es activado en las hojas por *OsMADS50* y *Ehd1*. *Ehd1* es a su vez inhibido por otros factores. (Komiya *et al.*, 2009)

## 1.4. VARIACIÓN NATURAL DE LOS GENES RESPONSABLES DE LA FLORACIÓN

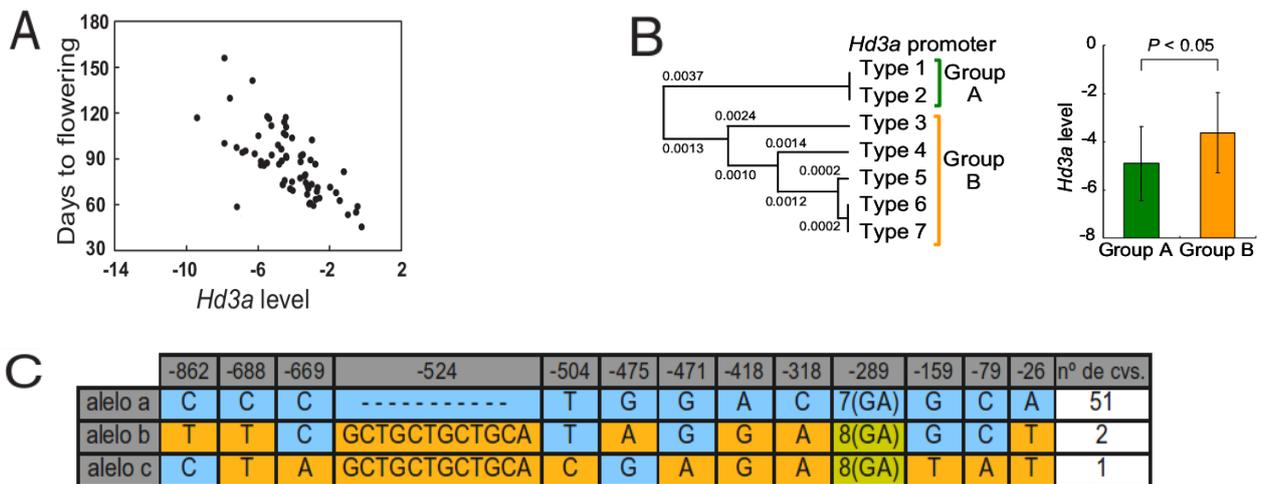
### 1.4.1. Genes principales:

*Hd3a* es el florígeno más estudiado. En un trabajo realizado previamente se analizaron los niveles de expresión de *Hd3a* en 64 cultivares de tipo *índica* (Takahashi *et al.*, 2009). En día largo los niveles de *Hd3a* fueron inapreciables, mientras que en día corto se observaron niveles elevados de mRNA, con un pico de expresión aproximadamente 30 días antes de la floración. Estos niveles de expresión estaban negativamente relacionados con la variación en el tiempo de floración: a mayor expresión menor tiempo de floración.

El análisis de la secuencia codificante reveló que la población poseía seis alelos de *Hd3a* que codificaban para seis proteínas cuya función estaba altamente conservada. Por tanto, estas variaciones en las secuencias no son las responsables de la variación del tiempo de floración. Se concluyó entonces que el tiempo de floración viene determinado por los niveles de expresión de *Hd3a*.

Para comprender a qué estaban asociadas estas variaciones de expresión se analizaron 2 kb de la región promotora del gen, dónde esperaban localizar las principales regiones reguladoras de la expresión. Se identificaron 7 tipos de secuencias que se dividieron en dos grupos de expresión: el grupo A compuesto de dos tipos de secuencias, presenta una expresión reducida frente a los alelos del grupo B. Como era de esperar, las plantas del grupo A florecieron de forma más tardía que las del grupo B.

En el Trabajo Fin de Carrera que realicé en el IVIA en 2010, analicé 1 kb de la región promotora de *Hd3a* en una colección de variedades *japónica*, encontrando 3 alelos en la población (Figura 3C). El alelo mayoritario, llamado a, coincidía con el tipo de región promotora 1 del estudio de Takahashi, incluyéndose en el grupo A. Los dos alelos restantes pertenecían al Grupo B. (Figura 3)



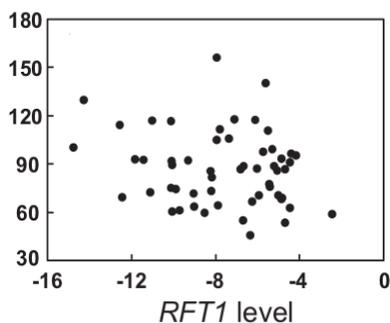
**Figura 3: Variación natural entre una población de arroz tipo *índica* del gen *Hd3a*.**

A) Correlación entre el tiempo de floración y los niveles de mRNA de *Hd3a*. La expresión se determinó a partir de hojas de 35 días de edad. Se aprecia la correlación negativa existente. (tomado de Takahashi *et al.*, 2009)

B) Agrupación de los 7 tipos de secuencias promotoras de *Hd3a* en función de sus niveles de expresión. Las regiones del grupo A presentan una mutación que dificulta la unión de activadores de transcripción.

C) Variaciones genéticas halladas en el estudio de 1 kb de la región promotora de *Hd3a* sobre variedades de tipo *japónica*. (Naranjo, TFC, 2010)

*RFT1* parece ser el principal responsable de la floración en arroz en condiciones de día largo. El estudio de su expresión en cultivares de tipo *índica* (Takahashi *et al.*, 2009) halló una correlación muy baja entre el nivel de expresión y el tiempo de floración en día corto. Sin embargo, trabajos posteriores realizados sobre individuos de floración en día largo determinaron que tiene un papel fundamental en estas condiciones (Komiya *et al.*, 2009).



**Figura 4:** Niveles de expresión de *RFT1* en relación al tiempo de floración en fotoperiodo de día corto. No se aprecia la existencia de correlación. (tomado de Takahashi *et al.*, 2009)

Teniendo estos conocimientos previos, en este trabajo se pretende ahondar en el conocimiento de las variaciones naturales de estos genes y de sus niveles de expresión en los cultivares empleados en la cuenca mediterránea dentro de un programa de mejora genética local.

#### **1.4.2. Otros genes de la ruta:**

Igual de importantes en la ruta de floración son los genes que controlan directamente la expresión de los florigenos. *Hdl* y *Ehd1* se encuentran en el mismo escalón de la ruta, inmediatamente anterior a *Hd3a* y *RFT1*. El equilibrio entre ambos genes es crucial para el inicio de la floración.

*Hdl* regula la expresión de *Hd3a* (Yano *et al.*, 2000). Bajo condiciones de día largo *Hdl* es un represor de *Hd3a*. *Hdl* también está regulando la expresión de *RFT1* como lo demuestra el descenso significativo del mRNA de *RFT1* en plantas con *Hdl* silenciado (Komiya *et al.*, 2008); aunque por mecanismos todavía desconocidos.

Un estudio genético realizado sobre cultivares de tipo *japónica* adaptados a latitudes norte (Naranjo, TFC, 2010), desveló que 23 de las 55 variedades analizadas presentaban alelos no funcionales de *Hdl*, debido a deleciones responsables de un cambio en el marco de lectura o bien de un codón de parada prematuro. Estas variaciones afectaban a un dominio CCT, responsable de su localización nuclear, por lo que las proteínas resultantes eran incapaces de migrar al núcleo donde realizan su acción represora. El resto de variedades poseían un dominio CCT intacto aunque tenían otras sustituciones que afectaban al dominio de unión a DNA (zinc-finger domain), lo que podría estar alterando la capacidad reguladora de HD1. No obstante no se halló correlación alguna entre la presencia de alelos funcionales de HD1 y el tiempo de floración.

*Ehd1* queda entonces como principal regulador de la ruta en estas variedades. EHD1 es una proteína altamente conservada (Takahashi *et al.*, 2009; Doi *et al.*, 2004), por lo que su influencia en el tiempo de floración deriva de sus niveles de expresión. El análisis de su región promotora permitió identificar cuatro tipos de regiones reguladoras (Naranjo, TFC, 2010). Las dos regiones más extendidas entre la población parecían mostrar diferentes niveles de expresión, no hallándose correlación entre estos y el tiempo de floración.

## 1.5. EL ECOTILLING COMO MÉTODO DE ESTUDIO

El EcoTILLING se desarrolló en 2003 a partir de una técnica llamada TILLING: Targeting Local Lesions in Genomes (McCallum *et al.*, 2000). El EcoTILLING se emplea para comparar fragmentos de ADN de diferentes individuos, normalmente variedades de un cultivo, y detectar las diferencias en la secuencia (Comai *et al.*, 2004). Para ello se emplea una endonucleasa de hinojo, llamada CEL1, que corta la doble hebra de ADN cuando detecta un error en el apareamiento de bases, debido o bien a cambios puntuales de nucleótidos (SNPs) o a pequeñas deleciones o inserciones (INDELS).

Los pasos a seguir para realizar un estudio mediante EcoTILLING consisten en:

- ♣ Definir el fragmento a estudiar y amplificar por PCR:

Se seleccionan los genes de interés en el estudio y se diseñan cebadores para la reacción de PCR que amplifiquen la región de forma exclusiva, es decir, que no anillen de forma inespecífica en una región genómica diferente a la de interés. La presencia de bandas inespecíficas en las amplificaciones por PCR dificultará en gran medida la interpretación de los resultados.

- ♣ Formación de heterodúplex:

Para que la endonucleasa CEL1 pueda detectar las diferencias en las secuencias amplificadas por PCR hay que generar dobles cadenas en las que cada cadena provenga de un individuo diferente. Para ello se desnaturaliza a 95°C una mezcla equilibrada de fragmentos de PCR de dos individuos. La temperatura se disminuirá de forma muy lenta para favorecer la unión de dos cadenas de parentales diferentes, llamados heterodúplex.

A pesar de ello, las cadenas homólogas (provenientes del mismo parental) tenderán a unirse entre sí formando homodúplex. Al final de la reacción tendremos, por tanto, una mezcla equilibrada de homodúplex y heterodúplex.

- ♣ Digestión con endonucleasa:

Se incuba la mezcla conteniendo los heterodúplex con CEL1 de hinojo. La endonucleasa actuará sobre los heterodúplex, cortando en tantos puntos como variaciones presenten las dos variedades para el fragmento en estudio.

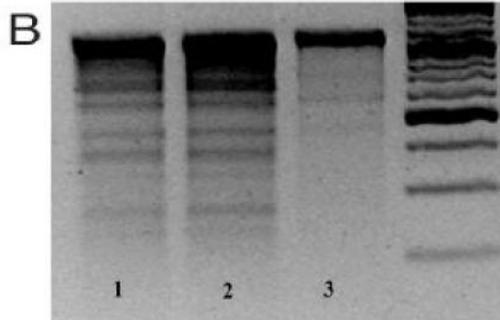
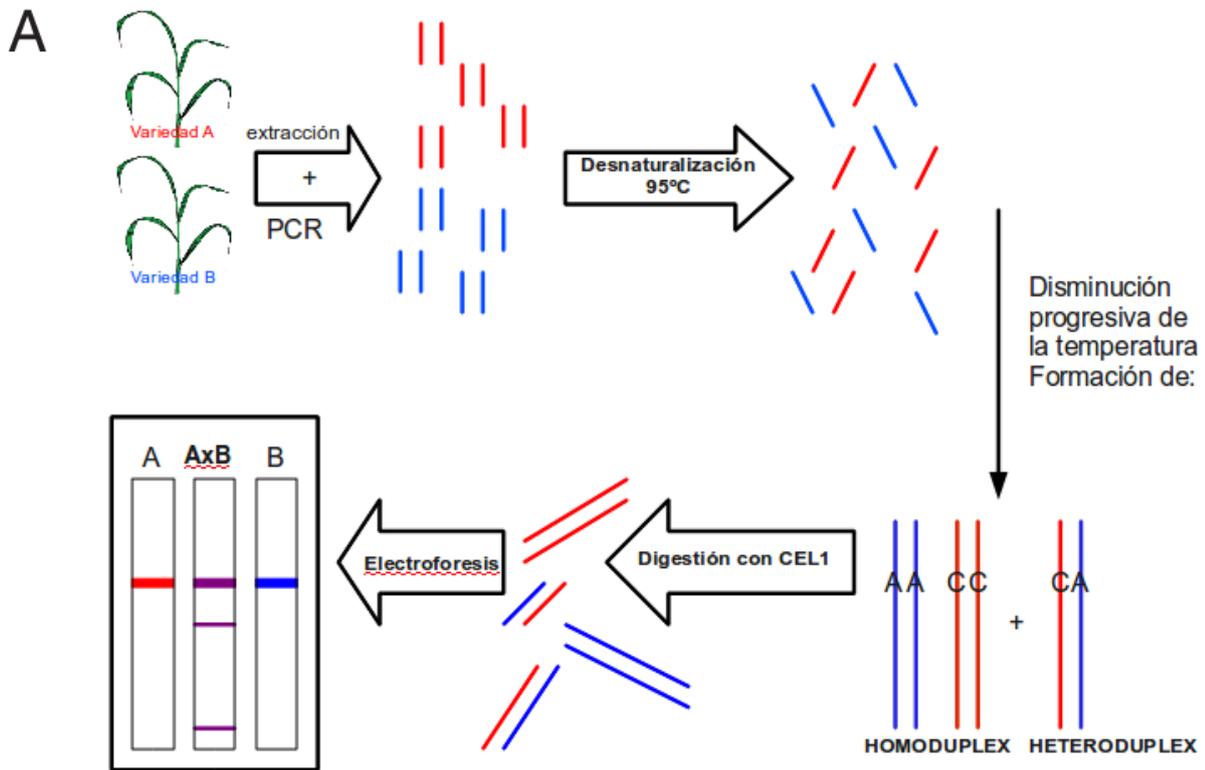
- ♣ Electroforesis de la digestión:

La digestión se visualizará en un gel de agarosa sometido a un campo eléctrico (Raghavan *et al.*, 2007). Los fragmentos se separarán en función de su tamaño, lo que permite identificar los alelos en función del número de bandas y de su tamaño.

El hecho de que el arroz sea una planta estrictamente autógama, y por tanto con tendencia a la homocigosis en todos sus loci, facilita este tipo de análisis. Si estudiáramos un fragmento en heterocigosis en alguno de los parentales no podríamos saber si las bandas generadas por CEL1 provienen de diferencias entre individuos o de heterodúplex formados a partir de los dos alelos de un mismo individuo.

Este método es rápido y poco costoso ya que sólo es necesario secuenciar un representante de cada grupo alélico. Se debe tener en cuenta que variaciones de secuencia muy próximas entre sí provocarán dos patrones de bandas indistinguibles, por lo que pueden pasar desapercibidos e identificarse erróneamente como el mismo alelo.

Este tipo de estudio puede constituir el primer paso en un programa de mejora clásica puesto que permite conocer la variación alélica con la que se cuenta para la elección de parentales con los que generar cruces. Además también facilita la selección ya que permite el rastreo de alelos y la heterocigosis de los genes de interés en las poblaciones descendientes.



**Figura 5: El EcoTILLING como método de estudio (Tomado de Naranjo, TFC, 2010)**

**A)** Pasos a seguir en un estudio por EcoTILLING.

**B)** Visualización de 3 digestiones de CEL1. Se aprecia que las dos primeras variedades presentan un patrón de bandas idéntico y diferente al de la tercera carrera. Se puede asumir que estas dos variedades poseen el mismo alelo.

# OBJETIVOS

El objetivo de este trabajo de investigación, llevado a cabo en el Centro de Genómica del Instituto Valenciano de Investigaciones Agrarias, es profundizar en el conocimiento de la variación natural de genes relevantes en el control de la floración de *Oryza sativa*, concretamente *Hd3a* y *RFT1*. El estudio se centrará en las variedades de arroz de tipo *japónica* empleadas en la cuenca mediterránea. Se pretende comprender los mecanismos que determinan la inducción de la floración en condiciones de día largo y que permitan acortar el ciclo, lo que reduciría la duración del cultivo en el campo y ayudaría así a mitigar posibles adversidades, como plagas o restricciones de agua.

Para conseguir este objetivo, se han planteado los siguientes objetivos parciales:

- 1) se analizará la variación natural de *Hd3a* y *RFT1* entre la población de arroz tipo *japónica* adaptada al clima mediterráneo. Para ello se empleará la técnica EcoTILLING. Se secuenciará ADN de cada uno de los alelos encontrados.
- 2) Paralelamente se estudiarán los niveles de expresión de ambos genes y
- 3) se estudiará la existencia de una relación entre los niveles de expresión, la variabilidad alélica encontrada y los tiempos de floración. Se realizarán análisis de correlación.

Los conocimientos aportados por este trabajo tienen como finalidad última la aplicación a un programa de mejora dirigido a acortar la fase vegetativa del arroz de tipo *japónica*.

# MATERIALES Y MÉTODOS:

## 3.1. EXTRACCIÓN DE ADN GENÓMICO:

Se cultivaron plantas a partir de semilla. Las semillas se sumergieron en agua durante 24 horas en oscuridad, para pre-hidratarlas y mejorar su germinación. Posteriormente se trasplantaron a semilleros en cámara de cultivo a 28°C y 75% de humedad relativa, fotoperiodo 12 horas de oscuridad y 12 horas de luz.

Las 57 variedades del estudio de EcoTILLING se cultivaron durante 10-12 días. Se recogieron muestras (2-3 gramos de peso fresco) para la extracción de ADN genómico que contenían la totalidad de la parte aérea de la planta (hojas y tallos), descartando los restos de la semilla y las raíces. La extracción se realizó con el kit comercial DNeasy Plant (Qiagen), siguiendo el protocolo indicado por el comerciante.

Posteriormente se cuantificó el ADN obtenido con el espectrofotómetro Nanodrop. Se descartaron las muestras con una absorbancia 260/280 inferior a 1,7.

## 3.2. CUANTIFICACIÓN DE LOS NIVELES DE EXPRESIÓN:

Se recogieron muestras de hoja de 77 días tras la siembra (DDS) cultivadas al aire libre bajo condiciones de fotoperiodo de día largo, entre los meses de marzo y mayo. Se extrajo la totalidad del ARN usando el kit RNeasy plant mini kit de Qiagen. El ARN extraído se trató con DNaseI (Invitrogen).

Se empleó una dilución a 75 ng/μl del RNA total para cuantificar la expresión con SYBR Green Master Mix (Applied Biosystem) con los siguientes cebadores (ver anexo II Y III):

<i>Hd3a</i> : D: cebador directo:	5'-GATGCACCAAGCCCAAGTGA-3'	Tm=57,9°C
C: cebador reverso:	5'-GGAACAGCACGAACACCA-3'	Tm=56,0°C
<i>RFT1</i> : a: cebador directo:	5'-ACCCTAACCTTAGGGAGTATCTACAC-3'	Tm=63,2°C
E: cebador reverso:	5'-GAAGTTCCTGGTGCTGAAGTTCT-3'	Tm=60,6°C

La reacción se llevó a cabo en el aparato Light Cycler de Roche Applied Science con las siguientes condiciones:

<i>Hd3a</i> :	activación Taq:	95°C	10 s	} 45 ciclos
	desnaturalización	95°C	8 s	
	anillamiento	59°C	8 s	
	extensión	72°C	8 s	
	melting	95°C	15 s	
		42°C	1 s	
	cooling	40°C	20 s	
<i>RFT1</i>	desnaturalización inicial:	95°C	10 s	} 45 ciclos
	desnaturalización	95°C	3 s	
	anillamiento	59°C	8 s	
	extensión	72°C	10 s	
	melting	95°C	15 s	
		42°C	1 s	
	cooling	40°C	20 s	

Se realizó una curva patrón con diluciones sucesivas de una de las variedades para cada gen. Se realizó el número de diluciones necesarias para obtener una curva patrón con un error y una eficiencia aceptables (cerca de 0 y 2, respectivamente).

### 3.3. EL ECOTILLING:

#### 3.3.1. Diseño de cebadores:

Los cebadores empleados para las PCR fueron sintetizados por Bonsai-Tech.

*Hd3a*: región promotora: 2148 pares de bases

prom 2: cebador directo: 5'-CACCCAGCTAATTGGCAGACA-3' Tm=59,8°C

promR: cebador reverso: 5'-ATAGGTGACCTTGAGGTTGGT-3' Tm=57,5°C

*RFTI*: región promotora y 5'UTR: 1500 pares de bases

H: cebador directo: 5'-AGCCTTTCTGTTTGTATTGCATC-3' Tm=54,0°C

I: cebador reverso: 5'-CGGCCATGTCAAATTAATAACCT-3' Tm=53,6°C

#### 3.3.2. Reacciones de PCR:

Para las reacciones de PCR se realizó una mezcla con:

- ▲ 150 ng ADN genómico
- ▲ 1 µl cebador directo (50 µM)
- ▲ 1 µl cebador reverso (50 µM)
- ▲ 4 µl dNTPS (0,25 µM cada uno)
- ▲ 5 µl tampón de reacción 10x
- ▲ 0,25 µl de polimerasa Taq (ExTakara) (5 U/ µl)

Se añadió agua hasta los 50 µl.

Las condiciones de las reacciones fueron:

*Hd3a* región promotora:

desnaturalización inicial	300 s	94°C	} 35 ciclos
desnaturalización	30 s	94°C	
anillamiento	30 s	56°C	
extensión	140 s	72°C	
extensión final	600 s	72°C	

*RFTI* región promotora:

desnaturalización inicial	300 s	94°C	} 35 ciclos
desnaturalización	30 s	94°C	
anillamiento	30 s	53°C	
extensión	90 s	72°C	
extensión final	600 s	72°C	

#### 3.3.3. Visualización de las amplificaciones por electroforesis:

Los fragmentos amplificados se separaron en geles de agarosa bajo un campo eléctrico. Los fragmentos se desplazan hacia el polo negativo a una velocidad que depende de su tamaño. Las bandas diferenciadas se detectan por la acción de GelRed, un agente químico intercalante que emite fluorescencia a 590 nm cuando se irradia con luz UV.

Los pasos a seguir en una electroforesis de ADN en gel de agarosa son los siguientes:

- ▲ Preparar un molde con un peine para pocillos, prestar atención a que la superficie sea totalmente horizontal para evitar zonas con un grosor del gel variable.
- ▲ Disolver 1 g de agarosa D-1 (Pronadisa) en 100 ml de TAE 1x (tris-acetato 40mM, EDTA

2mM, pH 8,0). Calentar en microondas hasta que entre en ebullición y se disuelva la agarosa por completo.

- ▲ Dejar enfriar hasta unos 65°C y añadir 1,5 µl de GelRed.
- ▲ Verter en el molde y esperar a que solidifique a temperatura ambiente (unos 30 minutos).
- ▲ Una vez sólido el gel, se retira el peine y se introduce en la cubeta de electroforesis con 1 litro de tampón TAE 1x.
- ▲ Se cargan 12 µl de cada muestra en los pocillos:
  - 10 µl de PCR
  - 2 µl de tampón de carga 6x (Ficoll 15%, azul de bromofenol 0,25%, xilencianol 0,25%).Se reserva un pocillo para el marcador de peso molecular de referencia, útil para medir el tamaño de las bandas. Se utilizó un preparado comercial de Fermentas (GeneRuler ADN Ladder) para medir bandas en el rango de 100-2000 pb.
- ▲ Someter el gel a 80V durante el tiempo necesario para la correcta separación de las bandas del marcador de peso molecular de referencia.
- ▲ Exponer a la luz UV y fotografiar.

#### 3.3.4. Formación de heterodúplex:

Se mezcló de forma equilibrada el ADN de una variedad a estudiar junto con el ADN de una variedad de referencia, de la que se conoce la secuencia de dicho fragmento:

- 8 µl ADN variedad "a"
- 8 µl ADN variedad "b"

La proporción de cada variedad se determinó en función de la eficiencia de la PCR para ese fragmento y esa variedad.

La mezcla se calentó a 95°C. La temperatura se disminuyó progresivamente para favorecer la formación de heterodúplex. Se forman así mismo homodúplex.

#### 3.3.5. Digestión con endonucleasa CEL1 de hinojo:

Se añadió a la mezcla de heterodúplex la enzima CEL1 de hinojo, junto con su correspondiente tampón:

- 1,5 µl de CEL1
- 2 µl de tampón para CEL1 (MgSO<sub>4</sub> 1M; Hepes 1M pH 7,5; KCl 2M; Triton X-100 10%; BSA 20 mg/ml)
- agua hasta 20 µl

Se incubó el conjunto a 45°C durante 50 minutos para la digestión por la enzima.

Una vez pasados los 50 minutos se detuvo la reacción mediante la adición de 5 µl de EDTA 0,225 mM.

#### 3.3.6. Visualización de la digestión en gel de agarosa:

Los geles de agarosa se prepararon siguiendo el mismo procedimiento que en el apartado 3.3.3, con 2 variaciones:

- TBE 1X (Tris; EDTA Na<sub>2</sub>; ácido bórico) reemplazando al TAE 1X
- agarosa MS-8 (Pronadisa) al 2,2%. Este tipo de agarosa es apropiada para una resolución de fragmentos de ADN inferiores a 1.200 pb.
- se utilizó el volumen completo de la reacción de digestión por CEL1:
  - 25 µl de digestión
  - 5 µl de tampón de carga 6x

El gel se sumergió en TBE 1X y se aplicó un voltaje de 100V hasta la completa separación de las bandas de ADN.

Se tomaron fotografías periódicas del gel bajo luz UV: transcurridos 40 minutos, 1 hora y 20 y 2 horas. Así se pudieron detectar bandas pequeñas que dejan de ser visibles según avanzan en el gel.

Para simplificar la interpretación de los datos se escogió Nipponbare como variedad de referencia. Se enfrentaron todas las variedades del estudio a Nipponbare. La ausencia de bandas indica que la variedad analizada comparte alelo con Nipponbare. Los patrones de bandas se clasificaron en grupos alélicos.

### 3.4. SECUENCIACIÓN DE ADN

Una vez establecidos los patrones de bandas de todas las muestras, se seleccionaron varios representantes de cada grupo alélico para su secuenciación.

El ADN de las muestras fue secuenciado por SECUGEN. El protocolo utilizado para el envío de muestras fue el siguiente:

- precipitación del ADN proveniente de PCR: añadir al volumen de la PCR 1/10V de NaAc 3M y 2,5 V de EtOH abs.
- dejar a -20°C durante al menos 2 horas
- centrifugar a 12.000 rpm durante 20 minutos
- eliminar sobrenadante y dejar secar los restos de alcohol al aire
- resuspender el pellet en 40 µl de agua y medir la concentración de ADN (Nanodrop).
- diluir la muestra hasta obtener la concentración exigida por la empresa de secuenciación. La concentración del cebador para la reacción de secuenciación fue de 10 ng/µl.

### 3.5. ANÁLISIS DE LAS SECUENCIAS DE ADN

Se visualizaron los cromatogramas mediante el programa informático Chromas2 y se eliminaron las indeterminaciones fácilmente interpretables.

La alineación de las secuencias se realizó online usando el programa ClustalW2 ([www.ebi.ac.uk/tools/clustalw2](http://www.ebi.ac.uk/tools/clustalw2)). Se comparó cada variedad del estudio con Nipponbare (*japónica*) y con IR64 (*índica*), esto permitió la clasificación de las variedades dentro de grupos alélicos.

Una vez formados los grupos se alinearon todas las secuencias incluidas dentro de cada uno y se reintrodujeron en ClustalW2 para comprobar su uniformidad.

A las alineaciones obtenidas mediante ClustalW2 se les aplicó un procesador de textos para resaltar la presencia de los SNPs.

Para determinar la localización de sitios de unión de factores de transcripción en las regiones promotoras en estudio se utilizó la base de datos PLACE (<http://www.dna.affrc.go.jp/PLACE/>) y se resaltaron los SNPs que afectaban a alguno de estos sitios.

### 3.6. TOMA DE DATOS DE LONGITUD DEL CICLO:

Se analizaron 10 plantas por variedad. Las semillas se sumergieron en agua durante 24 horas en oscuridad, para pre-hidratadas y mejorar su germinación. Posteriormente se sembraron en maceta en invernadero (39° 28'N) a una temperatura media de 24°C y 50% de humedad relativa, fotoperiodo natural durante el verano de 2011.

Como medida del tiempo de floración se calcularon los días desde que se siembra la semilla (sumergida en agua) hasta que la mitad de la primera panícula de cada planta emerge de la hoja bandera. Se realizó la media del tiempo de floración de seis plantas para establecer el dato de floración de una variedad. Como unidad de medida se estableció los días después de la siembra (DDS) en los que floreció cada variedad.

# RESULTADOS

## 4.1. COLECCIÓN DE VARIEDADES:

Las variedades seleccionadas para formar parte de este estudio fueron escogidas durante la realización del Trabajo Final de Carrera presentado por Laura Naranjo en Noviembre 2010: “Caracterización de la variación natural de genes reguladores de la floración en cultivares de tipo *japónica* de *Oryza sativa*”.

La colección consta de 52 variedades de arroz tipo *japónica* que reflejan el acervo genético representativo del arroz adaptado a fotoperiodos largos y latitudes norte. Entre estas se encuentran las variedades más cultivadas actualmente en el área mediterránea, así como variedades adaptadas a otras regiones, como Estados-Unidos o Japón. En la colección también se incluyeron dos cultivares de tipo *índica*: Kasalath, una variedad india con un tiempo de floración medio; e IR64, una variedad proveniente del International Rice Research Institute (IRRI, Filipinas) y utilizada frecuentemente como referencia en los estudios de genómica del arroz. También se añadieron tres variedades *japónica* con características destacadas en el tiempo de floración: Arros da Terra, Italiana Livorno y Pavlovsky. Estas variedades, portuguesa, italiana y rusa respectivamente, destacan por ser extremadamente precoces.

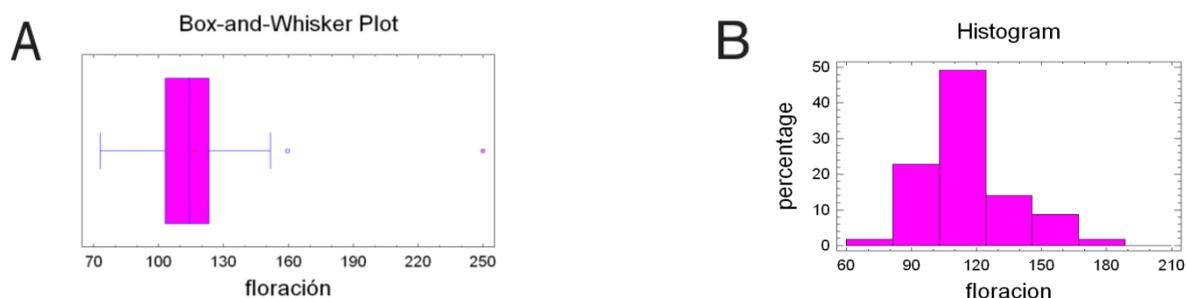
La colección se compone por tanto de 55 variedades de arroz *japónica* con tiempos de floración variables y adaptadas a diferentes fotoperiodos, y dos variedades *índicas* a modo de control ya que, debido a la gran distancia filogenética de los dos tipos de arroz, hay una mayor probabilidad de encontrar polimorfismos. Se espera que la variabilidad presente en los tiempos de floración de estas variedades pueda relacionarse con variabilidad genética en los genes seleccionados.

## 4.2. CARACTERIZACIÓN DEL TIEMPO DE FLORACIÓN:

La floración sobrevino entre los meses de Julio y Agosto de 2011. La fecha media de emergencia de panícula fue el 4 de Julio, lo que corresponde a un fotoperiodo de día largo (14h luz : 10h oscuridad).

La población floreció de manera uniforme entre los 73 DDS y los 200 DDS, con una media de 116,5 DDS y una desviación estándar de 22,3 DDS. La mayor parte de los cultivares floreció entre los 100 y los 125 DDS, distribuyéndose de forma normal (Figura 1). Los cultivares que florecieron antes de este intervalo fueron considerados tempranos y los que lo hicieron después se les consideró tardíos. Dos cultivares presentaron una floración anormalmente tardía: Agami, de origen africano, y Kinmaze, de origen asiático.

Los datos de floración se muestran en la tabla resumen del anexo I.



**Figura 1:** A) Diagrama de Box-Wishker de los datos de floración de las 57 variedades de la colección.

B) Distribución de densidad de los datos del tiempo de floración de los 57 cultivares del estudio. La población se aproxima a una distribución normal con una ligera asimetría positiva.

### 4.3. ANÁLISIS GENÉTICO:

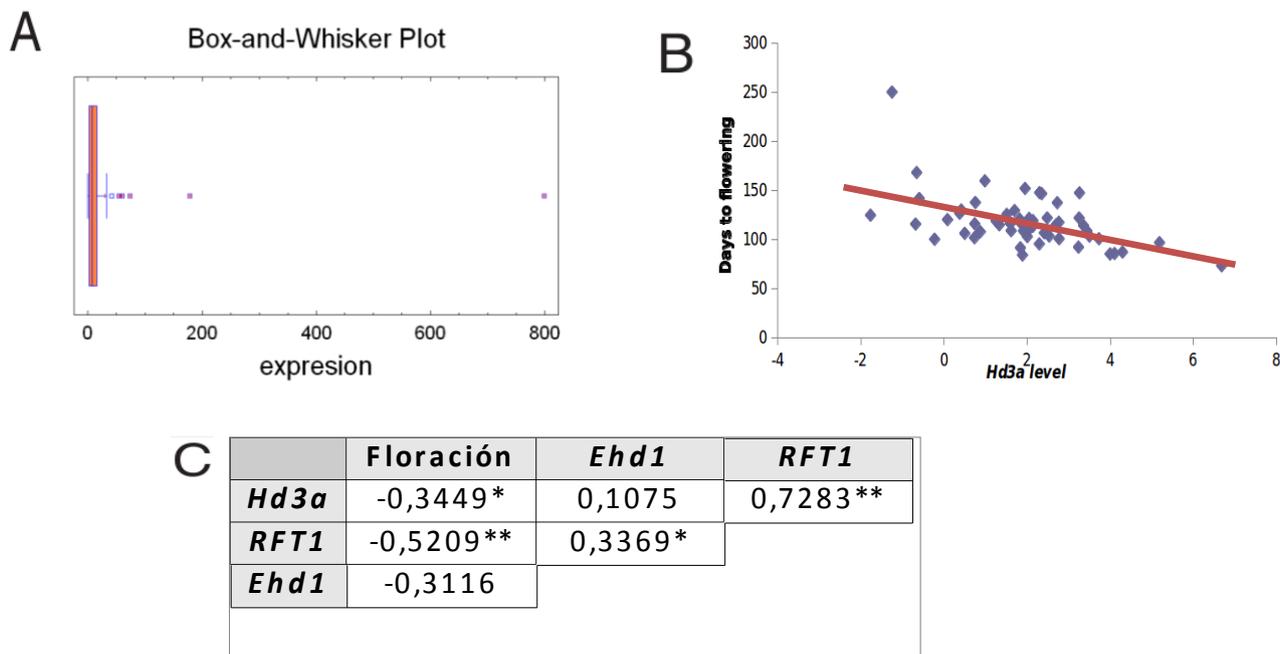
#### 4.3.1. Región promotora de *Hd3a*:

##### Nivel de expresión de *Hd3a*:

Los niveles de expresión de *Hd3a* son relativamente elevados respecto a lo esperado, teniendo en cuenta la acción represora de HD1 en fotoperiodos largos. La expresión es muy variable entre cultivares. La mayor parte de la población presenta una expresión relativa entre 0 y 50, habiendo cinco variedades con la expresión anormalmente elevada: Arroz da Terra, CRLB, Gigante Vercelli, Loto y Pavlovsky. Se considerarán anormales aquellos datos que se alejen más de dos desviaciones típicas de la media de la población. (ver Figura 2)

Las variedades que destacan claramente por un nivel de expresión muy por encima de la media presentan tiempos de floración muy tempranos. Es el caso de Arros da Terra, que presenta el mayor nivel de expresión (799) y un inicio de floración a los 73,0 DDS, esto es 43,5 días antes que la media de la población.

CRLB1 también destaca por un tiempo de floración precoz, 85,0 DDS. Aunque los niveles de expresión de *Hd3a* son más elevados que en el resto de la población (54,5) no alcanzan los niveles de Arros da Terra. Cabe indicar que la recolección de las hojas se realizó 77,0 DDS, por lo que Arros de Terra ya había florecido. Las dos variedades *índica* utilizadas mostraron unos niveles de expresión de *Hd3a* dentro del rango del resto de variedades *japónicas* analizadas.



#### **Figura 2:**

A) Diagrama de Box-Wishker de los datos de expresión de *Hd3a* a 77,0 DDS, se aprecia que el 50% de la población presenta una expresión entre 0 y 25. Unos valores entre 0 y 50 se consideran normales, sin embargo 5 variedades podrían ser consideradas anómalas puesto que su expresión se aleja demasiado de los valores medios de la población.

B) Gráfico que muestra los tiempos de floración asociados al logaritmo neperiano del nivel de expresión. Se aprecia la correlación negativa como una pendiente negativa de -0.29 en la recta de regresión.

C) Tabla de correlaciones (*Pearson's correlation coefficient*) entre niveles de expresión de los diferentes genes y el tiempo de floración. Los datos de expresión de *Ehd1* se obtuvieron en el laboratorio en un estudio anterior.

\* indica  $P < 0,01$ , \*\* indica  $P < 0,001$

Al comparar estos datos con los de floración vemos que existe una correlación negativa entre ambos factores. Aunque esta correlación es baja,  $-0,34$ , es consecuente con el fenómeno observado: a mayor expresión menor tiempo de floración. Esta correlación tan baja entre expresión de *Hd3a* y tiempo de floración no permite explicar ni predecir la longitud del ciclo en función de *Hd3a*. Hay casos en que un tiempo de floración precoz no va asociado a una elevada expresión de *Hd3a*. Es el caso de Giovanni Marchetti, que florece 84,0 DDS con unos niveles de *Hd3a* relativamente bajos: 6,62.

Estos datos indican la existencia de otro factor que regule la floración en día largo. Basándonos en estudios previos (Komiya *et al.*, 2008), *RFT1* podría explicar la variabilidad del tiempo de floración no asociado a *Hd3a*.

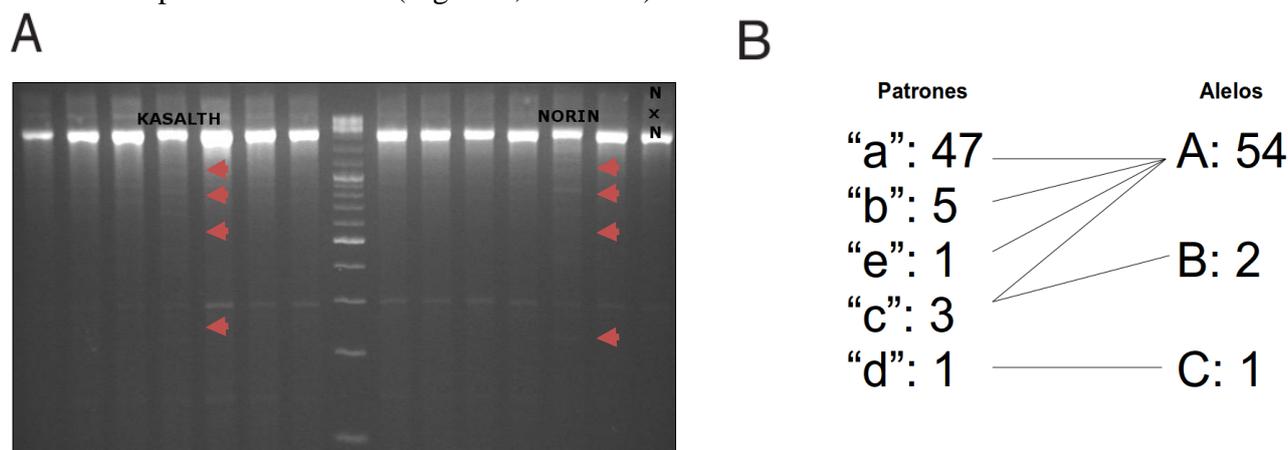
### **EcoTILLING de Hd3a:**

Se estableció la región de interés en las 2.000 pb anteriores al codón de iniciación de la transcripción. En esta región se hallan los elementos de unión a reguladores, ya descritos anteriormente (Takahashi *et al.*, 2009). En concreto hay un elemento de unión de tipo ARR1 sobre el que se han descrito variaciones en variedades de tipo *índica* y que podría ser responsable de cambios significativos en los niveles de expresión. Se localiza a  $-1.932$  pb. Los fragmentos amplificados fueron de 2.140 pb usando como cebadores los oligonucleótidos prom2 – promR (ver anexo II).

Las digestiones con CEL1 mostraron 5 patrones de bandas. La mayor parte de la población comparte el patrón con Nipponbare clasificándose en el grupo alélico “a”: 47 variedades no mostraban bandas al enfrentarlas a Nipponbare. Cinco variedades mostraron el mismo patrón, llamado “b”. El patrón “c” estaba presente en 3 variedades, mientras que los patrones “d” y “e” sólo aparecían en una variedad. (Figura 3B)

Se secuenciaron los productos de las amplificaciones por PCR de *Hd3a* con el cebador directo prom2, puesto que una variación relevante en el tiempo de floración se encuentra a  $-1932$  pb (Takahashi *et al.*, 2009). De este modo se aseguró la lectura de esta región.

Asimismo, se secuenciaron los productos de PCR de seis representantes del grupo “a” y todos los individuos del resto de grupos alélicos, puesto que los patrones de bandas eran de interpretación dudosa debido a un gran ruido de fondo. Las alineaciones de las secuencias mostraron la presencia de tres tipos de secuencias (Figura 3, Anexo II).



**Figura 3:** A) Electroforesis de las digestiones con CEL1 de los heterodúplex formados con cada variedad en estudio frente a Nipponbare. La carrera 4, Kasalath presenta un patrón de bandas similar al de la carrera 12, Norin8. La última carrera corresponde al control con ADN proveniente sólo de Nipponbare.

B) Distribución de las variedades de la colección según su patrón de bandas procedente de las digestiones de CEL1 para el gen *Hd3a* y clasificación según los alelos deducidos de las secuencias.

Los patrones de bandas “b”, y “e” resultaron deberse a ruido de fondo. El patrón de bandas “c”, con tres representantes, corresponde a la secuencia del alelo B. Un individuo clasificado en este grupo alélico, Ullal, resultó no presentar dichas variaciones, pasando a formar parte del alelo A. El patrón “d” mostró una secuencia similar al alelo B pero con algunas diferencias, se denominó como alelo C, presente únicamente en T65, una variedad taiwanesa de ciclo largo.

Todos los representantes del grupo “a”, como los de los grupos “b” y “e”, mostraron secuencias idénticas a las de Nipponbare, formando un alelo A compartido por 54 individuos.

Todos los SNPs hallados fueron anteriormente descritos por Takahashi et al., 2009. Entre nuestras variedades tenemos representantes de los dos grandes tipos de secuencias promotoras (ver Figura 3B Introducción). Nipponbare y las variedades con alelo A pertenecen a un primer grupo, de menor nivel de expresión según Takahashi. Los alelos B y C identificados en este trabajo forman parte del segundo grupo, que en fotoperiodo corto presenta mayor nivel de expresión.

Para determinar la importancia de estas variaciones en las secuencias se buscó en una base de datos online de elementos reguladores (<http://www.dna.affrc.go.jp/PLACE/>) la presencia de posibles sitios de unión factores de transcripción.

Los polimorfismos indicados en la Figura 4 coinciden con posibles sitios de unión de factores de transcripción. El cambio de una adenina por una guanina en posición -1824 podría afectar al reconocimiento de factores de transcripción similares a RAV1-A, de la familia *APETALA 2-LIKE*, relacionado con la formación de la flor. El resto de sitios de reconocimiento predichos señalan factores de transcripción que no están directamente relacionados con la floración, pero podrían indicar sitios de unión a otros factores que contengan dominios de reconocimiento de ADN similares.

La proteína EHD1 presenta un motivo de unión a ADN tipo GARP (*Golden2, Arabidopsis Response Regulator [ARR]*, y *Chlamydomonas regulatory protein of P-starvation acclimatization response [Psr1]*) en su sección media (Doi et al., 2004; Riechmann et al., 2000). Se hallaron tres sitios de unión potenciales para factores de tipo ARR1 en la región promotora de *Hd3a* que podrían ser claves en la unión de este activador. Estas regiones no presentaban variación, por lo que la activación de *Hd3a* por EHD1 no debería verse afectada por cambios en estas zonas.

## A

n	-1824	-1757	-1693	-1607	-1585	-1506	-1474	-1401	-1237	-1107	-862	-688	-669	-524	-504	-475	-471	-418	-318	-289	-159	-79	-26	n° cvs
Alelo a	A	C	C	G	C	A	C	A	G	-	C	C	C	.....	T	G	G	A	C	7(GA)	G	C	A	54
Alelo b	G	T	C	A	T	C	T	G	C	A	T	T	C	GCTGCTGCTGCA	T	A	G	G	A	8(GA)	G	C	T	2
Alelo c	G	T	T	G	C	C	T	G	C	A	C	T	A	GCTGCTGCTGCA	C	G	A	G	A	8(GA)	T	A	T	1

## B

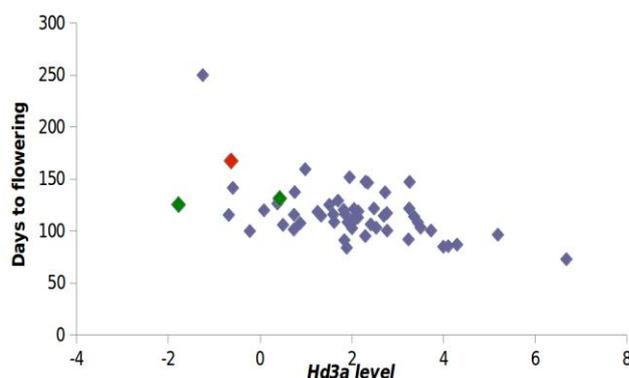
inicio	orientación	secuencia de reconocimiento	TF modelo	PLACE site #
-1824	(+)	CAACA	RAV1AAT	S000314
-1474	(-)	AGAAA	POLLEN1LELAT52	S000245
-1401	(+)	CAAT	CAATBOX1	S000028
-1237	(+)	ATAGAA	BOXIINNTPATPB	S000296
-1237	(+)	AGAAA	POLLEN1LELAT52	S000245
-1237	(+)	GRWAAW	GT1CONSENSUS	S000198

**Figura 4:** A) Mapa de las variaciones de la región promotora de *Hd3a* desde -2.000 pb al inicio de la transcripción. B) Predicción de sitios de reconocimiento por varios factores de transcripción en la secuencia del promotor de *Hd3a* dónde se han localizado polimorfismos en la región -2.000 a -860 pb. Se indica la secuencia de reconocimiento y la posición en el promotor.

### **Relación alelo *Hd3a*-floración:**

En la Figura 5 se señala la expresión de *Hd3a* de los alelos B y C (modificación de la Figura 2B). Los niveles de expresión del alelo A, presente de manera mayoritaria, fueron muy dispersos. El escaso número de individuos que poseen los alelos B y C mostraron una expresión baja comparada con la de los alelos A. Los datos de expresión obtenidos en variedades *índica* (Takahashi *et al.*, 2009), indican una expresión menor de los alelos del grupo A de promotores. Esto muestra una inversión de tendencia respecto a nuestro alelo A que induce una expresión mayor a los promotores del grupo B (alelos B y C) en el arroz de tipo *japónica*.

Las plantas que contenían el alelo B florecieron antes que el C. No obstante los pocos individuos que poseen los alelos B o C hacen imposible obtener una conclusión firme.



**Figura 5:** Gráfico que muestra los tiempos de floración asociados al logaritmo neperiano del nivel de expresión de *Hd3a*. Se muestra cada alelo hallado en el estudio por EcoTILLING con un color. Azul: alelo A, verde: alelo B, rojo: alelo C.

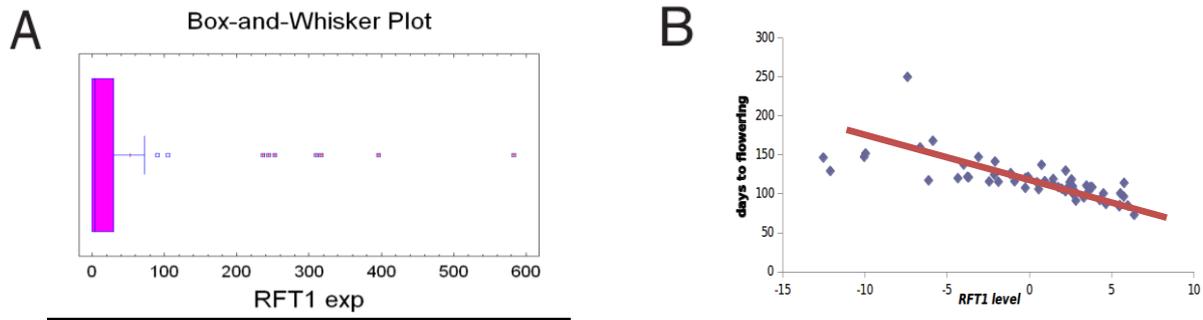
### **4.3.2. Región promotora de *RFT1*:**

#### **Nivel de expresión de *RFT1*:**

Los niveles de expresión de *RFT1* son también muy variables entre cultivares. El 50% de la población presenta una expresión relativa entre 0 y 25, siendo considerados como normales valores entre 0 y 65. Se considerarán anormales aquellos datos que se alejen más de dos desviaciones típicas de la media de la población. Hay nueve variedades que presentan datos anormalmente altos: Arroz da Terra, Carnice Precoce, CRLB1, Gigante Vercelli, Giovanni Marchetti, Gladio, Lemont, Loto y Pavlovsky (Figura 6).

Al analizar los datos de floración teniendo en cuenta la expresión de *RFT1* vemos la aparición de una correlación negativa -0,52, algo más elevada que en el caso de *Hd3a* (Figura 2). Esta correlación es moderada pero más cercana a uno que la de *Hd3a*, por lo que sería más atribuible la variación del tiempo de floración a la de la expresión de *RFT1*. La recta de regresión de los datos de floración enfrentados a los datos de expresión de *RFT1* presenta una pendiente negativa más marcada.

Variedades como Giovanni Marchetti que presentaban una floración temprana a pesar de tener bajos niveles de mRNA de *Hd3a* encuentran aquí la explicación a su precocidad. Esta es una de las variedades con una expresión de *RFT1* anómala, con un dato de 236.



**Figura 6: A)** Diagrama de Box-Wishker de los datos de expresión de *RFT1* a 77 DDS. El 50% de la población presenta una expresión entre 0 y 25. Unos valores entre 0 y 65 se consideran normales, habiendo nueve variedades con expresiones anormalmente altas.

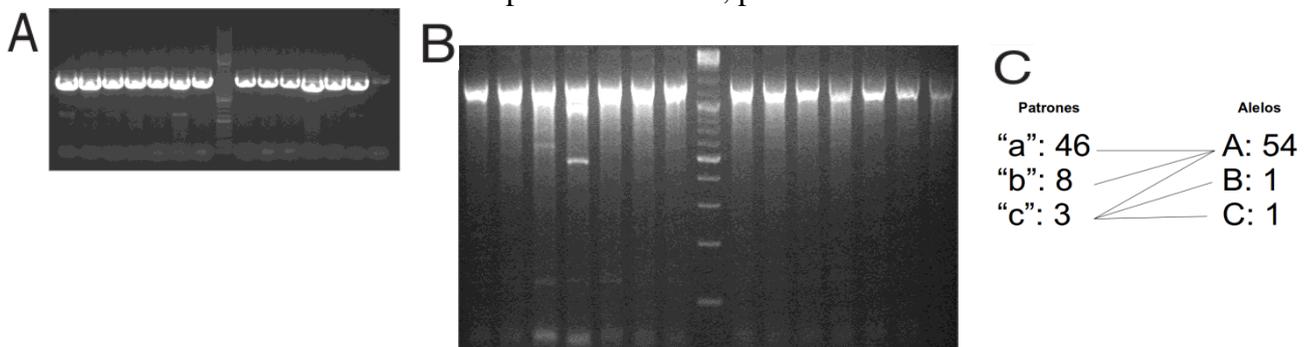
**B)** Gráfico que muestra los tiempos de floración asociados al logaritmo neperiano del nivel de expresión de *RFT1*. Se aprecia la correlación negativa como una ligera pendiente negativa de la recta de regresión que corresponde con una correlación de -0.52.

### *EcoTILLING de RFT1:*

En el caso de *RFT1* se estudiaron 1.200 pb de la región promotora, incluyendo la región 5'UTR. Algunos ADNs amplificados presentaron una banda inespecífica a 700 pb que tuvo que tenerse en cuenta para la posterior interpretación de las digestiones. La amplificación por PCR en las variedades Norin8 y Kasalath presentó ciertas dificultades. En el caso de Norin8 todos los intentos fueron infructuosos, por lo que se muestra esta variedad como una indeterminación.

La digestión de los heterodúplex por CEL1 de la región reguladora mostró 3 patrones de bandas diferentes. La ausencia de bandas resultó mayoritaria, siendo el patrón de Nipponbare el más extendido en la población. Se le denominó patrón “a”. El patrón llamado “b” agrupa 8 variedades, que presentaban bandas inespecíficas de PCR. El último patrón, “c”, estaba presente en 3 variedades, dos de ellas italianas.

Se secuenciaron los productos de las amplificaciones por PCR de *RFT1*, de algunos representantes de cada grupo, así como individuos con patrones de bandas confusos. Se emplearon como cebadores los oligonucleótidos H-I (ver anexo III). Las alineaciones de las secuencias mostraron 3 alelos, diferentes a los grupos de patrones de bandas. Todos los representantes del patrón “b” resultaron ser idénticos a Nipponbare, indicando que las bandas de las digestiones provenían, como se sospechaba, de reacciones inespecíficas de la PCR. El patrón “c” se dividió en tres tipos de secuencia, pasando una de sus variedades al alelo A.



**Figura 7: A)** PCR de la región promotora de *RFT1*, se aprecia una banda a 700 pb en algunas variedades.

**B)** Electroforesis de la digestión de heterodúplex de la región promotora de *RFT1*, de cada una de las variedades frente a Nipponbare. La última carrera corresponde a Nipponbare digerido con CEL1.

**C)** Distribución de las variedades de la colección según su patrón de bandas procedente de las digestiones de CEL1 para el gen *RFT1* y clasificación según los alelos deducidos de las secuencias.

Se procedió como en el caso anterior a la búsqueda de alteraciones en posibles sitios de unión a ADN. Mediante la base de datos PLACE se detectó la presencia de posibles sitios de unión a factores de transcripción, algunos de ellos podrían ser relevantes en la floración. Es el caso de 4 posibles sitios de unión a factores de tipo ARR1, definidos por la secuencia NGATT, que presenta una alta similitud con EHD1. Todos ellos estaban intactos en los 3 alelos hallados, por lo que la capacidad de reconocimiento de *Ehd1* no debería verse afectada.

**A**

posición	-1149	-1091	-1011	-934	-910	-879	-847	-777	-769	-728	-590	-480	-477	-468	-385	-337	n° cvs
Alelo A	G	G	C	T	G	G	G	A	G	T	G	G	-	-	T	--	54
Alelo B	A	C	T	C	A	A	C	G	A	A	A	G	-	-	T	--	1
Alelo C	G	G	C	T	G	G	G	A	G	T	G	-	A	C	A	GT	1

**B**

inicio	orientación	secuencia de reconocimiento	TF modelo	PLACE site #
-1091	(-)	CAAT	CAATBOX1	S000028
-934	(-)	TATAAAT	TATABOX2	S000109
-777	(+)	CANNTG	EBOXBNNAPA	S000144
-777	(+)	CATATG	CATATGGMSAUR	S000370
-769	(-)	GAGAC	SURECOREARSULTR11	S000499
-769	(+)	YGTGCWC	SEBFCONSSTPR10A	S000391
-769	(+)	TGTCTC	ARFAT	S000270
-728	(-)	ATAGAA	BOXIINTPATPB	S000296
-468	(-)	AGAAA	POLLEN1LELAT52	S000245

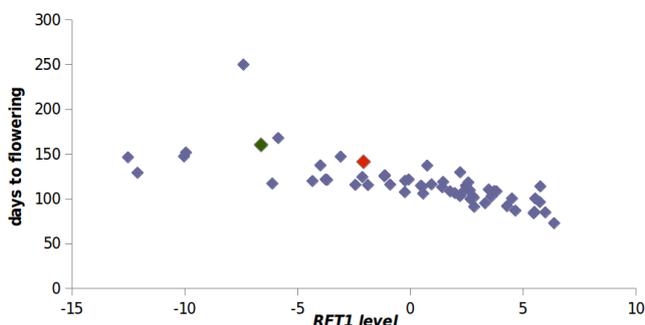
inicio	orien-tación	secuencia de reconocimiento	TF modelo	PLACE site #
-485	(-)	CATGCA	RYREPEATBNNAPA	S000264
-477	(+)	GAGAC	SURECOREATSULTR11	S000499
-477	(-)	TGTCTC	ARFAT	S000270
-468	(-)	AGAAA	POLLEN1LELAT52	S000245
-385	(-)	TATTAG	CPBCSPOR	S000491
-385	(+)	YACT	CACTFTPPCA1	S000449

**Figura 8: A)** Mapa de los polimorfismos de la región promotora de *RFT1*.

**B)** Predicción de sitios de reconocimiento por varios factores de transcripción en la secuencia del promotor de *RFT1* dónde se han localizado polimorfismos. Se indica la secuencia de reconocimiento y la posición en el promotor. El primer panel corresponde al alelo B, el segundo panel corresponde al alelo C.

### Relación alelo *RFT1*-floración:

Se localizaron los alelos en el gráfico de la Figura 6B (Figura 9). Los niveles de expresión del alelo A fueron muy dispersos. Los alelos B y C no presentaron niveles de expresión diferenciales respecto de A. Los pocos individuos con los alelos B y C hacen que no se puedan extraer conclusiones acerca de la influencia del alelo de *RFT1* en el nivel de expresión.



**Figura 9:** Gráfico que muestra los tiempos de floración asociados al logaritmo neperiano del nivel de expresión de *RFT1*. Se muestra cada alelo hallado en el estudio por EcoTILLING con un color. Azul: alelo A, verde: alelo B, rojo: alelo C.

#### **4.4. RELACIÓN *Hd3A-RFT1*:**

Los datos de expresión de ambos genes evidencian una correlación entre ellos. Las cinco variedades con expresiones anormalmente elevadas de *Hd3a* se encuentran entre las nueve variedades anómalas de *RFT1*. El análisis de ambas variables indica que existe una correlación positiva de 0.728 entre ambas: a mayor expresión de *Hd3a* mayor expresión de *RFT1* y viceversa (Figura 2).

# DISCUSIÓN

Este trabajo tiene como objetivo conocer aspectos relacionados con el control de la floración en condiciones de fotoperiodo de día largo, centrándose en los dos genes que cumplen la función de florígeno en arroz. La finalidad última de este conocimiento es el diseño de variedades mejor adaptadas al clima mediterráneo, promoviendo la precocidad de la floración.

El arroz fue domesticado en las regiones tropicales de China, a una latitud aproximada de 31°N. Los datos que se conocen apuntan a dos eventos de domesticación que dieron lugar a diferentes subgrupos, uno de los cuáles desembocó en cultivares que se adaptaron a zonas templadas alcanzando el límite norte de 45°N, dónde los días se alargan en verano y se acortan en invierno. Limitado por el frío invernal, en estos climas el arroz se ve forzado a completar su ciclo biológico entre la primavera y el verano, cuándo las condiciones de temperatura son adecuadas para el cultivo, y a florecer en condiciones de fotoperiodo de día largo. El resto de subgrupos, permaneció en los trópicos, dónde florecen en condiciones de fotoperiodo de día corto. Puesto que *Hd3a* y *RFT1* son los genes clave de la inducción de la floración, este trabajo se centró en conocer la variación alélica de estos genes en variedades adaptadas a latitudes norte, y de esta manera, tratar de entender el papel que han jugado en él los cambios adaptativos.

La floración de las variedades de la colección ocurrió en condiciones de fotoperiodo largo y se prolongó durante varias semanas, proporcionando tiempos de floración muy dispares. Esta alta variabilidad es esencial a la hora de analizar los factores que regulan la floración.

El gen *Hd3a* ha sido señalado como el principal responsable de la floración e identificado como el florígeno o la señal móvil que induce la floración. En condiciones inductoras de floración, es decir, en días cortos, sería esperable una expresión alta de *Hd3a*. Esto es lo que se observa en los estudios previos realizados con variedades *índica*, adaptadas a climas de fotoperiodo corto. Al contrario, la floración en arroz está inhibida en condiciones de día largo y por lo tanto no es esperable una expresión alta de *Hd3a* (Takahashi *et al.*, 2009). Sin embargo, el análisis de expresión llevado a cabo en nuestra colección de variedades *japónica* adaptadas a climas mediterráneos, muestra una expresión detectable e inusualmente alta de *Hd3a* en condiciones de fotoperiodo largo. Este hecho indica que la inhibición de la expresión de *Hd3a* en condiciones de día largo no se produce en las variedades de arroz adaptadas.

Tal y como predice la base genética de la floración la correlación entre los datos de expresión de *Hd3a* y el tiempo de floración es negativa, es decir, cuanto mayor es el nivel de expresión de *Hd3a*, antes se desencadena la transición floral. Sin embargo, este coeficiente de correlación es bajo, indicando que *Hd3a* no es el principal responsable de la floración en condiciones de fotoperiodo largo. Esta hipótesis viene reforzada por el hecho de que ciertas variedades no presentan los niveles de *Hd3a* esperables de acuerdo a su tiempo de floración. Como ejemplo, las variedades Gema y Ullal florecen al mismo tiempo, a pesar de que Ullal posee una expresión de *Hd3a* tres veces superior a la de Gema.

El análisis mediante EcoTILLING de la variabilidad genética en la región del promotor de *Hd3a* mostró una población altamente homogénea. 54 variedades de 57 presentaban la misma

región promotora que Nipponbare.

La correlación moderada entre los niveles de expresión de *Hd3a* y el tiempo de floración, junto con la homogeneidad alélica en la población hizo sospechar la influencia de otro factor que regulase en estas condiciones. Estudios previos (Komiya *et al.*, 2008) apuntaban a *RFT1*, el gen considerado como redundante de *Hd3a* y también identificado adicionalmente como florígeno, como el posible factor adicional en la regulación de la floración en condiciones de fotoperiodo largo.

El análisis de los niveles de expresión en relación al tiempo de floración desvela una correlación negativa mayor que en el caso de *Hd3a*, indicando una mayor influencia de *RFT1* en la inducción de la floración.

Al igual que *Hd3a*, la población presentaba una alta homogeneidad genética en la región promotora en el estudio por EcoTILLING. Se halló un alelo mayoritario idéntico a Nipponbare. El alelo A, presente en 54 variedades mostraba niveles de expresión similares al resto de alelos hallados. El alelo C inducía una expresión muy elevada de *RFT1*, presentando una floración muy precoz. La conservación tan alta de este gen puede indicar la importancia de su función.

En algunos de los cultivares analizados, los niveles de expresión de *Hd3a* bajos son compensados con niveles altos de *RFT1*, suficientes para explicar la inducción de la floración. Es el caso del cultivar temprano Giovanni Marchetti que compensa un bajo nivel de expresión de *Hd3a* con un alto nivel de expresión de *RFT1*. Tanto *Hd3a* como *RFT1* muestran correlación con el tiempo de floración, siendo ésta mayor en el caso de *RFT1*, lo que induce a pensar en un control conjunto por parte de ambos genes del tiempo de floración, siendo *RFT1* el que juega un papel predominante en condiciones de fotoperiodo de día largo.

Cabe destacar que la tendencia de la expresión de ambos genes muestra una correlación positiva y cercana a uno. El control ejercido por un mismo factor sobre ambos genes podría estar en el origen de esta correlación. El hecho de que *Hd3a* y *RFT1* se hallen separados únicamente por 11,5 kb en el genoma debido a un origen común por una duplicación en tándem (Kojima *et al.*, 2002) apoya la idea de elementos de regulación compartidos.

No hay muchos datos conocidos acerca de la regulación de la expresión de *RFT1*. Komiya *et al.*, 2009, proponen un control de la floración ejercido por *Ehd1* sobre ambos genes en condiciones de fotoperiodo largo (figura 2 Introducción). La presencia de sitios de reconocimiento de secuencias comunes a ambas regiones promotoras, como la presencia de un sitio de reconocimiento de secuencia de tipo ARR1 o CAAT-box, apoyan la hipótesis de una regulación común.

Sin embargo, una conexión en sus rutas de regulación, controladas por un mismo factor como EHD1, debería implicar una misma influencia sobre el tiempo de floración ejercida por ambos genes. No obstante *RFT1* presenta un grado de correlación mayor que *Hd3a* con el tiempo de floración. Los datos obtenidos confirman que el peso de la ruta de floración se ha desplazado hacia el florígeno *RFT1*, que en fotoperiodo largo pasa a ser el principal receptor de las señales inductoras de floración conocidas. *Hd3a* estaría jugando un rol complementario en fotoperiodos largos.

Las diferencias en la correlación con el tiempo de floración de ambos genes podría deberse a una influencia directa de uno de estos genes sobre la regulación del otro, o bien a la presencia de más factores involucrados en la regulación de los genes.

Los elementos reguladores que interaccionan con estas regiones podrían presentar variación y provocar las diferencias en el nivel de expresión de los genes. La variación de *Ehd1* podría ser

crucial para entender los cambios en la longitud del ciclo. Estudios anteriores en nuestro laboratorio indican que no hay correlación entre el tipo de alelo de *Ehd1* y la duración del ciclo en condiciones de fotoperiodo largo (Naranjo, TFC 2010). HD1 podría jugar un papel importante en estas circunstancias debido a su acción represora. Sin embargo, la mayor parte de estos cultivares presentan un alelo de HD1 no funcional y por lo tanto no puede inhibir la expresión de *Hd3a*, lo que hace pensar en la intervención de otros factores en la inducción de la floración.

Los fitocromos ejercen una gran influencia en la floración a través de la regulación de *Hd1*, cambiando su conformación. También se ha demostrado que modulan la expresión de *Ehd1*. Los mutantes de pérdida de función del gen *se5*, involucrado en la síntesis del cromóforo de los fitocromos, produce plantas con niveles de *Ehd1* altos y que presentan una floración muy temprana, del orden de 42 DDS, e insensible a fotoperiodo (Andrés *et al.*, 2009). Parece por tanto que *Ehd1* podría estar en el origen de las variaciones en el tiempo de floración de las variedades estudiadas.

Aunque nuestra hipótesis consiste en que las variaciones de expresión de *Hd3a* y *RFT1* son los reguladores de la floración, no hay que descartar que los procesos de adaptación a nuestra latitud pudiesen haber producido variaciones en las regiones codificantes de estos genes, introduciendo variabilidad en las proteínas con función de florígeno. Las diferencias en el tiempo de floración dentro de los alelos mayoritarios podrían deberse a diferentes funcionalidades de las proteínas. Un estudio de las regiones codificantes permitiría verificar este aspecto.

El empleo del EcoTILLING durante este trabajo nos ha permitido familiarizarnos con sus particularidades, detectando sus ventajas y sus inconvenientes.

La primera variante desarrollada del método EcoTILLING (Comai *et al.*, 2004) es muy resolutiva, ya que se emplean cebadores marcados con fluorescencia que se usan para detectar los fragmentos de ADN originados por la digestión, tras ser separados mediante electroforesis con geles de acrilamida. No obstante, la presencia de más de un polimorfismo en una misma región analizada es difícil de detectar ya que sólo están marcados los extremos. En el presente trabajo se utilizó una variante en la que no se emplea fluorescencia y la separación de los fragmentos generados en la digestión con CEL1 se realiza con geles de agarosa (Raghavan *et al.*, 2007). Este método de detección de polimorfismos es rápido y de bajo coste, pero presenta ciertas limitaciones, principalmente derivadas de la capacidad resolutiva del gel de agarosa, que no permite la detección de bandas inferiores a 100 pb, aproximadamente. Por tanto, el análisis de un fragmento de ADN con un gran número de polimorfismos resultaría en un número elevado de fragmentos de ADN cercanos al límite de resolución del gel. Tampoco se pueden detectar polimorfismos que disten pocas pares de bases, en los que la digestión con CEL1 produzca fragmentos muy pequeños. De hecho, en este trabajo se ha podido detectar mediante secuenciación polimorfismos diferentes en ADN proveniente de variedades que presentaban el mismo patrón de bandas. El uso de geles con mayor resolución, por ejemplo con acrilamida, implicaría aumentar el tiempo del análisis, eliminando la ventaja de la rapidez.

Por otro lado, otro aspecto dificultoso del EcoTILLING es la posible presencia de bandas inespecíficas, debidas a hibridaciones aleatorias de los cebadores en porciones del genoma diferentes a la secuencia objetivo, que hace la interpretación de los patrones de bandas algo más complicada. Este inconveniente se elude estableciendo una serie de controles adecuados y seleccionando cuidadosamente los cebadores para la reacción de PCR, que nos permitirán eliminar las bandas no asociadas a polimorfismos.

A pesar de estos inconvenientes, el EcoTILLING es un método óptimo para el análisis de un número elevado de muestras con fragmentos de ADN con pocos polimorfismos. En esta situación el método destaca por permitir la visualización directa del número y posición de polimorfismos de una misma secuencia en una gran cantidad de variedades de forma simultánea (tantas como carreras tenga el gel). Así se facilita la clasificación de las variedades en grupos alélicos, disminuyendo el número de fragmentos de ADN que hay que secuenciar dentro de cada grupo. Esta ventaja resultaba especialmente útil unos años atrás, cuando el método fue puesto a punto (Comai *et al.*, 2004). Hoy día el coste de las reacciones de secuenciación ha disminuido considerablemente de tal forma que se ha de sopesar la conveniencia del uso del EcoTILLING frente a la secuenciación directa de los genes, tras una amplificación por PCR. Cabe destacar además que la precisión de la secuenciación es insuperable, por lo que se presenta como la opción a elegir en futuros estudios de estas características.

# CONCLUSIONES

1. Se ha caracterizado el tiempo de floración de 57 variedades de *Oryza sativa* que se cultivan en clima templado. En condiciones de fotoperiodo largo, las plantas florecieron mostrando una distribución normal con una ligera asimetría positiva. La media de floración fue de 116,5 DDS, con una desviación típica de 22,3 DDS.
2. Se cuantificó la expresión de los genes responsables de la floración *Hd3a* y *RFT1*, a los 77 DDS. La correlación existente entre los niveles de expresión de *Hd3a* y *RFT1* con el tiempo de floración fue -0,3449 y -0,5209 respectivamente, lo que señala que el control de la floración en día largo recae sobre *RFT1*, complementado por *Hd3a*.
3. La región promotora de *Hd3a* está muy conservada. La variación genética de la región promotora de *Hd3a* se distribuyó en tres alelos, siendo uno de ellos mayoritario, el alelo A, presente en 54 variedades. Dos variedades no europeas compartían el alelo B. El alelo C pertenecía a una variedad taiwanesa.
4. La variación genética de la región promotora de *RFT1* se distribuyó en tres alelos. El alelo A, mayoritario, estaba presente en 54 variedades, el alelo B y el alelo C estaba presentes en dos variedades asiática tardías.
5. La baja tasa de variación de los promotores de *Hd3a* y *RFT1* y la correlación moderada con los niveles de expresión indican la implicación de ambos genes, actuando de forma sinérgica, en la inducción de la floración o la existencia de otros factores reguladores adicionales.
6. Los niveles de expresión de *Hd3a* y *RFT1* muestran una correlación alta indicando una conexión en la regulación de ambos genes.
7. El método de análisis de secuencia empleado, el EcoTILLING, facilitó la división de las variedades en grupos según su patrón de bandas, reduciendo la necesidad de secuenciación. No obstante el método presenta algunas limitaciones, una de ellas reside en la escasa capacidad de reconocimiento de polimorfismos distanciados unas pocas pares de bases.

# BIBLIOGRAFÍA

- Andrés F., Galbraith D., Talón M., Domingo C.** (2009) Analysis of *PHOTOPERIOD SENSITIVITY5* sheds light on the role of phytochromes in photoperiodic flowering in rice. - *Plant Physiology* 151, 681-690
- Baurle, I. and Dean, C.** (2006) The timing of developmental transitions in plants. - *Cell* 125, 655-664
- Comai L., Young K., Till B., Reynolds S., Greene E., Codomo C., Enns L., Johnson J., Burtner C., Odden A., Henikoff S.** (2004) Efficient discovery of DNA polymorphisms in natural populations by Ecotilling. - *The Plant Journal* 37, 778-786
- Corbesier L., Vincent C., Jang S., Fornara F., Fan Q., Searle I., Giakountis A., Farrona S., Gissot L., Turnbull C.** (2007) FT protein movement contributes to long-distance signaling in floral induction of *Arabidopsis*. - *Science* 316, 1030-1033
- Doi K., Izawa T., Fuse T., Yamanouchi U., Kubo T., Shimatani Z., Yano M. and Yoshimura A.** (2004) *Ehd1*, a B-type response regulator in rice, confers short-day promotion of flowering and controls *FT-like* gene expression independently of *Hd1*. - *Genes Dev.* 18, 926-936
- Fuller D. Q., Qin L., Zheng Y., Zhao Z., Chen X., Hosoya L. A., Sun G. P.** (2009) The Domestication Process and Domestication Rate in Rice: Spikelet Bases from the Lower Yangtze. - *Science* 323, 1607 – 1610
- Hagiwara W. E., Uwatoko A., Sasaki A., Matsubara K., Nagano H., Onishi K. and Sano Y.** (2009) Diversification in flowering time due to tandem *FT-like* gene duplication, generating novel Mendelian factors in wild and cultivated rice. - *Mol. Ecol.* 18, 1537-1549
- Hayama R., Yokoi S., Tamaki S., Yano M. and Shimamoto K.** (2003) Adaptation of photoperiodic control pathways produces short-day flowering in rice. - *Nature* 422, 719-722
- Imaizumi T. and Kay S. A.** (2006) Photoperiodic control of flowering: not only by coincidence. - *Trends Plant Sci.* 11, 550-558
- Izawa Takeshi** (2007) Daylength Measurements by rice plants in photoperiod short-day flowering. - *International Review of Cytology* 256, 191-222
- Kojima S., Takahashi Y., Kobayashi Y., Monna L., Sasaki T., Araki T. and Yano M.** (2002) *Hd3a*, a rice ortholog of the *Arabidopsis FT* gene, promotes transition to flowering downstream of *Hd1* under short-day conditions. - *Plant Cell Physiol.* 43, 1096-1105
- Komiya R., Ikegami A., Tamaki S., Yokoi S. and Shimamoto, K.** (2008) *Hd3a* and *RFT1* are essential for flowering in rice. - *Development* 135, 767-774
- Komiya R., Yokoi S., Shimamoto K.** (2009) A gene network for long-day flowering activates *RFT1* encoding a mobile flowering signal in rice. - *Development.* 136(20):3443-50
- Khush G. S.** (1997) Origin, dispersal, cultivation and variation of rice. - *Plant Molecular Biology* 35, 25–34

- Lee S., Kimy J., Han J. J., Han M. J. and An G.** (2004) Functional analyses of the flowering time gene *OsMADS50*, the putative *SUPPRESSOR OF OVEREXPRESSION OF CO 1/AGAMOUS-LIKE20 (SOC1/AGL20)* ortholog in rice. - *The Plant Journal* 38, 754-764
- Mannion A. M.** (1999) Domestication and the origins of agriculture: an appraisal - *Progress in Physical Geography* 23, 1, 37-56
- McCallum C. M., Comai L., Greene E. A., Henikoff S.** (2000) Targeting Induced Local Lesions IN Genomes (TILLING) for Plant Functional Genomics. - *Plant Physiology* 123, 439-442
- Mizoguchi T., Wright L., Fujiwara S., Cremer F., Lee K., Onouchi H., Mouradov A., Fowler S., Kamada H., Putterill J., and Coupland G.** (2005) Distinct Roles of *GIGANTEA* in Promoting Flowering and Regulating Circadian Rhythms in Arabidopsis - *The Plant Cell* 17, 8, 2255-2270
- Naranjo L.** (2010) Caracterización de la variación natural de genes reguladores de la floración en cultivares de tipo japónica de *Oryza sativa*. – Trabajo Fin de Carrera, ETSIAMN, UPV
- Raghavan C., Naredo M. E., Wang H., Atienza G., Liu B., Qiu., McNally K., Leung H.** (2007) Rapid method for detecting SNPs on agarose gels and its application in candidate gene mapping. - *Mol. Breeding* 19, 2, 87-101
- Riechmann J. L., Heard J., Martin G., Reuber L., Jiang C., Keddie J., Adam L., Pineda O., Ratcliffe O. J., Samaha R. R., Creelman R., Pilgrim M., Broun P., Zhang J. Z., Ghandehari D., Sherman B. K., Yu G.** (2000) *Arabidopsis* transcription factors: genome-wide comparative analysis among eukaryotes. - *Science* 15;290(5499):2105-10
- Samach A., Onouchi H., Gold S. E., Ditta G. S., Schwarz-Sommer Z., Yanofsky M. F. and Coupland G.** (2000) Distinct roles of *CONSTANS* target genes in reproductive development of *Arabidopsis*. - *Science* 288, 1613-1616
- Takahashi Y., Kosuke M. T., Shuji Y., Hideki I., Ko S.** (2009) Variations in Hd1 proteins, *Hd3a* promoters, and *Ehd1* expression levels contribute to diversity of flowering time in cultivated rice. - *PNAS Early Edition* [www.pnas.org/cgi/doi/10.1073/pnas.0812092106](http://www.pnas.org/cgi/doi/10.1073/pnas.0812092106)
- Xu X., Liu X., Ge S., Jensen J. D., Hu F., Li X., Dong Y., Gutenkunst R. N., Fang L., Huang L., Li J., He W., Zhang G., Zheng X., Zhang F., Li Y., Yu C., Kristiansen K., Zhang X., Wang J., Wright M., McCouch S., Nielsen R., Wang J., Wang W.** (2011) Resequencing 50 accessions of cultivated and wild rice yields markers for identifying agronomically important genes. - *Nat. Biotechnololy* 11;30(1):105-11
- Xue W., Xing Y., Weng X., Zhao Y., Tang W., Wang L., Zhou H., Yu S., Xu C., Li X.** (2008) Natural variation in *Ghd7* is an important regulator of heading date and yield potential in rice. - *Nat. Genet.* 40, 761-767
- Yano M., Katayose Y., Ashikari M., Yamanouchi U., Monna L., Fuse T., Baba T., Yamamoto K., Umehara Y., Nagamura Y.** (2000) *Hd1*, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the Arabidopsis flowering time gene *CONSTANS*. - *Plant Cell* 12, 2473-2484

# ANEXO I

Grupo	VARIETADES	Continente	Días hasta emer-gencia panícula	alelo pro-motor <i>Hd3a</i>	expresión relativ <i>Hd3a</i> x100	alelo pro-motor <i>RFT1</i>	expresión relativ <i>RFT1</i> x10
japónica	AGAMI	África	168,0	a	0,519	a	2,85x10 <sup>-3</sup>
japónica	AKIHIKARI	Asia	113,8	a	7,95	a	1,69
japónica	ARROS DA TERRA	Europa	73,0	a	799	a	583
japónica	AZUCENA	Asia	141,5	a	0,553	c	0,124
japónica	BAHÍA	Europa	121,8	a	12	a	0,925
japónica	BENDRET	Europa	119,0	a	8,52	a	4,27
japónica	BENISANTS	Europa	113,0	a	8,45	a	4,11
japónica	BOMBA	Europa	118,5	a	3,51	a	13
japónica	CALMOCHI	Norteamérica	100,0	a	0,8	a	14,3
japónica	CAMPINO	Europa	146,5	a	10,5	a	3,65x10 <sup>-6</sup>
japónica	CARNICE PRECOCE	Europa	100,5	a	41,7	a	253
japónica	CESARE	Europa	108,5	a	6,77	a	45,4
japónica	CORMORAN	Europa	115,8	a	2,1	a	0,0871
japónica	CRLB 1	Europa	85,0	a	54,5	a	396
japónica	CT 18	Sudamérica	110,5	a	7,77	a	32,2
japónica	DAEGUDO	Asia	91,3	a	6,3	a	16,9
japónica	DONGJIN	Asia	125,3	a	4,54	a	0,315
japónica	FONSA	Europa	95,3	a	9,92	a	27,4
japónica	GAVINA	Europa	103,3	a	12,6	a	15,7
japónica	GEMA	Sudamérica	114,8	a	3,79	a	1,59
japónica	GIGANTE VERCELLI	Europa	87,0	a	73,5	a	105
japónica	GIOVANNI MARCHETTI	Europa	84,0	a	6,62	a	236
japónica	GIZA 178	África	116,0	a	4,94	a	90,2
japónica	GLADIO	Europa	100,5	a	16	a	2,55
japónica	GLEVA	Europa	116,3	a	6,4	a	36,2
japónica	GUADIAMAR	Europa	102,8	a	7,43	a	0,0185
japónica	GUWEOLDO	Asia	137,5	a	2,13	a	0,407
japónica	HARRA	Australia	115,5	a	0,506	a	0,152
japónica	HWAYOUNG	Asia	107,7	a	2,4	a	0,78
índica	IR 64	Asia	129,3	a	5,48	a	5,58x10 <sup>-6</sup>
japónica	ITALICA LIVORNO	Europa	106,5	a	7,08	a	9,8
japónica	JSENDRA	Europa	109,8	a	7,57	a	13,9
índica	KASALATH	Asia	124,5	b	0,171	a	0,119
japónica	KINMAZE	Asia	200,0	a	0,287	a	6,09x10 <sup>-3</sup>
japónica	KOSHIHIKARI	Asia	106,0	a	1,65	a	1,76
japónica	L202	Norteamérica	120,0	a	1,09	a	0,013
japónica	LEDA	Europa	129,8	a	1,52	a	9,09
japónica	LEMONT	Norteamérica	114,0	a	28,8	a	317
japónica	LOTO	Europa	85,5	a	60,7	a	244
japónica	LTH	Asia	101,5	a	2,09	a	16,6
japónica	M 201	Norteamérica	121,8	a	25,9	a	0,0233
japónica	M 202	Norteamérica	108,5	a	31,2	a	5,82
japónica	MOROBBERKAN	África	147,5	a	10	a	4,38x10 <sup>-5</sup>
japónica	NIPPONBARE	Asia	147,3	a	26,1	a	0,0454
japónica	NORIN 8	Asia	126,5	b	1,46	-	0,322
japónica	PAVLOVSKY	Rusia	96,5	a	179	a	310
japónica	PEGONIL	Europa	117,3	a	15,9	a	2,19x10 <sup>-3</sup>
japónica	PUNTAL	Europa	120,3	a	6,2	a	0,79
japónica	RIPALLO	Europa	103,3	a	33,2	a	8,97
japónica	ROMOLO	Europa	106,5	a	11,2	a	7,17
japónica	SASANISHIKI	Asia	121,3	a	7,77	a	0,0249
japónica	SENIA	Europa	108,8	a	5,04	a	41,3
japónica	SEQUIAL	Europa	92,0	a	25,5	a	72,4
japónica	T 65	Asia	159,5	c	2,68	b	1,31x10 <sup>-3</sup>
japónica	T 67	Asia	137,3	a	15,3	a	2,1
japónica	ULLAL	Europa	114,8	a	14,9	a	11,7
japónica	ZHONGHUA 15	Asia	151,8	a	7,05	a	4,77x10 <sup>-5</sup>

# ANEXO 2

## *Hd3a*

### 1. Secuencia completa del gen *Hd3a* y de su región promotora:

>6 dna:chromosome chromosome:MSU6:6:2936905:2941553:1

TATAATTGTTATGCCATTGATTTATTGACACAGCTAGAGATGGTATCGGCTTATCTTGCC  
CGGACAAATCGAGCTAAGCAAGAACATTTACATAATAAGCAGGCATCAGATTAAGACAGC  
AACGCAGAAATTGCAGACAGAGAGAAATGCATGGAGGCAGCTAGCAAGCAGCCCCTGCCTGT  
TGGTTCCTACACGGTTTCAATCAAT **CACCAGCTAATTGGCAGACA**TTCAAACAAATCGC > Hd3a **prom2**  
AACACCGTAGTATAGTACTTTTAAATTCGTCTTCGAGATACGACTTTAATCATTAATTGCC  
TTACCTCAACTTTATGTTCCGATTACATATCATGTAATATTATAAAAAAATCTAATGGTG  
TAATAAATGTTATAAAAAAATAACTTAGATGCATAGTTCGGGATATATGCCCTCCGGTTTT  
TAATGTACAACACTTAAATTTTGGATGTGTGTTGTTTGAATCAAATCTGAAATAGTAT  
GTTACTTTTAAAGTGCCTTTAAAGATAAAACAAATCACAATAAAAAATACATATATTTTTTA  
AATAAGACAGATGATCAAGCATATATTCAAAGTCAACAGATCAGTCATACATTATTAGAA  
GAGAAATTTCTTGTATGCCCTTGAATTTAAGTCAATCTTTTCTATGCCCTTGAATTTTAC  
TCACTCTCTTGTATAGTTGTATACCTTTTAAATTATGATTTTCACTCTCTTTTACATCTTTT  
CGTTAGTTGACCGGTTCCCTTTAAAAATTAGAAGTTTGTGTACGAGACTATTATTTTCT  
GAGTTTGTGTGCGATGTACTATTTACAAAAAGAAATTAGATGTGCAAGAAAAATAATAT  
TTATTTAAAAACAATATGTGTAGCATATTTTTGTGATGTAACATATTTGCCACCTATTGTA  
TATATTATCTAAAAGAAATCCTATACAAGACCAATTTGTTTTTTTTCAATTATAAAAAAGC  
CGCACAAGCTTTTTTCATGAATATATTAATTTTTCTAAGTACAATACTTCCTCATAGAAAA  
AACTTATAATATTATAGTCCCACAATAAATTTACTATATTTCCAACGTTAGCATCCACAA  
AATTTCTAGGCATGTACAATTGTAAAATGACCTTTTTTACAGGCATTAACAGTTAACTAA  
CGGTACGGAAATGGTATAAAACGGGTTAAAAATAAAAATTCAGAGGTATATAAAAAATGAG  
CGATTAGCTAAACTTTTAGGAGTATAAAAGAAATTTCTTACTGAAAACCAAGAGGGAGTA  
CTACATATAACCTCATTTTT **TTACGGGTTGATAAGTTGC**GGAAAAACCAACAAATTAGCA > Hd3a **promF**  
ACAAATATGAGTAAAACCTTGTATACATGTGTCTTCTTAGCGATTTAAAAATCAATGCTGA  
AAATAAATTACAATAAAATTAATAAATCTCAAGATAATCTCTAAAATGTAGTTTTAAAAT  
TAAATTTTGATTGCGCATGATAAGAAAAAACAATAATGATGGGAGGCTATATCAACTGT  
CAAACCTGGCTAATTTAGAAAGACGACATCGAATTTCCATACAGATTGGAGGCAGCAAAGAG  
AGCCCGTCTCTGCCGGTGCATGATGCTCGATCATATCCCATCTCTCCTCACTTCA  
TCATCAACAAGAACAAGAGGAACCTATATAGCTGCAAGATCTAGACTGAACTACTATAGC  
AACTAACTGTACTGTAGCTAGATTACGCTTAGCTATAGCTGCTGCTGCATCTATCTGTAA  
ATTCCAACACGACGCTGACTGCTGCAAGCTAGCTTGGCCGGCTAGCTAGATAGCTAGCT  
AGCTCAAAAAGAGAAAGCTTTGCGAAATAGGAGTAACTAGCTAGCTCTAGCTAGGGCGCCA  
TCGATGCGTGACCACCGAGCTCGCCTCGTCCACACGTACAGGAAGACGATGCAGAAAGCG  
GCCGGCATGCAGTACACTGACCGAGCTAAGAGAGAGAGAGATGATATTATTCTTCTGC  
AGCTAAATTAAGTAAAGTTGGACATGGACATGGACATAGTAATTTTGCATGGCCATCA  
TCTTGCCCTCCTATATAAAGCGCCATCTCACCTCTCAACCACAGCTCGATCCAGCAGCCC  
**TGCACCACACACAGTTCAGCTAGCAGATCACCTAGCTAGATAGCTGCCTCTATCACAGTA**  
**TATTTGCTCCCTGCAACTTGCTGCTGCTGCAATAGCTAGCAGCTGCAGCTAGTAAGCAA**  
**ACTATAAACCTTCAGGGTTTTTTTGCAAGATCGATGGCCGGAAGTGGCAGGGACAGGGACC** -28  
**CTCTTGTGGTTGGTAGGGTTGTGGGTGATGTGCTGGACGCGTTTCGTCCGGAGC** < Hd3a **promR**  
**ACAAGGTCACCTA**GGCTCCAAGACCGTGTCCAATGGCTGCGAGCTCAAGCCGTCCATGG  
**TCACCACACAGCTTAGGGTCGAGTCCGGGCAATGACATGAGGACATTTCTACACCCTTG**  
TATGTGAGCTCTACCATGTGTGCTAGTTGGTGCAGAGACCAGAAGTTATACTTTCTTTCA  
TTATTATATTATAGAAAAATTTGTTTGGTACTTAACCCAAGATGACTTTAAAATGCATTA  
ATTTGATGTTTGTGTCATGGTTTTTTTGTGGTGTGTACCTGAAG**GTGATGGTAGACCAGATG** > Hd3a **C**  
**CACCAAGCCCAAGTGA**CCCTAACCTTAGGGAGTATCTACATTGGTAAGCACACTAATGTT  
AAGCTAGCGCACTTGTTTTCATGCAATGATCAAAATAATCACTGCTAGCTGATTTACATA  
CATAACCACATCTTGACTGCATGATGGGTGATCATCTTAATTTGTTTGTCCACAG**GTTG**  
**GTCACTGATATTCCTGGTACTACTGCAGCGTCATTTG**GTGTCAGTATTAACGATCAAATGG  
TTTAGATTGATT **TGATTTCCCACTTAAATACATTGC**AGTATAAAACTAACAGAGGTAAT < Hd3a **A**  
TGTCTTTCCATACTATTGTTTGGGGGGGGGGGGCAGGGGGCAGTGAGTTGTTGAGCATT  
TCCCTCACTCGGAGCCCCAGTTTGGCCGATTAATTTCCCATCATCCTTACCACCAC  
AAAAAAAAAAGTGATTTGGAGAAGGATACGCTCCTCCAATAATTTTTTAACAAGTTCACC

ACCACAACGAGGCACCGACAACCGCCATGAGTCCACGCACCCTGCAGCTAGTCGTGCCAC  
 CAGCACCACAGTTAGCCACACCGCGAGCAATCCACGCCATCACTTGCCGCTGTGAGTCCA  
 CGATCACACGCCCTCGATCAGCCAGCTGCCATGATATGTCGGCTTGGGCTTGGATTCTGCA  
 ACATACCATGTCCCCATCGATGGCACCACACTCTAATTCGCCAACACAACACTGAAGGGA  
 TCGGATCGGGAGGGCTAGCTCACTCAAGATAGTGACGGGATGATTCGATGGGAATCCAT  
 GGGGAACAAAGCACAATAATCATAAATGTCTTTTCAATTCATGGCAAATTCCTAAAG  
 CAATACTCTCTCCATTTCTATAATGTAGGTTGCACACTCATTTCAGATTCAACTTTCA  
 AAACATTTGACTAACGATTTGTATAATATAAATTTGAATTTTATTTGATAGAAATAATATT  
 ATTGGATTGATATTTCAATATACTTTTCATATGGTTATAATTTTGTACTACAAACTTTAT  
 AGTATATTAGAAATTATAAATAAAAAATTAGTTTTGTAGACCGTACATGCCAAGTTTGAC  
 CATACCATATTATGAAAAAGAGGGAGTAGTAACCTTGTATATTTGCAATATATACTCC  
 CTCGGTTTCTAAATATCTGACGCTGTTGACTTTTTTAAACATGTTTGACCGTTCGTCTTAT  
 TAAAAATAAGTAATTATTAATTTCTTTTCCATCATTGATTTATTGTTAAATATATTTT  
 TATGCAGACATATAGTTTTACATATTTTATAAAGTTTTTGAATAAGAAGAATGATCGTC  
 AAATATTTAGAAACGAGGGGAGTATATTGCTAGTAAATTTTACTGTAAAGTCATGTTTCT  
 TCATGTGTCTTTTTTAGAACATTTCTTCTTCATGTGTTCCAATCAATTTTTTCGATCTCC  
 ACCACCCTACCACAGTTGTATGTGTAGTAGCTAGCTGTTGCACATGCTCATGTTAATCT  
 GAATCTGTCCATGGTTTTTTTGCAGGGCAAGAGGTGATGTGCTACGAGAGCCCAAGGCCAA  
**CCATGGGGATCCACCGCTGGTGTTCGTGCTGTTCCAGCAGCTGGGGCGTCAGACAGTGT** < Hd3a **D**  
**ACGCGCCCGGGTGGCGTCAGAACTTCAACACCAAGGACTTCGCCGAGCTCTACAACCTCG**  
**GCTCGCCGGTCCCGCCGCTACTTCAACTGCCAGCGCAGGCAGGCTCCGGCGGCAGGA**  
**GGTCTACCCCTAGCTAACGATGATCCCGATCGATCTGCTGCATGCTCACTATCATCATC**  
**CAGCATGCTATACATTGCAGTTTCAGCAATTTGAAATGATTTCTCGACACACAACATATAT**  
**ATGATGGTGTAAATTAATTATGCAATTTAAATAGCTGAGCAAGGCTAAGGTCTCTCATGAAG**  
 CTAGCTTGTCTATATATATATATATATATATAGTGAAGTGTGCAATAAGCTGCAAG  
 TATATAAGACTGGATTTAAGGAGCTAATTAACATAACCATGCATCATGTATATGTGAGAT  
 GCAGTCGTGCTCTTGCATCCAAGGATACATACAGCATATGCATGCAATATTGGTATCATA  
 TGCAGCGGCTGCTAGCTACTCCTAGATCAAGCTATACAGCAGAATGTAAACATATATGTG  
 TAGCAGTACTTATACGTAGCACATCGATCTGCACATGTTGGATGTAGCAAGAAATTGCTA  
 TGAATATAAGTAGAGATGTGCTTAATATCAAATGTGTATCACATGTGAATCTATCAGCT  
 GGGCATATATAGTGTCTTTTCCAGGCTTCCATCTCACTGTCTATCTCGCCCCCACTGAA  
 TATGCCAGTTTAAAGGACAGGTACTACGTGCTGATGATATGATATGATGGTGAAGGGAC  
 TACAGCTAGGGAGTATCAATCAGAACCCATGTTTTGGAAAAGCTGATACGGATCGATATT  
 GATCATATGCATGCGTAGAAGAATTCATCGTCAGAAAACGCGTACGATT

## 2. Alineación de las secuencias promotoras de *Hd3a*:

-1995

Rómolo	-----TATAGTACTTTTAAATTCGTCTTCGAGATACGACTTTAATCATTAATTGCC	50
NIPPON	AACACCGTAGTATAGTACTTTTAAATTCGTCTTCGAGATACGACTTTAATCATTAATTGCC	60
IR64	----CCGTAGTATAGTACTTTTAAATTCGTCTTCGAGATACGACTTTAATCATTAATTGCC	56
Kasalath	-----AAGTACTTTTAAATTCGTCTTCGAGATACGACTTTAATCATTAATTGCC	48
Norin	-----TAAAGTACTTTTAAATTCGTCTTCGAGATACGACTTTAATCATTAATTGCC	49
LTH	-----GTACTTTTAAATTCGTCTTCGAGATACGACTTTAATCATTAATTGCC	45
T65	-----TATAGTACTTTTAAATTCGTCTTCGAGATACGACTTTAATCATTAATTGCC	49
	*****	
Rómolo	TTACCTCAACTTTATGTTTCGGATTACATATCATGTAATATTATAAAAAATCTAATGGTG	110
NIPPON	TTACCTCAACTTTATGTTTCGGATTACATATCATGTAATATTATAAAAAATCTAATGGTG	120
IR64	TTACCTCAACTTTATGTTTCGGATTACATATCATGTAATATTATAAAAAATCTAATGGTG	116
Kasalath	TTACCTCAACTTTATGTTTCGGATTACATATCATGTAATATTATAAAAAATCTAATGGTG	108
Norin	TTACCTCAACTTTATGTTTCGGATTACATATCATGTAATATTATAAAAAATCTAATGGTG	109
LTH	TTACCTCAACTTTATGTTTCGGATTACATATCATGTAATATTATAAAAAATCTAATGGTG	105
T65	TTACCTCAACTTTATGTTTCGGATTACATATCATGTAATATTATAAAAAATCTAATGGTG	109
	*****	
Rómolo	TAATAAATGTTATAAAATAAACTTAGATGCATAGTTCCGGATATATGCCCTCCGGTTTT	170
NIPPON	TAATAAATGTTATAAAATAAACTTAGATGCATAGTTCCGGATATATGCCCTCCGGTTTT	180
IR64	TAATAAATGTTATAAAATAAACTTAGATGCATAGTTCCGGATATATGCCCTCCGGTTTT	176
Kasalath	TAATAAATGTTATAAAATAAACTTAGATGCATAGTTCCGGATATATGCCCTCCGGTTTT	168
Norin	TAATAAATGTTATAAAATAAACTTAGATGCATAGTTCCGGATATATGCCCTCCGGTTTT	169
LTH	TAATAAATGTTATAAAATAAACTTAGATGCATAGTTCCGGATATATGCCCTCCGGTTTT	165
T65	TAATAAATGTTATAAAATAAACTTAGATGCATAGTTCCGGATATATGCCCTCCGGTTTT	169
	*****	

-1824

Rómolo TAATGTACAACACTTAAATTTTGGATGTGTGTTTCGTTTTGAATCAAATCTGAAATAGTAT 230
NIPPON TAATGTACAACACTTAAATTTTGGATGTGTGTTTCGTTTTGAATCAAATCTGAAATAGTAT 240
IR64 TAATGTACAACACTTAAATTTTGGATGTGTGTTTCGTTTTGAATCAAATCTGAAATAGTAT 236
Kasalath TAATGTACGACACTTAAATTTTGGATGTGTGTTTCGTTTTGAATCAAATCTGAAATAGTAT 228
Norin TAATGTACGACACTTAAATTTTGGATGTGTGTTTCGTTTTGAATCAAATCTGAAATAGTAT 229
LTH TAATGTACAACACTTAAATTTTGGATGTGTGTTTCGTTTTGAATCAAATCTGAAATAGTAT 225
T65 TAATGTACGACACTTAAATTTTGGATGTGTGTTTCGTTTTGAATCAAATCTGAAATAGTAT 229
\*\*\*\*\*

-1757

Rómolo GTTACTTTTAAAGTGCCTTTAAAGATAAAAACAAATCACAATAAAAATACATATATTTTTTA 290
NIPPON GTTACTTTTAAAGTGCCTTTAAAGATAAAAACAAATCACAATAAAAATACATATATTTTTTA 300
IR64 GTTACTTTTAAAGTGCCTTTAAAGATAAAAACAAATCACAATAAAAATACATATATTTTTTA 296
Kasalath GTTACTTTTAAAGTGCCTTTAAAGATAAAAACAAATCACAATAAAAATACATATATTTTTTA 288
Norin GTTACTTTTAAAGTGCCTTTAAAGATAAAAACAAATCACAATAAAAATACATATATTTTTTA 289
LTH GTTACTTTTAAAGTGCCTTTAAAGATAAAAACAAATCACAATAAAAATACATATATTTTTTA 285
T65 GTTACTTTTAAAGTGCCTTTAAAGATAAAAACAAATCACAATAAAAATACATATATTTTTTA 289
\*\*\*\*\*

-1693

Rómolo AATAAGACAGATGATCAAGCATATATTCAAAGTCAACAGATCAGTCATACATTATTAGAA 350
NIPPON AATAAGACAGATGATCAAGCATATATTCAAAGTCAACAGATCAGTCATACATTATTAGAA 360
IR64 AATAAGACAGATGATCAAGCATATATTCAAAGTCAACAGATCAGTCATACATTATTAGAA 356
Kasalath AATAAGACAGATGATCAAGCATATATTCAAAGTCAACAGATCAGTCATACATTATTAGAA 348
Norin AATAAGACAGATGATCAAGCATATATTCAAAGTCAACAGATCAGTCATACATTATTAGAA 349
LTH AATAAGACAGATGATCAAGCATATATTCAAAGTCAACAGATCAGTCATACATTATTAGAA 345
T65 AATAAGACAGATGATCAAGCATATATTCAAAGTCAACAGATCAGTCATACATTATTAGAA 349
\*\*\*\*\*

-1607

Rómolo GAGAAATTTCTTGTATGCCCTTGAATTTAAGTCAATCTTTTCTATGCCCTTGAATTTTAC 410
NIPPON GAGAAATTTCTTGTATGCCCTTGAATTTAAGTCAATCTTTTCTATGCCCTTGAATTTTAC 420
IR64 GAGAAATTTCTTGTATGCCCTTGAATTTAAGTCAATCTTTTCTATGCCCTTGAATTTTAC 416
Kasalath GAGAAATTTCTTGTATGCCCTTGAATTTAAGTCAATCTTTTCTATGCCCTTGAATTTTAC 408
Norin GAGAAATTTCTTGTATGCCCTTGAATTTAAGTCAATCTTTTCTATGCCCTTGAATTTTAC 409
LTH GAGAAATTTCTTGTATGCCCTTGAATTTAAGTCAATCTTTTCTATGCCCTTGAATTTTAC 405
T65 GAGAAATTTCTTGTATGCCCTTGAATTTAAGTCAATCTTTTCTATGCCCTTGAATTTTAC 409
\*\*\*\*\*

-1585

Rómolo TCACTCTCTTGTATAGTTGTATACCTTTTAATTATGATTTTCATCTCTTTTACATCTTTTC 470
NIPPON TCACTCTCTTGTATAGTTGTATACCTTTTAATTATGATTTTCATCTCTTTTACATCTTTTC 480
IR64 TCACTCTCTTGTATAGTTGTATACCTTTTAATTATGATTTTCATCTCTTTTACATCTTTTC 476
Kasalath TCACTCTCTTGTATAGTTGTATACCTTTTAATTATGATTTTCATCTCTTTTACATCTTTTC 468
Norin TCACTCTCTTGTATAGTTGTATACCTTTTAATTATGATTTTCATCTCTTTTACATCTTTTC 469
LTH TCACTCTCTTGTATAGTTGTATACCTTTTAATTATGATTTTCATCTCTTTTACATCTTTTC 465
T65 TCACTCTCTTGTATAGTTGTATACCTTTTAATTATGATTTTCATCTCTTTTACATCTTTTC 469
\*\*\*\*\*

-1506

-1474

Rómolo CGTTAGTTGACCGGTTCCCTTTAAAAATTAGAAGTTTGTGTACGAGACTATTATATTTCT 530
NIPPON CGTTAGTTGACCGGTTCCCTTTAAAAATTAGAAGTTTGTGTACGAGACTATTATATTTCT 540
IR64 CGTTAGTTGACCGGTTCCCTTTAAAAATTAGAAGTTTGTGTACGAGACTATTATATTTCT 536
Kasalath CGTTAGTTGACCGGTTCCCTTTAAAAATTAGAAGTTTGTGTACGAGACTATTATATTTCT 528
Norin CGTTAGTTGACCGGTTCCCTTTAAAAATTAGAAGTTTGTGTACGAGACTATTATATTTCT 529
LTH CGTTAGTTGACCGGTTCCCTTTAAAAATTAGAAGTTTGTGTACGAGACTATTATATTTCT 525
T65 CGTTAGTTGACCGGTTCCCTTTAAAAATTAGAAGTTTGTGTACGAGACTATTATATTTCT 529
\*\*\*\*\*

Rómolo GAGTTTGTGTGCGATGTACTATTTACAAAAAGAAATTAGATGTGCAAGAAAAATAATAT 590
NIPPON GAGTTTGTGTGCGATGTACTATTTACAAAAAGAAATTAGATGTGCAAGAAAAATAATAT 600
IR64 GAGTTTGTGTGCGATGTACTATTTACAAAAAGAAATTAGATGTGCAAGAAAAATAATAT 596
Kasalath GAGTTTGTGTGCGATGTACTATTTACAAAAAGAAATTAGATGTGCAAGAAAAATAATAT 588
Norin GAGTTTGTGTGCGATGTACTATTTACAAAAAGAAATTAGATGTGCAAGAAAAATAATAT 589
LTH GAGTTTGTGTGCGATGTACTATTTACAAAAAGAAATTAGATGTGCAAGAAAAATAATAT 585
T65 GAGTTTGTGTGCGATGTACTATTTACAAAAAGAAATTAGATGTGCAAGAAAAATAATAT 589
\*\*\*\*\*

-1401

Rómolo	TTATTTAAAACAATATGTGTAGCATATTTTTGTGATGTAACATATTTGCCACCTATTGTA	650
NIPPON	TTATTTAAAACAATATGTGTAGCATATTTTTGTGATGTAACATATTTGCCACCTATTGTA	660
IR64	TTATTTAAAACAATATGTGTAGCATATTTTTGTGATGTAACATATTTGCCACCTATTGTA	656
Kasalath	TTATTTAAAACGATATGTGTAGCATATTTTTGTGATGTAACATATTTGCCACCTATTGTA	648
Norin	TTATTTAAAACGATATGTGTAGCATATTTTTGTGATGTAACATATTTGCCACCTATTGTA	649
LTH	TTATTTAAAACAATATGTGTAGCATATTTTTGTGATGTAACATATTTGCCACCTATTGTA	645
T65	TTATTTAAAACGATATGTGTAGCATATTTTTGTGATGTAACATATTTGCCACCTATTGTA	649

\*\*\*\*\*

Rómolo	TATATTATCTAAAAGAAATCCTATACAAGACCAATTGTTTTTTTCAATTATAAAAAAGC	710
NIPPON	TATATTATCTAAAAGAAATCCTATACAAGACCAATTGTTTTTTTCAATTATAAAAAAGC	720
IR64	TATATTATCTAAAAGAAATCCTATACAAGACCAATTGTTTTTTTCAATTATAAAAAAGC	716
Kasalath	TATATTATCTAAAAGAAATCCTATACAAGACCAATTGTTTTTTTCAATTATAAAAAAGC	708
Norin	TATATTATCTAAAAGAAATCCTATACAAGACCAATTGTTTTTTTCAATTATAAAAAAGC	709
LTH	TATATTATCTAAAAGAAATCCTATACAAGACCAATTGTTTTTTTCAATTATAAAAAAGC	705
T65	TATATTATCTAAAAGAAATCCTATACAAGACCAATTGTTTTTTTCAATTATAAAAAAGC	709

\*\*\*\*\*

-1237

Rómolo	CGCACAAAGCTTTTTTCATGAATATATTTAAATTTTCTAAGTACAATACTTCCATAGAAAA	770
NIPPON	CGCACAAAGCTTTTTTCATGAATATATTTAAATTTTCTAAGTACAATACTTCCATAGAAAA	780
IR64	CGCACAAAGCTTTTTTCATGAATATATTTAAATTTTCTAAGTACAATACTTCCATAGAAAA	776
Kasalath	CGCACAAAGCTTTTTTCATGAATATATTTAAATTTTCTAAGTACAATACTTCCATAGAAAA	768
Norin	CGCACAAAGCTTTTTTCATGAATATATTTAAATTTTCTAAGTACAATACTTCCATAGAAAA	769
LTH	CGCACAAAGCTTTTTTCATGAATATATTTAAATTTTCTAAGTACAATACTTCCATAGAAAA	765
T65	CGCACAAAGCTTTTTTCATGAATATATTTAAATTTTCTAAGTACAATACTTCCATAGAAAA	769

\*\*\*\*\*

Rómolo	AACTTATAATATTATAGTCCCACAATAAATTAATTTACTATATTTCCAACGTTAGCATCCACAA	830
NIPPON	AACTTATAATATTATAGTCCCACAATAAATTAATTTACTATATTTCCAACGTTAGCATCCACAA	840
IR64	AACTTATAATATTATAGTCCCACAATAAATTAATTTACTATATTTCCAACGTTAGCATCCACAA	836
Kasalath	AACTTATAATATTATAGTCCCACAATAAATTAATTTACTATATTTCCAACGTTAGCATCCACAA	828
Norin	AACTTATAATATTATAGTCCCACAATAAATTAATTTACTATATTTCCAACGTTAGCATCCACAA	829
LTH	AACTTATAATATTATAGTCCCACAATAAATTAATTTACTATATTTCCAACGTTAGCATCCACAA	825
T65	AACTTATAATATTATAGTCCCACAATAAATTAATTTACTATATTTCCAACGTTAGCATCCACAA	829

\*\*\*\*\*

Rómolo	AATTTCTAGGCATGTACAATTGTAAAATGACCTTTTTTACAGGCATTAACAGTTAACTAA	890
NIPPON	AATTTCTAGGCATGTACAATTGTAAAATGACCTTTTTTACAGGCATTAACAGTTAACTAA	900
IR64	AATTTCTAGGCATGTACAATTGTAAAATGACCTTTTTTACAGGCATTAACAGTTAACTAA	896
Kasalath	AATTTCTAGGCATGTACAATTGTAAAATGACCTTTTTTACAGGCATTAACAGTTAACTAA	888
Norin	AATTTCTAGGCATGTACAATTGTAAAATGACCTTTTTTACAGGCATTAACAGTTAACTAA	889
LTH	AATTTCTAGGCATGTACAATTGTAAAATGACCTTTTTTACAGGCATTAACAGTTAACTAA	885
T65	AATTTCTAGGCATGTACAATTGTAAAATGACCTTTTTTACAGGCATTAACAGTTAACTAA	889

\*\*\*\*\*

-1107

Rómolo	CGGTA-CGGAAATGGTATAAAAACGGGTAAAATTTAAAATTCAGAGGTATATAAAAAATGA	949
NIPPON	CGGTA-CGGAAATGGTATAAAAACGGGTAAAATTTAAAATTCAGAGGTATATAAAAAATGA	959
IR64	CGGTA-CGGAAATGGTATAAAAACGGGTAAAATTTAAAATTCAGAGGTATATAAAAAATGA	954
Kasalath	CGGTA-CGGAAATGGTATAAAAACGGGTAAAATTTAAAATTCAGAGGTATATAAAAAATGA	947
Norin	CGGTA-CGGAAATGGTATAAAAACGGGTAAAATTTAAAATTCAGAGGTATATAAAAAATGA	948
LTH	CGGTA-CGGAAATGGTATAAAAACGGGTAAAATTTAAAATTCAGAGGTATATAAAAAATGA	944
T65	CGGTA-CGGAAATGGTATAAAAACGGGTAAAATTTAAAATTCAGAGGTATATAAAAAATGA	948

\*\*\*\*\*

Rómolo	GCGATTAGCTAAACTTTTAGGAGTATAAAAAGAAATTCCTTACTGAAAACAGAGGGAGT	1009
NIPPON	GCGATTAGCTAAACTTTTAGGAGTATAAAAAGAAATTCCTTACTGAAAACAGAGGGAGT	1019
IR64	GCGATTAGCTAAACTTTTAGGAGTATAAAAAGAAATTCCTTACTGAAAACAGAGGGAGT	1012
Kasalath	GCGATTAGCTAAACTTTTAGGAGTATAAAAAGAAATTCCTTACTGAAAACAGAGGGAGT	1006
Norin	GCGATTAGCTAAACTTTTAGGAGTATAAAAAGAAATTCCTTACTGAAAACAGAGGGAGT	1008
LTH	GCGATTAGCTAAACTTTTAGGAGTATAAAAAGAAATTCCTTACTGAAAACAGAGGGAGT	1004
T65	GCGATTAGCTAAACTTTTAGGAGTATAAAAAGAAATTCCTTACTGAAAACAGAGGGAGT	1008

\*\*\*\*\*

Rómolo ACTACATATAACCTCATTTTTTTACGGGTTGATAAGTTGCGGAAAAACCAACAAATTAGC 1066  
 NIPPON ACTACATATAACCTCATTTTTTTACGGGTTGATAAGTTGCGGAAAAACCAACAAATTAGC 1079  
 IR64 ACTACATATAACCTCATTTTTTTACGGGTTGATAAGTTGCGGAAAAACCAACAAATTAGC 1072  
 Kasalath ACTACATATAACCTCATTTTTTTACGGGTTGATAAGTTGCGGAAAAACCAACAAATTAGC 1066  
 Norin ACTACATATAACCTCATTTTTTTACGGGTTGATAAGTTGCGGAAAAACCAACAAATTAGC 1068  
 LTH ACTACATATAACCTCATTTTTTTACGGGTTGATAAGTTGCGGAAAAACCAACAAATTAGC 1063  
 T65 ACTACATATAACCTCATTTTTTTACGGGTTGATAAGTTGCGGAAAAACCAACAAATTAGC 1068  
 \*\*\*\*\*

Rómolo AACAAATATGAGTAAACTTGTATACATGTGTCTTCTTAGCGATTTAAAAATCAATGCTG 1123  
 NIPPON AACAAATATGAGTAAACTTGTATACATGTGTCTTCTTAGCGATTTAAAAATCAATGCTG 1139  
 IR64 AACAAATATGAGTAAACTTGTATACATGTGTCTTCTTAGCGATTTAAAAATCAATGCTG 1132  
 Kasalath AACAAATATGAGTAAACTTGTATACATGTGTCTTCTTAGCGATTTAAAAATCAATGCTG 1126  
 Norin AACAAATATGAGTAAACTTGTATACATGTGTCTTCTTAGCGATTTAAAAATCAATGCTG 1128  
 LTH AACAAATATGAGTAAACTTGTATACATGTGTCTTCTTAGCGATTTAAAAATCAATGCTG 1123  
 T65 AACAAATATGAGTAAACTTGTATACATGTGTCTTCTTAGCGATTTAAAAATCAATGCTG 1128  
 \*\*\*\*\*

-815

Rómolo AAAATAAATTACAATAAAATTA AAAATCTCAAGATAATCTCTAAAATGTAGTTTTAA 1181  
 NIPPON AAAATAAATTACAATAAAATTA AAAATCTCAAGATAATCTCTAAAATGTAGTTTTAA 1197  
 IR64 AAAATAAATTACAATAAAATTA AAAATCTCAAGATAATCTCTAAAATGTAGTTTTAA 1190  
 Kasalath AAAATAAATTACAATAAAATTA AAAATCTCAAGATAATCTCTAAAATGTAGTTTTAA 1183  
 Norin AAAATAAATTACAATAAAATTA AAAATCTCAAGATAATCTCTAAAATGTAGTTTTAA 1188  
 LTH AAAATAAATTACAATAAAATTA AAAATCTCAAGATAATCTCTAAAATGTAGTTTTAA 1181  
 T65 AAAATAAATTACAATAAAATTA AAAATCTCAAGATAATCTCTAAAATGTAGTTTTAA 1186  
 \*\*\*\*\*

# ANEXO III: *RFT1*

## 1. Secuencia completa de gen *RFT1* y su región promotora: (cv. Nipponbare)

>6 dna:chromosome chromosome:MSU6:6:2923724:2927575:1

CAAACACTACTAAAATGTATTCAATTGTGGAGTCAATGAAGCTAGTGATTTCGGTGGGCGT  
AACTGGTATGAATTCTCCTCTGGCTATATATGCAATTAATCCACATTCAAACAATTAGTT  
TTTCAATAACAGAAATCAACCCAAGCCGCTCTACATCATAGTGAATATACATCATAATGA  
ATAATGAGTACTAGCCTTTTTCATCATCTAAAAAGACAGGAATAACATTTGGCTGGCCGGT  
CGGCTATCGCGAATGCAATCTGATATTCTCGCACCCAGTCTTGCTCTGTCCTGGTGGCCA  
TCCCTGTCCAATAATGCTTAGCCATGCATCCATCAAGAAAATTCCTTTATCAATTCATCT  
AGTATATATTCCTTTTTGAAAGTTTGCATTTCAAATGTACGACACCATTAATTTTTTCGC  
ATAACGTTTGACCATTTGTCTTATTAATATATATATAAATATCATTACTCTCTCTGTT  
TCATATTATAAGTCGTTTTTTTTCTTAGTTAAACTTTTTTAATTATGTTTAACCAAGTTT  
ATAGAAAAAATTAGCAACATTTATGATACCAAAC TAGTTTTATTAATTTAACACTGTAT  
ATACTTTGATAATATGTTTGTCTTATTTAAAATATTACTATATTTATCTTTAAACTTG  
ATCAAGTTTAAAGAAAAGTTTGACTAAGGAAAAAGTCAAATCTTATAATATAATAAAAA  
CGTAGGGAGTATTTTGTGTTACTTATTTTATCATCAAACAAAATATAATCTCAACTTAT  
ACTTTACATATAATAATAATATATAATAAATATTCATTTATTTTGTGTGTTTAGCTTTTCT  
GCCGCCGGCCGGAATAATGTCT**AGCCTTTCTGTTTGTATTGCATC**ATGCATTTTGATA  
TCTGCAAGTCGATCTGGAAGCTAGTTCCTCCACATATGCATGCATGCTTTGCTTCTCT  
GGCCTGGCCACCACACCTTTCAAGACAAATCTATGGATATGTACTATCTTTTTTTCTTAT  
TTTTTTTAACTAAGCATTTTGAGAAGTTTCAAAGTTATCGATCTCAGAATTTTTTAAAA  
GCTTAACCAAACATTGTATGAGTTTTGGAATCGGAAATATACCGAACACACCCGCAATTG  
TAGAATAGGTATGGTGAATAATTAATGAGAGATGAAGCTTGTAACCGTTAGGCCAGCTA  
TAAGTTTGGCCATGCCAGATCAGCTACAGTACCATCTGATTCTATCGTAACAATAGAATC  
TACATTGTACCCTAGAGTATGTAATTTAATTGTGCACAACAATATTAGTAGCTATAGGGC  
CTGTTTGGGGGAGCTTCTAGCTGCATCAACTTCTCCAGAATCAAAGCTCTCAAATAG  
TTCAGCTGTTGGTCCAGATTTCAAAGCTATAGTTATAGAATTTATAAAAGAACTAAAA  
GCCAGAAGCTAGGTAACCAGCTTTTCCAAACTCTCAGAGCCTGGTTACCAAACAGCTGTT  
CTTAACTCTGCCAAACATGCCATAGTTAGCTTCAATTTTTACTGCTAGCTTCAATTTAT  
TTTGGACGGGTTAGTACATATGTATCTGTCTCGAAATCGCCTCTGTAAATGGGGTCAGT  
TTCAGCACTTCTATCGACATGAAGAAGATAGAAAGGGCCATGCGAGAGAGAGTGTGTGAG  
AAAGAGAAGATTCTACCGGCACAACAAGTGTGTAATGCATAAAGAAAAGATTCTACT  
GGGCACATCATAGTAAGGCTATTTCTCAGCTCAATAACTAGGATGGTGTCCATAGCATTAA  
ATAAGCTGCTACCTAATATGAAAAATGATGTGGCAAGTGAGTAAATGAGGAAAGAGAAGA  
AAACAAGGCTTGCATGAGACATGGTTTCTACACAACATTTGAAGACATCATGTGAGATAA  
ATTTAATTTAAGCATGAAAAGTGGTGTTTGTATTGAAAAGTAGTGTCTAATACTAGTT  
TATTGATGATGTGAAGTTTATGAAAACACCCTGTAGTGTCTTGGGTTGGTACTGCCTAA  
AGGGCCCGGCTAGCTATATAAACTGTACAGATGCAGAGAAGCTGCACAACACTAGTGCATC  
**CCTGCTACCTGTTTGGCTAGCTTAACCTTCTGACATCTATCCTCTGGATTGAACGGCAGG**  
**AGATACCTAAGCTAGCTAGCAATCTCTATCGATCTGTTTGTTTACATGTTTCAGTTAAAGG**  
**TTACTGAGAAATGCCTAGAGTTTTTCCGGCTAGCTTCATAAGTTAGTGGGTTAGCTGACC**  
**TAGATTCAAAGTCTAATCCTTTTATTTATTTGATATTAGATATCCTAACGTTTTTATGTTA**  
**GAGGTTATTAATTTGACATGGCCG**GCAGCGGCAGGGACGATCCTCTTGTGGTTGGCAGGA  
**TTGTGGGTGATGTGCTGGATCCATTCGTCGGATCACTAACCTCAGTGTGAGCTATGGTG**  
**CAAGGATCGTCTCCAATGGCTGCGAGCTCAAGCCGTCCATGGTGACCCAACAGCCCAGGG**  
**TCGTGGTGGTGGCAATGACATGAGGACGTTCTACACACTC**GTACGGATCATATCTTGGA  
TGCAGAGACCCACCAGAAGTTATTTAATTACTTTTCATTAATTATCATAAAACTAGACTAT  
AAATTATATTTTTTACATGGATGCATGTTAATTTTGTGTGGCTTACGTAATACTAATT  
ACCTACAG**GTGATGGTAGACCCGGATGCTCCGAGCCAAGCAACCCTAACCTTAGGGAGT**  
**ATCTACACTG**TAGGCACCGATCAGATATGTTAGCTAGCTAATTGTATACATTTCGTCTTA  
GAATATAAAGAATTTGGACCGGATTTCTCAATCATTCGCGTTGATTCTCTTTGTCTGTCTG  
TAG**GCTGGTCACCGATATTTCCGGTACCCTGGAGCAACATTTG**GTGTCAGTAACTAGTAT  
ATATATATAGTACTCATATCAATTTTCGATGTAACAGCAACATATGTGGCAGTTCCATGAA  
TTTTTATTACCTTGGTCTACCCCATATATATACTTCAAATAATGCATAATGACAAAATT

>RFT1 **H**

< RFT1 **I**

ATACTCCCTCCGTATTTTAAATGTATGACGCCGTTGACTTTTTTAACCAACATTTGACCATT  
 CGTCTTATTTAAATTTTTTATGCAAATACAAAAATACTTATGTTCATGCTTAAAGAACATT  
 TGATGATAAAATCAAGTCACAATAAAAATAAATGATAATTACATAATTTTTTTGAATAAGAC  
 AAAAAGTCAAACGTTTGTAAAAAGTCAATGTCGTCATACATTAAAATACGGAGGGAGTA  
 TCTATTTTGTCAAAAATTTTCGTCAGATTTGAAGGATAGGGCTGTACTTCTTCTATGCCA  
 AAATGGAGGGTGGTGTATCCCTCAAAACTTGTAGAACAGCCACTGTTGATATATATATGG  
 TGAATAACGTAGATAAATAAATTGATGCAGGGCAAGAGGTGATGTGCTACGAGAGCCCAA  
**GGCCAACCATGGGGATCCACCGCTGGTGTTCGTCGCTTCCAGCAGCTGGGGCGTCAGA**  
**CGGTGTACGCACCGGGTGGCGCCAGAACTTCAGCACCAGGAACCTCGCCGAGCTCTACA**  
**ACCTCGGCTCGCCGGTCCGACCGTCTACTTCAACTGCCAGCGCGAGGCCGGCTCCGGCG**  
**GCAGGAGGGTCTACCCCTAGCTACGCTACGCATGCCACCGGCTCCATGCATGCAGCAGC**  
**TATAGCTAAGCTGAGACCTGCCCTAGCTGTATA**TGCATGCAGGCTCTTAAGTGTGTATATA  
 TAAATAAATAAATGTATATCTACTATTATAAAAATTATAAAAATATAAAGATGCTAGCGTT  
 CATGCATTAGCT

**2. Alineación de las secuencias promotoras de *RFT1*:**

kasalath	-----	-----TA	35
AZUCENA	-----	-----NNNNNNNNNNNNNNNNNNNNNTCTGGAGCTA	30
KINMAZE	-----	-----NNNNNNNNNNNNNNNNNNNNCTGCAAGTCGATCTGGAAGCTA	35
T65	-----	-----TTTTGCATATCTGCAAGTCGATCTGGAAGCTA	33
Giza	-----	-----CATGCATTTTGCATATCTGCAAGTCGATCTGGAAGCTA	38
Rómolo	-----	-----ATTTTGCATATCTGCAAGTCGATCTGGAAGCTA	33
NIPPON	AGCCTTTCTGTTTGTATTGCATCATGCATTTTGCATATCTGCAAGTCGATCTGGAAGCTA		60
			**
kasalath	GTTCTCCTCCACATATGCATGCATGCTTTGCTTCTCTGGCCTGGCCACCACACCTTTCAA		95
AZUCENA	GTTCTCCTCCACATATGCATGCATGCTTTGCTTCTCTGGCCTGGCCACCACACCTTTCAA		90
KINMAZE	GTTCTCCTCCACATATGCATGCATGCTTTGCTTCTCTGGCCTGGCCACCACACCTTTCAA		95
T65	GTTCTCCTCCACATATGCATGCATGCTTTGCTTCTCTGGCCTGGCCACCACACCTTTCAA		93
Giza	GTTCTCCTCCACATATGCATGCATGCTTTGCTTCTCTGGCCTGGCCACCACACCTTTCAA		98
Rómolo	GTTCTCCTCCACATATGCATGCATGCTTTGCTTCTCTGGCCTGGCCACCACACCTTTCAA		93
NIPPON	GTTCTCCTCCACATATGCATGCATGCTTTGCTTCTCTGGCCTGGCCACCACACCTTTCAA		120
	*****		
kasalath	GACAAATCTATGGATATGTAATCTTTTTTCTTATTTTTTTTAACTAAGCATTTTGA		155
AZUCENA	GACAAATCTATGGATATGTAATCTTTTTTCTTATTTTTTTTAACTAAGCATTTTGA		150
KINMAZE	GACAAATCTATGGATATGTAATCTTTTTTCTTATTTTTTTTAACTAAGCATTTTGA		155
T65	GACAAATCTATGGATATGTAATCTTTTTTCTTATTTTTTTTAACTAAGCATTTTGA		153
Giza	GACAAATCTATGGATATGTAATCTTTTTTCTTATTTTTTTTAACTAAGCATTTTGA		158
Rómolo	GACAAATCTATGGATATGTAATCTTTTTTCTTATTTTTTTTAACTAAGCATTTTGA		153
NIPPON	GACAAATCTATGGATATGTAATCTTTTTTCTTATTTTTTTTAACTAAGCATTTTGA		180
	*****		
kasalath	GAAGTTTCAAAGTTATCGATCTCAGAATTTTAAAAGCTTAACCAAACATTGTATGAGT		215
AZUCENA	GAAGTTTCAAAGTTATCGATCTCAGAATTTTAAAAGCTTAACCAAACATTGTATGAGT		210
KINMAZE	GAAGTTTCAAAGTTATCGATCTCAGAATTTTAAAAGCTTAACCAAACATTGTATGAGT		215
T65	GAAGTTTCAAAGTTATCGATCTCAGAATTTTAAAAGCTTAACCAAACATTGTATGAGT		213
Giza	GAAGTTTCAAAGTTATCGATCTCAGAATTTTAAAAGCTTAACCAAACATTGTATGAGT		218
Rómolo	GAAGTTTCAAAGTTATCGATCTCAGAATTTTAAAAGCTTAACCAAACATTGTATGAGT		213
NIPPON	GAAGTTTCAAAGTTATCGATCTCAGAATTTTAAAAGCTTAACCAAACATTGTATGAGT		240
	*****		
kasalath	TTTGGAAATCGGAAATATACCGAACACACCGCGAATTGTAGAATAGGTATGGTGAAAAATT		275
AZUCENA	TTTGGAAATCGGAAATATACCGAACACACCGCGAATTGTAGAATAGGTATGGTGAAAAATT		270
KINMAZE	TTTGGAAATCGGAAATATACCGAACACACCGCGAATTGTAGAATAGGTATGGTGAAAAATT		275
T65	TTTGGAAATCGGAAATATACCGAACACACCGCGAATTGTAGAATAGGTATGGTGAAAAATT		273
Giza	TTTGGAAATCGGAAATATACCGAACACACCGCGAATTGTAGAATAGGTATGGTGAAAAATT		278
Rómolo	TTTGGAAATCGGAAATATACCGAACACACCGCGAATTGTAGAATAGGTATGGTGAAAAATT		273
NIPPON	TTTGGAAATCGGAAATATACCGAACACACCGCGAATTGTAGAATAGGTATGGTGAAAAATT		300
	*****		
		-1149	
kasalath	AATGAGAGATGAAGCTTGTAAAACGTTAGGCCAGCTATAAGTTTTGCCATGCCAGATCAG		335
AZUCENA	AATGAGAGATGAAGCTTGTAAAACGTTAGGCCAGCTATAAGTTTTGCCATGCCAGATCAG		330
KINMAZE	AATGAGAGATGAAGCTTGTAAAACGTTAGGCCAGCTATAAGTTTTGCCATGCCAGATCAG		335
T65	AATGAGAGATGAAGCTTGTAAAACGTTAGGCCAGCTATAAGTTTTGCCATGCCAGATCAG		333
Giza	AATGAGAGATGAAGCTTGTAAAACGTTAGGCCAGCTATAAGTTTTGCCATGCCAGATCAG		338
Rómolo	AATGAGAGATGAAGCTTGTAAAACGTTAGGCCAGCTATAAGTTTTGCCATGCCAGATCAG		333
NIPPON	AATGAGAGATGAAGCTTGTAAAACGTTAGGCCAGCTATAAGTTTTGCCATGCCAGATCAG		360
	*****		

kasalath CTACAGTACCATCTGATTCTATCGTAACAATAGAATCTACATTGTACCCTAGAGTATGTA 395  
 AZUCENA CTACAGTACCATCTGATTCTATCGTAACAATAGAATCTACATTGTACCCTAGAGTATGTA 390  
 KINMAZE CTACAGTACCATCTGATTCTATCGTAACAATAGAATCTACATTGTACCCTAGAGTATGTA 395  
 T65 CTACAGTACCATCTGATTCTATCGTAACAATAGAATCTACATTGTACCCTAGAGTATGTA 393  
 Giza CTACAGTACCATCTGATTCTATCGTAACAATAGAATCTACATTGTACCCTAGAGTATGTA 398  
 Rómolo CTACAGTACCATCTGATTCTATCGTAACAATAGAATCTACATTGTACCCTAGAGTATGTA 393  
 NIPPON CTACAGTACCATCTGATTCTATCGTAACAATAGAATCTACATTGTACCCTAGAGTATGTA 420  
 \*\*\*\*\*

kasalath ATTTAATTGTGCACAACAATATTAGTAGCTATAGGGCCTGTTGGGGGAGCTTCTAGCTG 455  
 AZUCENA ATTTAATTGTGCACAACAATATTAGTAGCTATAGGGCCTGTTGGGGGAGCTTCTAGCTG 450  
 KINMAZE ATTTAATTGTGCACAACAATATTAGTAGCTATAGGGCCTGTTGGGGGAGCTTCTAGCTG 455  
 T65 ATTTAATTGTGCACAACAATATTAGTAGCTATAGGGCCTGTTGGGGGAGCTTCTAGCTG 453  
 Giza ATTTAATTGTGCACAACAATATTAGTAGCTATAGGGCCTGTTGGGGGAGCTTCTAGCTG 458  
 Rómolo ATTTAATTGTGCACAACAATATTAGTAGCTATAGGGCCTGTTGGGGGAGCTTCTAGCTG 453  
 NIPPON ATTTAATTGTGCACAACAATATTAGTAGCTATAGGGCCTGTTGGGGGAGCTTCTAGCTG 480  
 \*\*\*\*\*

kasalath CATCAACTTCTCCAGAATCAAAGCTCTCAAATAGTTCAGCTGTTGGTCCAGATTCTA 515  
 AZUCENA CATCAACTTCTCCAGAATCAAAGCTCTCAAATAGTTCAGCTGTTGGTCCAGATTCTA 510  
 KINMAZE CATCAACTTCTCCAGAATCAAAGCTCTCAAATAGTTCAGCTGTTGGTCCAGATTCTA 515  
 T65 CATCAACTTCTCCAGAATCAAAGCTCTCAAATAGTTCAGCTGTTGGTCCAGATTCTA 513  
 Giza CATCAACTTCTCCAGAATCAAAGCTCTCAAATAGTTCAGCTGTTGGTCCAGATTCTA 518  
 Rómolo CATCAACTTCTCCAGAATCAAAGCTCTCAAATAGTTCAGCTGTTGGTCCAGATTCTA 513  
 NIPPON CATCAACTTCTCCAGAATCAAAGCTCTCAAATAGTTCAGCTGTTGGTCCAGATTCTA 540  
 \*\*\* \*\*\*\*\*

kasalath AAAAGCTATAGTTATAGAATTTATAAAGAAGCTAAAAGCCAGAAGCTAGGTAACCAGCTT 575  
 AZUCENA AAAAGCTATAGTTATAGAATTTATAAAGAAGCTAAAAGCCAGAAGCTAGGTAACCAGCTT 570  
 KINMAZE AAAAGCTATAGTTATAGAATTTATAAAGAAGCTAAAAGCCAGAAGCTAGGTAACCAGCTT 575  
 T65 AAAAGCTATAGTTATAGAATTTATAAAGAAGCTAAAAGCCAGAAGCTAGGTAACCAGCTT 573  
 Giza AAAAGCTATAGTTATAGAATTTATAAAGAAGCTAAAAGCCAGAAGCTAGGTAACCAGCTT 578  
 Rómolo AAAAGCTATAGTTATAGAATTTATAAAGAAGCTAAAAGCCAGAAGCTAGGTAACCAGCTT 573  
 NIPPON AAAAGCTATAGTTATAGAATTTATAAAGAAGCTAAAAGCCAGAAGCTAGGTAACCAGCTT 600  
 \*\*\*\*\*

kasalath TTCCAAACTCTCAGAGCCTGGTTACCAAACAGCTGTTCTTAACTCTGCCAAACATGCCCA 635  
 AZUCENA TTCCAAACTCTCAGAGCCTGGTTACCAAACAGCTGTTCTTAACTCTGCCAAACATGCCCA 630  
 KINMAZE TTCCAAACTCTCAGAGCCTGGTTACCAAACAGCTGTTCTTAACTCTGCCAAACATGCCCA 635  
 T65 TTCCAAACTCTCAGAGCCTGGTTACCAAACAGCTGTTCTTAACTCTGCCAAACATGCCCA 633  
 Giza TTCCAAACTCTCAGAGCCTGGTTACCAAACAGCTGTTCTTAACTCTGCCAAACATGCCCA 638  
 Rómolo TTCCAAACTCTCAGAGCCTGGTTACCAAACAGCTGTTCTTAACTCTGCCAAACATGCCCA 633  
 NIPPON TTCCAAACTCTCAGAGCCTGGTTACCAAACAGCTGTTCTTAACTCTGCCAAACATGCCCA 660  
 \*\*\*\*\*

kasalath TAGTTAGCTTCATTATTTTACTGCTAGCTTCATTTATTTTGGACGGGGTTAGTACATATG 695  
 AZUCENA TAGTTAGCTTCATTATTTTACTGCTAGCTTCATTTATTTTGGACGGGGTTAGTACATATG 690  
 KINMAZE TAGTTAGCTTCATTATTTTACTGCTAGCTTCATTTATTTTGGACGGGGTTAGTACATATG 695  
 T65 TAGTTAGCTTCATTATTTTACTGCTAGCTTCATTTATTTTGGACGGGGTTAGTACATATG 693  
 Giza TAGTTAGCTTCATTATTTTACTGCTAGCTTCATTTATTTTGGACGGGGTTAGTACATATG 698  
 Rómolo TAGTTAGCTTCATTATTTTACTGCTAGCTTCATTTATTTTGGACGGGGTTAGTACATATG 693  
 NIPPON TAGTTAGCTTCATTATTTTACTGCTAGCTTCATTTATTTTGGACGGGGTTAGTACATATG 720  
 \*\*\*\*\* \*\*

kasalath TATCTGTCTCGAAATCGCCTCTGTAATGGGGTCAGTTTCAGCACTTCTATCGACATGAA 755  
 AZUCENA TATCTGTCTCGAAATCGCCTCTGTAATGGGGTCAGTTTCAGCACTTCTATCGACATGAA 750  
 KINMAZE TATCTGTCTCGAAATCGCCTCTGTAATGGGGTCAGTTTCAGCACTTCTATCGACATGAA 755  
 T65 TATCTGTCTCGAAATCGCCTCTGTAATGGGGTCAGTTTCAGCACTTCTATCGACATGAA 753  
 Giza TATCTGTCTCGAAATCGCCTCTGTAATGGGGTCAGTTTCAGCACTTCTATCGACATGAA 758  
 Rómolo TATCTGTCTCGAAATCGCCTCTGTAATGGGGTCAGTTTCAGCACTTCTATCGACATGAA 753  
 NIPPON TATCTGTCTCGAAATCGCCTCTGTAATGGGGTCAGTTTCAGCACTTCTATCGACATGAA 780  
 \*\*\*\*\*

kasalath GAAGATAGAAAGGGCCATGCGAGAGAGAGTGTGTGAGAAAGAGAAGATTCTACCGGGCAC 815  
 AZUCENA GAAGATAGAAAGGGCCATGCGAGAGAGAGTGTGTGAGAAAGAGAAGATTCTACCGGGCAC 810  
 KINMAZE GAAGATAGAAAGGGCCATGCGAGAGAGAGTGTGTGAGAAAGAGAAGATTCTACCGGGCAC 815  
 T65 GAAGATAGAAAGGGCCATGCGAGAGAGAGTGTGTGAGAAAGAGAAGATTCTACCGGGCAC 813  
 Giza GAAGATAGAAAGGGCCATGCGAGAGAGAGTGTGTGAGAAAGAGAAGATTCTACCGGGCAC 818  
 Rómolo GAAGATAGAAAGGGCCATGCGAGAGAGAGTGTGTGAGAAAGAGAAGATTCTACCGGGCAC 813  
 NIPPON GAAGATAGAAAGGGCCATGCGAGAGAGAGTGTGTGAGAAAGAGAAGATTCTACCGGGCAC 840  
 \*\*\*\*\*

```

kasalath AACAAAGTGTGTAATGCATAAAGAAAAGATTCTACTGGGCACATCATAGTAAGGCTATT 875
AZUCENA AACAAAGTGTGTAATGCATAAAGAAAAGATTCTACTGGGCACATCATAGTAAGGCTATT 870
KINMAZE AACAAAGTGTGTAATGCATAAAGAAAAGATTCTACTGGGCACATCATAGTAAGGCTATT 875
T65 AACAAAGTGTGTAATGCATAAAGAAAAGATTCTACTGGGCACATCATAGTAAGGCTATT 873
Giza AACAAAGTGTGTAATGCATAAAGAAAAGATTCTACTGGGCACATCATAGTAAGGCTATT 878
Rómolo AACAAAGTGTGTAATGCATAAAGAAAAGATTCTACTGGGCACATCATAGTAAGGCTATT 873
NIPPON AACAAAGTGTGTAATGCATAAAGAAAAGATTCTACTGGGCACATCATAGTAAGGCTATT 900
*****
-590
kasalath CTCAGCTCAATAACTAGGATGGTGTCCATAGCATTAAATAAGCTGTACCTAATATGAAA 935
AZUCENA CTCAGCTCAATAACTAGGATGGTGTCCATAGCATTAAATAAGCTGTACCTAATATGAAA 930
KINMAZE CTCAGCTCAATAACTAGGATGGTGTCCATAGCATTAAATAAGCTGTACCTAATATGAAA 935
T65 CTCAGCTCAATAACTAGGATGGTGTCCATAGCATTAAATAAGCTGTACCTAATATGAAA 932
Giza CTCAGCTCAATAACTAGGATGGTGTCCATAGCATTAAATAAGCTGTACCTAATATGAAA 938
Rómolo CTCAGCTCAATAACTAGGATGGTGTCCATAGCATTAAATAAGCTGTACCTAATATGAAA 933
NIPPON CTCAGCTCAATAACTAGGATGGTGTCCATAGCATTAAATAAGCTGTACCTAATATGAAA 960
****
-485 -477
kasalath AATGATGTGGCAAGTGAAGTAAATGAGGAAAGAGAGAAAACAAGGCTTTGCATGAGA-CA 994
AZUCENA AATGATGTGGCAAGTGAAGTAAATGAGGAAAGAGAGAAAACAAGGCTTTGCATGAGA-CA 989
KINMAZE AATGATGTGGCAAGTGAAGTAAATGAGGAAAGAGAGAAAACAAGGCTTTGCATGAGA-CA 994
T65 AATGATGTGGCAAGTGAAGTAAATGAGGAAAGAGAGAAAACAAGGCTTTGCATGAGA-CA 991
Giza AATGATGTGGCAAGTGAAGTAAATGAGGAAAGAGAGAAAACAAGGCTTTGCATGAGA-CA 997
Rómolo AATGATGTGGCAAGTGAAGTAAATGAGGAAAGAGAGAAAACAAGGCTTTGCATGAGA-CA 992
NIPPON AATGATGTGGCAAGTGAAGTAAATGAGGAAAGAGAGAAAACAAGGCTTTGCATGAGA-CA 1019
*****
-468
kasalath TGGTTTC-TACACAACATTGAAGACATCATGTGAGATAAAATTTAATTTAAGCATGAAA 1051
AZUCENA TGGTTTCCTACACAACATTGAAGACATCATGTGAGATAAAATTTAATTTAAGCATGAAA 1047
KINMAZE TGGTTTC-TACACAACATTGAAGACATCATGTGAGATAAAATTTAATTTAAGCATGAAA 1053
T65 TGGTTTC-TACACAACATTGAAGACATCATGTGAGATAAAATTTAATTTAAGCATGAAA 1049
Giza TGGTTTC-TACACAACATTGAAGACATCATGTGAGATAAAATTTAATTTAAGCATGAAA 1054
Rómolo TGGTTTC-TACACAACATTGAAGACATCATGTGAGATAAAATTTAATTTAAGCATGAAA 1049
NIPPON TGGTTTC-TACACAACATTGAAGACATCATGTGAGATAAAATTTAATTTAAGCATGAAA 1076
***
-385
kasalath AAGTGGTGTTTGTATTGAAAAAGTAGTGTCTAATACTAGTTTATTGATGATGTGAAGTTT 1111
AZUCENA AAGTGGTGTTTGTATTGAAAAAGTAGTGTCTAATACTAGTTTATTGATGATGTGAAGTTT 1107
KINMAZE AAGTGGTGTTTGTATTGAAAAAGTAGTGTCTAATACTAGTTTATTGATGATGTGAAGTTT 1113
T65 AAGTGGTGTTTGTATTGAAAAAGTAGTGTCTAATACTAGTTTATTGATGATGTGAAGTTT 1109
Giza AAGTGGTGTTTGTATTGAAAAAGTAGTGTCTAATACTAGTTTATTGATGATGTGAAGTTT 1114
Rómolo AAGTGGTGTTTGTATTGAAAAAGTAGTGTCTAATACTAGTTTATTGATGATGTGAAGTTT 1109
NIPPON AAGTGGTGTTTGTATTGAAAAAGTAGTGTCTAATACTAGTTTATTGATGATGTGAAGTTT 1136
*****
-337
kasalath ATGAAAACACTACCGTGTAGTGT--CTTGGGTTGGTACTGCCTAAAGGGGCCGGCTAGCTA 1168
AZUCENA ATGAAAACACTACCGTGTAGTGTCTTCTTGGGTTGGTACTGCCTAAAGGGGCCGGCTAGCTA 1166
KINMAZE ATGAAAACACTACCGTGTAGTGT--CTTGGGTTGGTACTGCCTAAAGGGGNNNNNNANNNN 1170
T65 ATGAAAACACTACCGTGTAGTGT--CTTGGGTTGGTACTGCCTAAAGGGGCCGGCTAGCTA 1166
Giza ATGAAAACACTACCGTGTAGTGT--CTTGGGTTGGTACTGCCTAAAGGGGCCGGCTAGCTA 1172
Rómolo ATGAAAACACTACCGTGTAGTGT--CTTGGGTTGGTACTGCCTAAAGGGGCCGGCTAGCTA 1166
NIPPON ATGAAAACACTACCGTGTAGTGT--CTTGGGTTGGTACTGCCTAAAGGGGCCGGCTAGCTA 1193
*****

```