

TESIS DE MAESTRÍA

PEDECIBA-BIOLOGÍA

Subárea Genética

“Utilización de microsatélites génicos como estrategia para el análisis de tolerancia a frío en arroz”

Eliana Monteverde

MONTEVIDEO

URUGUAY

2011

Orientador: Pablo Speranza

AGRADECIMIENTOS

Quisiera agradecer en primer lugar a mi tutor Pablo Speranza y a la Ing. Agr. Victoria Bonnacarrère por su valiosa dirección y apoyo durante todo mi trabajo.

También quisiera agradecer a mis compañeros de laboratorio de la Unidad de Biotecnología de INIA Las Brujas, particularmente a Silvia Garaycochea y Leticia Chao, así como a Juan Rosas, Pedro Blanco y Fernando Pérez de Vida de INIA Treinta y Tres.

A mi familia y amigos por su apoyo incondicional.

Este trabajo fue apoyado en parte por una beca de Maestría de la Agencia Nacional de Investigación e Innovación (ANII).

RESUMEN

En Uruguay, una de las principales limitantes en la producción de arroz es la ocurrencia de bajas temperaturas tanto durante las etapas tempranas del desarrollo de la planta, como durante la fase reproductiva. La temperatura es un factor determinante en la distribución geográfica de la gran mayoría de las especies cultivables. Existen dos subespecies de arroz cultivado, *indica* y *japonica*, las cuales difieren principalmente en su tolerancia a las bajas temperaturas. La subespecie *indica* es de origen tropical y es fundamentalmente más susceptible al frío que la subespecie *japonica*, de origen templado. Uno de los objetivos del programa de mejoramiento de arroz del INIA (Instituto Nacional de Investigación Agropecuaria) es mejorar el germoplasma disponible de manera de poder desarrollar nuevos cultivares de amplia adaptación ambiental, específicamente tolerantes a bajas temperaturas. La utilidad de los microsatélites para evaluar la variación genética dentro de un germoplasma de estudio ha sido bastante probada. Un alto número de microsatélites han sido localizados en regiones transcritas del genoma, lo que sugiere un posible rol funcional en la expresión y/o regulación génica. De este modo, las diferentes variantes alélicas podrían determinar cambios fenotípicos en un organismo. Este trabajo de tesis propone identificar marcadores de tipo microsatélites asociados a genes relacionados con la respuesta a frío en arroz en etapas tempranas del desarrollo, así como evaluar la relación entre la variación alélica de los microsatélites en dichos genes de respuesta a frío, y su nivel de expresión en situaciones de estrés térmico. A partir de revisiones bibliográficas y búsquedas en bases de datos de ESTs y de QTLs, se seleccionó un conjunto de 38 genes candidatos de resistencia a frío para los cuales se determinó la presencia de microsatélites en su secuencia. A partir de una colección de germoplasma consistente en 131 variedades de arroz del programa de mejoramiento genético de arroz del INIA pertenecientes a ambas subespecies *indica* y *japonica*, se analizó su variabilidad alélica con el conjunto de 38 marcadores de tipo microsatélites funcionales previamente seleccionados. De los 38 marcadores probados 16 resultaron polimórficos, de los cuales 6 mostraron una marcada diferencia de distribución entre ambas subespecies. Se buscó asociación de estos marcadores con respuesta a frío, medida como fluorescencia de la clorofila (Fv/Fm), mediante análisis discriminante. Entre los 16 marcadores polimórficos, 6 fueron capaces de discriminar con un alto porcentaje de clasificación entre clases fenotípicas predefinidas. En la segunda parte, se realizó un estudio de expresión mediante PCR en tiempo real de 6 genes candidatos de resistencia a frío, los cuales contenían secuencias de microsatélites asociadas, con el objetivo de explorar la posible relación entre las variantes alélicas con diferencias en su expresión. Para 4 de los 6 genes analizados se encontró una congruencia entre el alelo identificado y su nivel de expresión en distintos genotipos. Los resultados obtenidos en este trabajo permitirán direccionar futuros abordajes donde mediante análisis de polimorfismos de un mayor número de marcadores y la aplicación de técnicas de asociación más robustas, como es el mapeo asociativo, en un mayor número de genotipos, puedan identificarse un mayor número de marcadores asociados a frío en arroz. Por otro lado, el hecho de que 5 de los 6 genes analizados mostraran patrones similares de expresión entre genotipos tolerantes a frío, los convierte en un conjunto de genes candidatos interesante para futuros estudios de asociación.

Tabla de contenido

AGRADECIMIENTOS	ii
RESUMEN	iii
Tabla de contenido	iv
Introducción General	1
Identificación del problema.....	1
Regulación genética de la respuesta a frío en arroz	2
Microsatélites génicos.....	4
Abundancia y distribución de los microsatélites génicos.....	4
Dinámica evolutiva de los microsatélites génicos.....	5
Objetivos.....	6
CAPÍTULO 1.....	8
Identificación de marcadores génicos asociados a tolerancia a frío en plántulas de arroz usando análisis de discriminante.....	8
Introducción.....	8
Materiales y métodos.....	12
Material vegetal	12
Genotipado	12
Fenotipado.....	14
Resultados	17
Discusión.....	34
Bibliografía.....	39
CAPÍTULO 2.....	42
Análisis de expresión.....	42
Introducción.....	42
Cuantificación de la expresión génica	43
Genes candidatos asociados a la respuesta a frío en arroz	45
Objetivos.....	46

Materiales y métodos.....	47
Resultados y discusión.....	52
I. Validación de genes de referencia.....	52
II. Cuantificación de la expresión génica.....	58
Bibliografía.....	71
Conclusiones finales.....	75
Bibliografía General	77

Introducción General

Identificación del problema

Los factores abióticos como la sequía, la alta salinidad, y las bajas temperaturas tienen efectos perjudiciales sobre el crecimiento y desarrollo de las plantas, y por tanto limitan la productividad en especies cultivadas. La temperatura es un factor determinante en la distribución geográfica de la gran mayoría de las especies cultivables. Las plantas difieren en su tolerancia a las temperaturas de enfriamiento (0-15 °C) o de congelamiento (< 0 °C). Aquellas que son de origen templado son tolerantes al enfriamiento, pudiendo aumentar dicha tolerancia a través de un proceso conocido como aclimatación al frío. Por otro lado las plantas de origen tropical o subtropical, como el arroz, no son tolerantes a las bajas temperaturas y por tanto son incapaces en gran medida de aclimatarse al frío (Zhu et al., 2007). Por este motivo, el cultivo de arroz en regiones de clima templado, está fuertemente condicionado por las bajas temperaturas. Se estima que los daños causados por el frío afectan a siete millones de hectáreas de arroz por año en todo el mundo (Kariya, 2003), y casi un millón de hectáreas son potencialmente afectadas en América Latina (Cruz et al., 2001).

Existen dos subespecies de arroz cultivado, *indica* y *japonica*, las cuales difieren principalmente en cuanto a su tolerancia a las bajas temperaturas. La subespecie *indica*, cuyo centro de origen constituye el sudeste de Asia, está mayormente adaptada a clima tropical. Por otro lado, la subespecie *japonica* está principalmente adaptada a clima moderado, siendo su centro de origen el sur de China (Londo, 2006).

En el cultivo de arroz, las temperaturas por debajo de los 15 °C son potencialmente perjudiciales dependiendo del estadio de desarrollo de la planta, siendo la fase de germinación y desarrollo de las plántulas, y la fase reproductiva las dos etapas más críticas (Yoshida, 1981). Durante la fase de germinación, los principales daños observados son el retraso y la disminución del porcentaje de germinación. Durante la etapa de plántula los daños incluyen retraso en el crecimiento, baja estatura y decoloración en las hojas, mientras que durante el período reproductivo los principales síntomas del frío son esterilidad, excersión incompleta de las panojas y floración retrasada (Yoshida, 1981). Durante este período existe también un alto requerimiento de luz solar para que se desarrollen adecuadamente los órganos reproductivos (Stansel, 1975).

En Uruguay, la fase reproductiva coincide con el mes de febrero, durante el cual es frecuente la ocurrencia de temperaturas inferiores a 15 °C, y la disponibilidad de luz se reduce significativamente en comparación al mes de enero. Desde el punto de vista agronómico sería conveniente adelantar la siembra al mes de setiembre, con el objetivo de ubicar la floración en enero cuando la temperatura y luminosidad son más apropiadas. Sin embargo, durante este mes las temperaturas registradas son bajas para que se dé una correcta germinación y desarrollo de la plántula, impidiendo un correcto establecimiento del cultivo (Deambrosi et al., 1997). Esto último justifica la importancia de la generación de cultivares capaces de desarrollar tolerancia a las bajas temperaturas desde las primeras fases de su desarrollo, y poder así obtener un mejor provecho de los recursos climáticos durante el mes de enero.

De este modo, el conocimiento sobre los mecanismos bioquímicos y moleculares de respuesta a las bajas temperaturas es fundamental para la identificación de los genes vinculados a la tolerancia al frío en las etapas iniciales del desarrollo del cultivo de arroz. Estudios previos han identificado diversos grupos de genes involucrados en la respuesta temprana al frío (Cheng et al., 2007; Yun et al., 2010), para muchos de los cuales es posible determinar la presencia de microsatélites mediante el uso de las bases de datos de genoma de arroz disponibles (Gramene, Rice Annotation Project). La presencia de microsatélites polimórficos en genes transcritos sugiere su posible rol funcional y/o regulatorio, por lo tanto la búsqueda de asociaciones entre dichos polimorfismos y la variación a nivel fenotípico sería de gran interés, así como su posible relación con los patrones de expresión que dichos genes presenten en genotipos contrastantes.

En esta tesis se propone identificar marcadores funcionales (SSR génicos) y analizar la expresión de genes involucrados en la respuesta al frío durante las etapas tempranas del desarrollo (plántula), en genotipos de arroz contrastantes respecto a su grado de tolerancia al frío. Dado que estos genes contienen microsatélites en su secuencia, se realizarán estudios sobre su variación alélica y se analizará la correspondencia con su nivel de expresión mediante ensayos de PCR cuantitativo.

Regulación genética de la respuesta a frío en arroz

Las bajas temperaturas inducen cambios en la homeostasis celular activando cascadas de señalización que conducen a la expresión de genes de respuesta a frío. Tanto en plantas sensibles como resistentes al frío, la membrana celular constituye el primer sitio de daño. Principalmente en plantas sensibles, el frío induce una alteración en la fluidez de la membrana plasmática (Murata and Los, 1997) lo cual provoca un aumento de la permeabilidad de la misma. Este aumento de permeabilidad puede

afectar los canales de Ca^{2+} provocando la liberación de Ca^{2+} al medio intracelular (Plieth et al., 1999). Estos cambios en los niveles de Ca^{2+} citosólico son detectados por proteínas censoras de Ca^{2+} (quinasas y fosfatasa), las cuales desencadenan las cascadas de fosforilación, cuyo objetivo final será la activación de las proteínas involucradas en la protección celular y de los factores de transcripción que controlan la expresión de los genes de respuesta a frío específicos (Xiong et al., 2002). Los factores de transcripción son los principales reguladores, capaces de controlar la transcripción de grandes grupos de genes (regulones), a través de su unión específica a los elementos *cis* ubicados en los promotores de sus genes blanco (Nakashima et al., 2009).

La mayoría de los regulones de respuesta a frío fueron identificados por primera vez en *Arabidopsis*. El análisis de los promotores de varios genes COR en *Arabidopsis* indica que estos genes están regulados tanto por vías dependientes como independientes del ácido abscísico (ABA) bajo condiciones de estrés salino, hídrico y bajas temperaturas. Sin embargo su expresión inducida por frío es en gran medida independiente del ABA (Shinozaki and Yamaguchi-Shinozaki, 2000). Una de las vías independientes de ABA más estudiadas en *Arabidopsis* es un conjunto de factores de transcripción conocidos como el regulón controlado por la proteína DREB1 (*Dehydration-responsive element binding protein 1*)/CBF (*C-repeat binding factor*) (Cook, 2004; Fowler, 2002; Yamaguchi-Shinozaki and Shinozaki, 1992).

En aquellas plantas que no son capaces de aclimatarse al frío se han encontrado genes de respuesta a frío que contienen sitios de unión a DREB/CBF en sus promotores, indicando que muchos mecanismos de respuesta comunes entre plantas de origen tropical y templado (Rabbani et al., 2003; Yazaki, 2004; Zhang et al., 2004).

En arroz, se han encontrado cuatro homólogos a DREB1/CBF: OsDREB1A, OsDREB1B, OsDREB1C, y OsDREB1D; los cuales presentaron una rápida respuesta al estrés por frío (Dubouzet et al., 2003). De todos modos, la organización jerárquica de las vías que componen este regulón, y sus componentes aún no han sido del todo definidos (Dubouzet et al., 2003).

En un análisis transcriptómico en plántulas de arroz *japonica*, Cheng, et al. (2007) identificaron 121 genes de respuesta temprana a frío. Entre estos genes encontraron un factor de transcripción, ROS-bZIP 1, el cual era además inducido por la aplicación de H_2O_2 exógeno. Los autores concluyeron que este factor de transcripción controlaría la expresión de un nuevo regulón de respuesta a frío independiente de ABA y de CBF/DREB, el cual respondería al aumento de los niveles de ROS durante la respuesta temprana a frío.

Recientemente, en un estudio transcriptómico que involucró análisis de *microarrays* y arquitectura de promotores, Yun, et al. (2010) lograron revelar los componentes y la organización de la red de regulación a nivel transcripcional de la

respuesta temprana al frío en arroz del ecotipo *japonica*. Los autores encontraron que la activación temprana de esta red sería desencadenada por los ROS, principalmente el H₂O₂, los cuales constituyen las primeras señales fisiológicas de respuesta a frío (Yun et al., 2010). Los genes objetivo de estos mecanismos tempranos constituyen varias clases de factores de transcripción del tipo bzip, ERF y MYB. Las otras respuestas independientes de ROS, como los factores de transcripción tipo DREB/CBF, RAP2 y RAV1, y aquellas vías dependientes de ABA fueron activadas más tardíamente (Yun et al., 2010).

Microsatélites génicos

Los microsatélites o SSR (del inglés, *Simple Sequence Repeats*) son secuencias cortas de ADN repetido en tándem. Constituyen uno de los tipos más variables de secuencias en el genoma de los organismos vivos, y su polimorfismo radica en el número de repetidos en vez de su secuencia primaria (Ellegren, 2004). Debido a sus altos niveles de polimorfismo, su carácter codominante, y a su aparente evolución neutral, los microsatélites han constituido los marcadores moleculares más utilizados para el mapeo genómico y estudios de genética de poblaciones (Jarne and Lagoda, 1996; Powell et al., 1996).

Si bien los microsatélites han sido tradicionalmente considerados como “ADN basura” (carentes de información genética relevante), surgen crecientes evidencias sobre sus efectos a nivel fenotípico y molecular. Numerosos análisis sobre la composición y variación del número de repetidos de los SSR en diferentes genomas permiten apoyar la hipótesis de que los microsatélites serían una fuente importante de variación adaptativa, jugando un papel significativo en la evolución de los organismos (Kashi and King, 2006; Li, 2004).

Abundancia y distribución de los microsatélites génicos en vegetales

Las herramientas de análisis genómico y la secuenciación de ESTs (del inglés *Expressed Sequence Tags*) a gran escala, han permitido evaluar la abundancia y distribución de los microsatélites en regiones transcritas del genoma de varias especies. Si bien estas características han sido bien estudiadas en bacterias y en animales (Gur-Arie et al., 2000; Katti et al., 2001; Pupko and Graur, 1999; Toth, 2000), la información disponible en vegetales es limitada. De todos modos existen trabajos donde mediante abordajes bioinformáticos logran determinar la frecuencia y

composición de los microsatélites génicos en diferentes genomas vegetales (Lagercrantz et al., 1993; Lawson and Zhang, 2006b).

En un análisis comparativo de secuencias genómicas y de ESTs en *Arabidopsis*, arroz, soja, maíz y trigo, Morgante, et al. (2002) encontraron que en estas especies las frecuencias de los microsatélites eran más elevadas en las regiones transcritas, especialmente las porciones no codificantes, que en el resto del genoma. En un análisis *in silico* de microsatélites en *Arabidopsis* y arroz, Lawson & Zhang (2006b) encontraron que los microsatélites presentaban distintos patrones de distribución en diferentes regiones génicas (intrones, exones, 5' y 3'UTR) entre y dentro de los genomas de ambas especies. La presencia de microsatélites en regiones codificantes también presenta preferencias en cuanto a su composición nucleotídica. Tanto en *Arabidopsis* como en arroz, se encontró que el repetido AG/CT era el más frecuente en ESTs (Lawson and Zhang, 2006b), así como en el resto de las monocotiledóneas y dicotiledóneas estudiadas (Morgante et al., 2002). En cuanto a la composición de trinucleótidos ésta fue más variable entre especies siendo más frecuente el repetido CCG/CGG en arroz y mayoritariamente en regiones codificantes que en el resto del genoma (Morgante et al., 2002).

Dinámica evolutiva de los microsatélites génicos

Al igual que los microsatélites presentes en regiones no transcritas, los microsatélites génicos presentan mayores tasas de mutación que otras secuencias no repetidas del genoma. Se han presentado varios modelos que explican la dinámica evolutiva de los microsatélites. El modelo más aceptado en la actualidad propone que las distribuciones en equilibrio de las longitudes de los microsatélites son el resultado de un balance entre eventos de deslizamiento durante la replicación y mutaciones puntuales (Kruglyak et al., 1998). Mientras que el deslizamiento favorece el crecimiento, las mutaciones puntuales romperían los arreglos largos en dos o más unidades de menor tamaño. Esta dinámica evolutiva sería común a todo tipo de microsatélites ya sean asociados o no a regiones transcritas, pero los microsatélites génicos estarían sometidos a una mayor presión selectiva que el resto debido a su importancia funcional en la regulación de la expresión y función génica.

Dado que los microsatélites génicos pueden afectar la función de un gen no puede negarse su rol evolutivo. Varios trabajos proponen que en virtud de sus cualidades mutacionales y funcionales, los microsatélites génicos son capaces de proveer una fuente continua de variación cuantitativa capaz de facilitar una rápida evolución adaptativa (Fondon and Garner, 2004; Kashi and King, 2006). Se ha propuesto que la capacidad de algunos organismos de incorporar un mayor número de repetidos, podría

otorgarles una adaptación más rápida a los diversos factores de estrés generados por el ambiente (Marcotte et al., 1998; Trifonov, 2003; Wren et al., 2000). Trifonov (2003) sugiere que una de las posibles respuestas provocadas por los cambios ambientales constituye la variación en el número de copias de los repetidos en tándem, suposición que ha sido apoyada por estudios realizados en procariotas (Jackson et al., 1998) y vegetales (Schmidt and Mitter, 2004). Esta variación de naturaleza aleatoria ocurriría en la mayoría de los microsatélites del genoma del organismo sometido a algún tipo de estrés ambiental, siendo seleccionados aquellos cambios capaces de modificar la actividad de un gen -o grupo de genes- relacionada con la tolerancia al estrés, otorgándole una ventaja adaptativa (Trifonov, 2003).

Las evidencias sobre la dependencia entre la expresión y otras funciones génicas, y el número de copias de sus microsatélites asociados son cada vez más frecuentes (Li, 2004). Como fue mencionado anteriormente, las diferentes condiciones de estrés a las que puede estar sometido un organismo, incluidas las bajas temperaturas, serían capaces de inducir variaciones en el número de repetidos de los microsatélites génicos. Estas variaciones podrían afectar o bien la función, o la regulación de la expresión de dichos genes. Esto sería especialmente importante para el caso de los genes de respuesta a frío, ya que la presencia de ciertas variantes alélicas podría otorgar ventajas adaptativas (tolerancia) a aquellos organismos portadores. Por lo tanto, el análisis de la relación entre ciertas variantes alélicas con la expresión de los genes de resistencia a frío sería de gran importancia para lograr un mejor conocimiento sobre los mecanismos de tolerancia a estos tipos de estrés.

Objetivos

Este trabajo de tesis propone en primer lugar identificar marcadores SSR asociados a genes relacionados con la respuesta a frío en arroz en etapas tempranas del desarrollo, así como evaluar la relación entre la variación alélica de los microsatélites en dichos genes de respuesta a frío, y su nivel de expresión en situaciones de estrés térmico.

El cuerpo general de este trabajo se divide en dos capítulos. En el primer capítulo se realizará un estudio sobre la presencia de microsatélites (a través de las bases de datos de secuencias de arroz), en genes de respuesta a frío en arroz previamente seleccionados mediante búsquedas bibliográficas. Con estos marcadores identificados se realizará el genotipado de 131 genotipos de arroz pertenecientes al programa de mejoramiento genético de arroz del INIA (Instituto Nacional de Investigación Agropecuaria). Con estos resultados se estudiará la distribución de la diversidad

genética, y la estructura de las poblaciones de arroz analizadas. En paralelo se harán mediciones fenotípicas de estos mismos genotipos en respuesta a frío, las cuales servirán para la búsqueda de asociaciones entre los marcadores identificados y las características fenotípicas.

En el segundo capítulo, se realizará un estudio de expresión mediante PCR en tiempo real de un subconjunto de genes seleccionados a partir de los resultados obtenidos en el primer capítulo, con el objetivo de analizar la posible relación entre las variantes alélicas con diferencias en su expresión. Estos estudios se realizarán sobre tres genotipos de arroz uruguayo, los cuales difieren en cuanto a su capacidad de tolerancia a las bajas temperaturas. Para cumplir con este propósito, se realizará previamente un estudio de validación de genes de referencia para estrés por frío para las variedades seleccionadas.

CAPÍTULO 1

Identificación de marcadores génicos asociados a tolerancia a frío en plántulas de arroz usando análisis de discriminante

Introducción

La temperatura es un factor determinante en la distribución geográfica de la gran mayoría de las especies cultivables. Las plantas de origen tropical o subtropical, como el arroz, son particularmente susceptibles a las bajas temperaturas. En Uruguay, una de las principales limitantes en la producción de arroz es la ocurrencia de bajas temperaturas tanto durante las etapas tempranas del desarrollo de la planta, como durante la fase reproductiva.

Existen dos subespecies de arroz cultivado, *indica* y *japonica*, las cuales difieren principalmente en cuanto a su tolerancia a las bajas temperaturas. La subespecie *indica* es de origen tropical y es fundamentalmente susceptible al frío en comparación a la subespecie *japonica*, de origen templado. En un estudio sobre la estructura de la diversidad genética dentro de *Oryza sativa* mediante marcadores moleculares nucleares y cloroplásticos, Garris et al. (2005) lograron distinguir 5 subpoblaciones dentro de *O. sativa*: *indica*, *aus*, *tropical japonica*, *temperate japonica* y *aromatic*. Las subpoblaciones *indica* y *aus* conformarían el grupo *indica*, mientras que *tropical japonica*, *temperate japonica* y *aromatic* constituyen el grupo *japonica*, existiendo un alto grado de diferenciación genética entre ambos grupos (Garris et al., 2005).

El modelo actual sobre el proceso de domesticación de *O. sativa* establece que *indica* y *japonica* fueron domesticados a partir de poblaciones divergentes de la especie ancestral, *O. rufipogon*, lo cual sumado al gradual pasaje desde un hábito reproductivo alógamo a uno autógamo durante el proceso de domesticación, explican el alto grado de diferenciación genética entre ambas (Caicedo et al., 2007; Garris et al., 2005). Sin embargo, en las etapas tempranas de este proceso un cierto grado de flujo

génico, aunque restringido, permitió la introgresión y fijación de un conjunto de alelos de domesticación comunes entre ambas subespecies (Kovach et al., 2007; Sang and Ge, 2007).

Los principales efectos del daño causado por el frío durante las fases tempranas de desarrollo del cultivo incluyen retraso de la germinación y desarrollo de las plántulas, mientras que durante la etapa reproductiva los principales síntomas del frío son esterilidad, excursión incompleta de las panojas y floración (Yoshida, 1981). Esto justifica la importancia de mejorar el germoplasma disponible de manera de poder desarrollar nuevos cultivares de amplia adaptación ambiental, especialmente las bajas temperaturas.

Los estreses ocasionados por condiciones ambientales como la sequía, la alta salinidad o las bajas temperaturas inducen una serie de respuestas fisiológicas y moleculares. Muchas de estas respuestas inducidas son comunes a estos tres tipos de estrés, lo cual es una consecuencia de la activación de genes compartidos (Ishitani et al., 1997; Rabbani et al., 2003; Shinozaki and Yamaguchi-Shinozaki, 2000). Las bajas temperaturas inducen la expresión de genes denominados COR (COR=Cold Responsive), la mayoría de los cuales son regulados a nivel transcripcional. Estos genes pueden clasificarse principalmente en dos grupos (Shinozaki and Yamaguchi-Shinozaki, 2000): un grupo de genes codifica para productos que conducen directamente a la protección y reparación de los daños celulares, y otro grupo de genes cuyos productos se encargan de regular la expresión génica y de la transducción de señales.

Los mayores progresos en el entendimiento de las respuestas al estrés por frío vienen de la aplicación de las tecnologías de análisis genómico a nivel estructural y funcional. Los estudios transcriptómicos y proteómicos han aportado un mayor conocimiento global sobre los grupos de genes expresados en condiciones de bajas temperaturas. Mediante el uso de estas técnicas, varios estudios han logrado dilucidar en gran parte algunos de los genes que se expresan o inhiben bajo condiciones de estrés por frío en plántulas de arroz (Cheng et al., 2007; de los Reyes et al., 2003; Rabbani et al., 2003; Yun et al., 2010).

Otro gran avance en ésta área constituye el continuo aumento de la información genómica estructural y funcional sobre el cultivo de arroz, la cual está disponible en bases de datos de acceso público. Actualmente existen diversas bases de datos de secuencias de EST para especies de la familia Poaceae (incluyendo cultivos como arroz), que pueden ser analizadas para detectar microsatélites tanto en regiones codificantes como en regiones no traducidas de los genes (5' ó 3' UTRs). Esto permite analizar la presencia de microsatélites en genes de resistencia a frío en arroz previamente identificados y de este modo estudiar su posible función.

La utilidad de los microsatélites para evaluar la variación genética dentro de un germoplasma de estudio ha sido bastante probada. El uso de microsatélites génicos con este fin podría potenciar el rol de los marcadores genéticos analizando la variación en genes de función conocida (Andersen and Lübberstedt, 2003; Varshney et al., 2005). La variación en el número de repetidos de los microsatélites dentro de genes de función conocida puede ser testada tanto por asociación con variación a nivel fenotípico como por una función biológica determinada (Ayres et al., 1997).

El análisis discriminante es una técnica estadística multivariada desarrollada por Fisher (1936), la cual es capaz de analizar si existen diferencias entre una serie de grupos en los que se divide una población, con respecto a un conjunto de variables cuantitativas. Este tipo de análisis se ha utilizado previamente en aplicaciones agronómicas con el objetivo de combinar datos obtenidos a partir de marcadores moleculares, con el desempeño fenotípico de individuos o grupos de individuos de modo de identificar marcadores informativos (Bonamico et al., 2010; Mcharo et al., 2004; Zhang et al., 2005). En estos casos el método implica la generación de dos clases compuestas por líneas puras seleccionadas con valores fenotípicos contrastantes, y con perfiles genotípicos caracterizados mediante marcadores moleculares. De este modo, el análisis ha permitido identificar aquellos marcadores capaces de diferenciar entre clases fenotípicas con mayor precisión. Los resultados de este análisis se expresan como porcentaje de correcta clasificación, el cual es una medida de la habilidad de los marcadores seleccionados para asignar cada individuo a su clase correspondiente. Si estos valores de porcentaje de clasificación son elevados, podría inferirse una asociación entre el/los marcadores y las características fenotípicas en estudio (Mcharo et al., 2004; Zhang et al., 2005). En un trabajo sobre identificación de marcadores candidatos de tipo SSR y RFLP, asociados a características agronómicas en arroz, Zhang et al. (2005) encontraron que para los 12 caracteres fenotípicos medidos pudieron encontrarse asociaciones con un alto porcentaje de clasificación. A su vez, estos marcadores fueron detectados en regiones cercanas a QTLs para dichas características agronómicas, previamente identificados. A partir de estos resultados, los autores reconocen la utilidad del análisis discriminante para encontrar asociaciones entre marcadores y caracteres fenotípicos de tipo cuantitativos.

Para un correcto estudio de asociación es necesario contar con una buena medición de las características fenotípicas. Uno de los principales efectos fisiológicos de las bajas temperaturas en plántulas de arroz, lo constituye la fotoinhibición generada por la inhibición del transporte de electrones a través del fotosistema II (PSII), ocasionado por la acumulación de especies reactivas del oxígeno (ROS) (Jeong et al., 2002). Dicho daño se produce principalmente por la inhibición de los mecanismos de reparación del daño en el PSII (Nishiyama et al., 2006). Trabajos anteriores sostienen que la medición de la máxima eficiencia cuántica (Fv/Fm) constituye el mejor indicador de fotoinhibición y por tanto constituye un buen método

de evaluación fenotípica de la respuesta a frío (Hirotsu, 2005; Maxwell and Johnson, 2000).

En este trabajo se identificaron genes candidatos de respuesta a frío con microsatélites en su secuencia, para lo cual se realizó una revisión bibliográfica y posterior análisis de presencia de microsatélites mediante bases de datos de acceso público. Aquellos marcadores polimórficos fueron utilizados para analizar la estructura de 131 líneas puras de arroz provenientes del programa de mejoramiento de arroz de INIA (Instituto Nacional de Investigación Agropecuaria). Posteriormente se analizó la capacidad de estos marcadores candidatos de clasificar correctamente a estos individuos dentro de clases fenotípicas, predefinidas en base a respuesta a frío, mediante análisis discriminante.

Objetivo general: Identificar marcadores funcionales asociados a tolerancia a frío en arroz en etapas tempranas del desarrollo mediante análisis discriminante.

Objetivos específicos

Identificación y selección de genes candidatos involucrados en la respuesta a frío portadores de microsatélites polimórficos en su secuencia.

Genotipado con los microsatélites génicos del germoplasma elite del programa de mejoramiento de arroz de INIA (Instituto Nacional de Investigación Agropecuaria).

Búsqueda de asociación entre la respuesta a frío en germoplasma elite del programa de mejoramiento de arroz de INIA y marcadores microsatélites génicos

Hipótesis

Objetivo específico 1: Dada la elevada presencia de microsatélites génicos en vegetales reportada en trabajos anteriores, es esperable que parte sino la mayoría de los genes de respuesta a frío en arroz contengan microsatélites polimórficos en su secuencia, principalmente en las regiones transcritas no codificantes (3' y 5'UTRs e intrones).

Objetivo específico 2: Sería posible encontrar una relación entre la presencia de ciertos alelos de microsatélites en genes de respuesta a frío en determinadas variedades de arroz, con su capacidad de tolerancia a nivel fenotípico.

Materiales y métodos

Material vegetal

Se seleccionaron un total de 131 genotipos pertenecientes al programa de mejoramiento nacional de arroz de INIA (Instituto Nacional de Investigación Agropecuaria). Estos genotipos constituyen líneas puras las cuales aún siguen en proceso de mejoramiento. Entre estas 131 líneas 72 constituyen genotipos *indica*, donde 43 de estos provienen del programa de mejoramiento de arroz del FLAR (Fondo Latinoamericano para el Arroz de Riego) y el resto pertenecen al INIA; el resto de las 59 líneas constituyen genotipos *japonica* y todos pertenecen al INIA. Las semillas correspondientes a cada genotipo fueron embebidas en agua destilada estéril por 10 min, y luego esterilizadas superficialmente mediante inversiones consecutivas en etanol 70% por 3 minutos e hipoclorito de sodio al 50% 3 veces por 15 min cada vez. Las semillas fueron posteriormente enjuagadas 3 veces con agua destilada estéril, y colocadas en placas de Petri sobre gasa embebida en agua e incubadas a 30 °C por 4 días para favorecer la germinación.

Genotipado

Extracciones de ADN

El ADN de las muestras fue extraído utilizando el protocolo de FAO/IAEA (2002), con pequeñas modificaciones, a partir de 0,1 g hoja fresca de cada variedad, de acuerdo al método CTAB (Bromuro de cetiltrimetilamonio) (M.G.Murray and W.F.Thompson, 1980). La muestra fresca fue macerada con mortero en nitrógeno líquido, a la cual se agregó 0,7 ml de buffer de extracción (CTAB + β -mercaptoetanol). Esta mezcla se incubó a 65 °C durante 20 min. Luego, se dejó enfriar a temperatura ambiente y se agregó a cada muestra 1 volumen de cloroformo-isoamílico (24:1), mezclándose suavemente por inversión. Posteriormente se centrifugó a 16000 x g por 20 min y se transfirió la fase acuosa (500 μ L) a un nuevo tubo. En el siguiente paso se agregó 0,7 volúmenes de isopropanol puro y frío, mezclando cuidadosamente. Se dejó reposar por 45 min a 4 °C y después se centrifugó a 3000 x g por 20 min y se descartó el sobrenadante. Se lavó el *pellet* de ADN, con 0,5 mL de etanol 70 % y centrifugó a 3000 x g por 5min. El lavado del *pellet* se repitió dos veces, posteriormente se secó en estufa a 35 °C por 30 min. Se resuspendió el *pellet* en 300 μ L de agua ultrapura, constituyendo la solución de stock reservada a -20 °C para su posterior utilización. La calidad del ADN se determinó mediante electroforesis en geles de agarosa al 1 %, en

solución de tinción con bromuro de etidio. La concentración de ADN fue cuantificada mediante NanoDrop ND8000 (Thermo Scientific). La concentración de la solución de trabajo para las reacciones de amplificación por PCR se ajustó a 100 ng/μL.

Amplificación por PCR

Aquellos marcadores encontrados en los genes seleccionados fueron utilizados para el genotipado de la población en estudio. Para la amplificación de los marcadores SSR, se diseñaron cebadores forward extendidos por el agregado de una secuencia M13 (5'-TGTTAAACGACGGCCAGT-3') en su extremo 5', y los cebadores M13 utilizados fueron marcados con los fluoróforos 6-FAM y HEX. Las reacciones de PCR fueron llevadas a cabo en un volumen final de 25 μL. Las concentraciones de los reactivos fueron las siguientes: buffer 1X (Fermentas, Thermo Scientific), 2.5 mM de Mg²⁺, 0.12 μM de dNTPs (Fermentas), 2.2 μM de cebador M13, 2.2 μM de cebador reverse, 0,56 μM de cebador forward y 0,8 U de enzima Taq polimerasa (Fermentas, Thermo Scientific) y 100 ng de ADN. La amplificación fue llevada a cabo en un termociclador Eppendorf (Eppendorf AG, Hamburgo, Alemania). El ciclado consistió de una desnaturalización inicial a 94 °C por 5 min, seguido por un ciclo de 30 s a 94 °C, 30 s a 65 °C, y 30 s a 72 °C. Luego la temperatura de alineamiento fue disminuyendo de a 1 °C por 10 ciclos, siguiendo con 30 ciclos de 30 s a 94 °C, 30 s a 55 °C y 30 s a 72 °C, seguido por 8 ciclos de 30 s a 94 °C, 30 s a 53 °C y 30 s a 72 °C. El ciclado finalizó con una extensión final de 5 min a 72 °C.

Detección de fragmentos

Los productos de PCR fueron sometidos a electroforesis en gel de agarosa al 2% para confirmar amplificación. Los productos amplificados fueron enviados en placas de 96 en formamida al servicio de secuenciación de la Universidad de Yale (DNA Analysis Facility). Estos resultados fueron analizados con el programa Peak Scanner versión 1.0 (Applied Biosystems).

Análisis de diversidad

Se construyeron matrices con el tamaño de todos los alelos de cada marcador para cada genotipo. Se calcularon las frecuencias alélicas para cada marcador analizado mediante el paquete para Excel Genalex 6.4 (www.anu.edu.au/BoZo/GenAlEx/). Se realizaron análisis moleculares de la varianza (AMOVA) (Excoffier et al., 1992), para describir la partición de la diversidad genética entre y dentro de poblaciones, utilizando el mismo programa. En primer lugar se realizó un análisis tomando en

cuenta la estructura entre ecotipos (*indica/japonica*), mientras que para el segundo análisis se realizó considerando los tres grupos delimitados por programa de mejoramiento (*indica*-INIA, *indica*-FLAR y *japonica*).

Estructura de las poblaciones

Para analizar la estructura de las poblaciones se utilizó el programa Structure versión 2.3.3 (Pritchard et al., 2000). Este programa permite asignar individuos en K grupos o poblaciones, en base a su genotipo para múltiples loci mediante un abordaje Bayesiano (Pritchard et al., 2000). Este programa calcula el logaritmo de la probabilidad de los datos (LnPD) dado por la ecuación $\Pr(X/K)$ para cada valor de K, permitiendo la estimación del número de poblaciones más probable.

En primer lugar se creó un nuevo conjunto de parámetros, para lo cual se eligió el modelo de admixia (sin información *a priori* sobre poblaciones) y se mantuvo el resto de los parámetros por defecto. Todas las simulaciones fueron llevadas a cabo con 100,000 corridas *burn-in*, seguidas por 500,000 generaciones MCMC. El resto de los parámetros fueron establecidos por defecto.

Para establecer el valor de K más apropiado se analizaron varios números de subpoblaciones (2-9) con 10 iteraciones cada una. Se calculó el promedio de LnPD para cada valor de K con su correspondiente desvío y se graficaron. Con estos resultados se realizaron las simulaciones con un nuevo conjunto de parámetros, incorporando información sobre poblaciones mediante la opción USEPOPINFO.

Fenotipado

Tratamientos de frío

Una vez germinadas, las plántulas fueron luego crecidas en un sistema de hidroponia en medio líquido Murashige & Skoog (Murashige and Skoog, 1962), en cámara de crecimiento con un régimen de alternación de períodos de luz de 16 horas y períodos de oscuridad de 8 horas, y con $300 \mu\text{E m}^{-2} \text{s}^{-1}$ de irradiación en la superficie foliar, con una temperatura de 28 °C. Luego de los 13 días, de crecimiento, las plántulas alcanzan el estadio de crecimiento V3 (Counce et al., 2000), y fueron luego sometidas a 10 °C para el tratamiento de frío, por 48 horas. Las plantas control fueron mantenidas bajo las mismas condiciones a 28 °C.

Medidas de FV/FM

Las mediciones de fluorescencia de la clorofila fueron realizadas por triplicado usando plantas control y plantas tratadas con frío, con un fluorímetro FMSI Hansatech (King's Lynn Instrument Ltd., King's Lynn, UK). Las hojas fueron sometidas a oscuridad por 30 minutos antes de realizar el análisis. Se tomaron los datos durante la exposición por 0,8 segundos a 8000 μmol fotones $\text{m}^{-2}\text{s}^{-2}$ de luz actínica. La eficiencia cuántica máxima del PSII (F_V/F_M), la cantidad cuántica intrínseca medida durante las iluminaciones (F_V/F_M), y el *quenching* fotoquímico (qP) fueron calculados automáticamente por el fluorímetro. La eficiencia del PSII (P) fue calculada como el producto de F_V/F_M y qP (Verhoeven et al., 1997).

Análisis discriminante

Para buscar asociación entre marcadores y fenotipo se utilizaron análisis discriminantes. Con los datos obtenidos a partir del genotipado se elaboró una matriz binaria, para presencia y ausencia de cada alelo por marcador. Este análisis se realizó a través del algoritmo "vecino más cercano" (*k-nearest neighbor-IBK*), el que pertenece al tipo de análisis discriminante no paramétrico implementado en WEKA versión 3.5.8 (Waikato Environment for Knowledge Analysis, Univ. of Waikato, NZ, <http://www.cs.waikato.ac.nz/~ml/Weka/>)

El algoritmo *k-nearest neighbor* clasifica cada instancia del conjunto (sobre las que se quiere tener una predicción), en base al conjunto de referencia definido (aquellas instancias para las que se conoce su clase) de acuerdo a los siguientes criterios: 1) encuentra los k vecinos más cercanos dentro de las instancias del conjunto de referencia, y 2) predice la clase según la del vecino más cercano (si $k > 1$, utiliza la clase más frecuente), por ejemplo, elige la clase que es más común entre los k vecinos más cercanos (Witten and Frank, 2005).

Determinar cuál es el número de vecinos más cercanos óptimo, tiene directa relación con el tamaño del conjunto de entrenamiento. Si el conjunto de entrenamiento es grande, es intuitiva la elección de más de un vecino ($k > 1$), pero es peligroso cuando el número de instancias es pequeño. Es posible demostrar que cuando k y el número n de instancias tienden a infinito de tal forma que $k/n \rightarrow$ cero, la probabilidad del error se aproxima al mínimo teórico para el conjunto de datos (Witten and Frank, 2005).

Los resultados del análisis se expresan como porcentaje de clasificación correcta para cada variedad de acuerdo a su correspondiente clase fenotípica. Para medir la

capacidad de acierto de la clasificación se utilizó el coeficiente estadístico Kappa (Cohen, 1960).

$$\text{Kappa} = \frac{\text{Pr}(a) - \text{Pr}(e)}{1 - \text{Pr}(e)}$$

Donde $\text{Pr}(a)$ es la proporción de aciertos observados y $\text{Pr}(e)$ es la probabilidad esperada de aciertos por azar. Este coeficiente permite estimar si la clasificación obtenida puede ser producto del azar ($k \leq 0$) así como el grado de robustez del clasificador utilizado (mayor cuanto más cercano a 1).

También se procedió a seleccionar entre las instancias evaluadas aquellas con mayor poder discriminante, para lo cual se utilizó el método *best first* implementado en WEKA. Este método de búsqueda de variables recorre todo el espacio de atributos (en este caso alelos de cada marcador analizado), asigna un valor de desempeño a cada una de las variables y devuelve una lista ordenada según la medida del desempeño (poder discriminante), la cual está expresada en porcentajes.

Resultados

Selección de genes candidatos y microsatélites génicos

Se realizó una revisión bibliográfica sobre trabajos donde se aislaron genes de resistencia a frío en etapa de plántula, de la cual se eligió un total de 218 genes (Tabla 1). Para dicha clasificación funcional se siguió el criterio de clasificación propuesto por Yun et al. (2010).

Tabla 1: Genes de respuesta a frío en arroz, expresados en etapas tempranas del desarrollo e identificados mediante revisiones bibliográficas. Se muestran los genes agrupados por función, y el porcentaje de representación de cada grupo.

Función	Total	Porcentaje
Señalización y regulación de la respuesta	5	2,29%
Defosforilasas	2	0,91%
Quinasas	3	1,38%
Defensa celular	72	33%
Genes asociados a estrés oxidativo	49	22,50%
Reparación de componentes celulares dañados	3	1,38%
Respuesta a estrés (DHN,LEA)	20	9,12%
Ajuste y sustento fisiológico	18	8%
Biogénesis celular y crecimiento	4	1,58%
Traducción	3	1,38%
Metabolismo energético	11	5,04%
Factores de transcripción	123	56%
Tipo WRKY	16	7,30%
Tipo bhlh	29	13,00%
Tipo NAC	21	9,60%
Tipo bzip	11	5,04%
Tipo AP2/ERF	29	13,00%
Tipo MYB	17	8,06%

La mayor proporción de genes de respuesta a frío encontrados en la bibliografía constituyen factores de transcripción (56%). Entre el resto de los genes, aquellos que codifican para proteínas relacionadas con la defensa celular frente a situaciones de estrés tienen mayor representación (33%), dentro de los cuales aquellos vinculados a procesos regulatorios del metabolismo oxidativo (peroxidasas, catalasas, componentes del sistema glutatión, tiorredoxinas, etc.), son los más abundantes (22,5%).

Estos 218 genes fueron analizados en cuanto a la presencia de microsatélites en su secuencia a través de las bases de datos de acceso público (<http://www.gramene.org/>, <http://rice.plantbiology.msu.edu/>) (Tabla 2). La mayoría de los genes que contenían SSRs en su secuencia resultaron nuevamente ser factores de transcripción (73,68%). Por otro lado, los tri-nucleótidos resultaron ser el tipo de repetido más prevalente, representando un 64,7% del total. Los di-nucleótidos fueron el segundo tipo de repetido más frecuente (27,77%), mientras que sólo un tetra-nucleótido fue encontrado (RM22034). Todos los repetidos localizados en regiones traducidas (exones) resultaron ser tri-nucleótidos, y con un alto contenido GC (tabla 3) salvo uno, el RM5353, el cual consiste en un di-nucleótido.

Tabla 2: Genes candidatos con microsatélites génicos. Se muestra el nombre del marcador asociado (Gramene), el motivo, su localización, así como los cebadores utilizados para amplificarlos.

Gen	Accesión	Marcador SSR asociado	Motivo	Localización	Cebador Forward	Cebador Reverse
Genes adaptativos						
Low-temperature induced 6a	LOC_Os07g44180	RM22034 ^P	(ATTA)5	3'UTR	CCAGTTTATCTTCTGCACCTTCTCG	TCTTTGAGCAGATGGCTAACAGG
Phosphatase 2C	LOC_Os01g40094	RM11239 ^P	(TG)15	intrón	GGAGTTAGTGCAAGCAATTCAGG	TGGTCAGGTAGACCACAGATCC
Gluthation peroxidase	LOC_Os03g24380	RM14978 ^P	(TC)10	intrón	TATCTGCAGGTGCGTGAAATGG	GCATATAGAGCGAGTAAGCGAGAGG
Auxin induced protein	LOC_Os01g13030	RM6651 ^P	(GTG)8	intrón	AAGCATCAAGTAGGCAGCACAGC	GGAAGGAGGGAGATGAGATGAGC
Ferritina	LOC_Os11g01530	RM1761 ^P	(AT)16	intrón	ATCTTACACTGTGCAGGTTGTGC	GGAACTTCCTGGTGAATGAGC
Early-response to dehydration-like	LOC_Os03g23010	RM14960 ^P	(CGA)7	exón	GTCAAGACCACCGCTGGTTCC	CGACGAGGAGGTTATGGGTCTGG
Alternative oxidase	LOC_Os04g51150	RM17440 ^M	(CGC)8	exón	CTTCTCCGCTGCAAGGAACG	TTTGCAAATCGCAACCATCTCG
Heat-shock protein 70	LOC_Os11g47760	RM144 ^P	(ATT)11	intrón	CATGTTGTGCTTGTCTACTGC	AGCTAGAGGAGATCAGATGGTAGTGC
Low temperature viability protein	LOC_Os03g25620	RM15041 ^M	(GCC)7	5'UTR	CTGTTTACGAGCTCCTCTCCAAATCC	GACGGTGTAGTCGTTGCTGTCC
AMP binding protein	LOC_Os02g32490	RM13335 ^P	(GT)10	intrón	TATGCCAAGAGGAATCCTGAAGC	GCACTCACACTGATCTGGACAGG
Factores de transcripción						
OsWRKY1v2	LOC_Os01g14440	RM6516 ^M	(CCG)8	exón	AAGCATCAAGTAGGCAGCACAGC	GGAAGGAGGGAGATGAGATGAGC
OsWRKY32	LOC_Os02g53100	RM8016 *	(AT)22	intrón	CAAGCATTATCAGTTGGCTTCC	GTGCTAGACGACGTTCTCAAACC
OsWRKY5	LOC_Os05g04640	RM17863 ^M	(TA)11	intrón	GGACGTCATGATTTGTGCTTGG	GTCATTGGCTTGCAAGATACACC
OsWRKY8v2	LOC_Os05g50610	RM19165 ^M	(CTG)7	exón	GTGTAGAGATGAAGAGGAGGCATCG	CGTGAGCTACCACATGCACAGG
bhlh37	LOC_Os01g11910	RM10381*	(CCG)10	intrón	TGGAGATGGCAAAGAGGAATTGTGG	GAGAGTTCGCCGTCGCTTGG
bhlh29	LOC_Os02g12820	RM12729 ^M	(AT)33	exón	TGGAGCAGCTCAAGATAGTAGACC	TATTTGGATTCCGTTCCAGACG
bhlh148	LOC_Os03g53020	RM15892 ^P	(AAG)7	5'UTR	TATCCCGGTTTCAATATTTCC	ACGACTCCATTTGCATCTCTTCC

Gen	Accesión	Marcador SSR asociado	Motivo	Localización	Cebador Forward	Cebador Reverse
bhlh41	LOC_Os03g59670	RM16111 *	(CCG)8	exón	GACGGAGATGAACGTCGAGACC	CAGTTACTGTCAACAAGGCAGGTAGG
bhlh18	LOC_Os03g51580	RM15841 *	(GCG)7	exón	CAGTACACGGAGCAGCAGATCG	CCCAGCTGGTGTTCACCTTCG
bhlh87	LOC_Os08g38080	RM5353 *	(TC)13	exón	GTCGCTGGATGCAGCAAAGTGG	GCCGAAGCGTTGCTCTCTACTCC
NAC	LOC_Os01g60020	RM6547 ^P	(GCT)9	5'UTR	ATGACACACCACCACCAACG	GGAAGCGCAGCATCAACAAAGG
NAC	LOC_Os01g66490	RM11987 ^M	(GGA)7	exón	AACTAGCCCTAGCGACGGATGC	AGGTCATCCTCATGCCCATCTACC
NAC	LOC_Os12g41680	RM28690 ^M	(CT)17	intrón	CTTCTCCTCCAAGGTATGTATGG	GCATATTTGCAGCTTGATAGGG
NAC	LOC_Os11g05614	RM26081 ^M	(TCA)7	3'UTR	GTGAAATGGGATCGGATTCATGG	GATCGATCAACCGGAGATGACG
ROS_bZIP1	LOC_Os08g43090	RM23550 ^P	(TA)35	3'UTR	GAAGGATGATGGTATCCATGC	AAACACGATGCCTCCTATTCAGG
bZIP	LOC_Os12g06520	RM27564 ^M	(GGA)7	exón	CTTCTTCGAGGGTTTCTCGTTGC	CTCCTCTTATGCCACCCTTCC
bZIP	LOC_Os08g38020	RM23346 ^M	(GTG)7	exón	GCAGCGCCATGTTGATGAGG	CCTCCAAACCTCCCTTCTTCG
bZIP	LOC_Os02g03960	RM12403 ^P	(GGC)8	exón	GGATCTGAAGAGGTCGTCGAAGG	ACATCCTCTCCTGCATCAACACC
bZIP	LOC_Os05g37170	RM18776 ^P	(GA)13	3'UTR	CTCCAGGAGGGTACAAATCTGC	CCATTGGAACATAGCAAGTGATCG
AP2 domain containing protein	LOC_Os05g28350	RM18462 *	(CCG)7	exón	GCATGGAATCGGTCACTTGAGG	GTCTTCCACGGCTGAAACACAGG
Ethylene-responsive transcription factor	LOC_Os06g11860	RM19630 ^M	(GCC)7	exón	CACGCCGCGGTACATCTTGC	CGAACACCACCCTCCTTCTCG
AP2 domain containing protein	LOC_Os04g44670	RM17273 ^M	(CCG)7	5'UTR	GCTCCGCCAATCTTCTTGAGC	GCGCTAGCTGCGTTTCTTCTCC
B3 DNA binding domain containing protein	LOC_Os01g04750	RM10114 ^P	(GAG)7	5'UTR	TGGTCGTCTTCTCCTGCAACTCC	GATCAACTCGTCTCCTCCATTGC
AP2 domain containing protein	LOC_Os03g08500	RM14490 ^M	(CGG)7	exón	GGAGACGACGGCTAGAGCCTAGAAGG	GCCGACCAGTACAACCACTTCAACG
MYB family transcription factor	LOC_Os04g40420	RM17139 ^P	(AT)13	intrón	TATCTGTTGGTGCAGTTGATCC	AGCTGCCCTACTTGATATGTGG
MYB family transcription factor	LOC_Os09g36730	RM215 ^P	(CT)16	3'UTR	GAGCAGCAAGAGCAGCAGAGG	CATGCTCGACTTCAGAAGCTTGG

*: No amplificó, M: Monomórfico, P: Polimórfico

Por su parte, el tipo de repetido más frecuente entre los di-nucleótidos fue AT, mientras que el tri-nucleótido más frecuente fue CCG (Tabla 3), el cual codifica para el aminoácido prolina (22,73%) (Tabla 4). Los siguientes aminoácidos más frecuentes fueron la alanina (18,18%), la arginina (13,65%), y la glicina (13,64%) (Tabla 4).

Tabla 3: Número y frecuencia de los repetidos analizados para genes ligados a la tolerancia a frío.

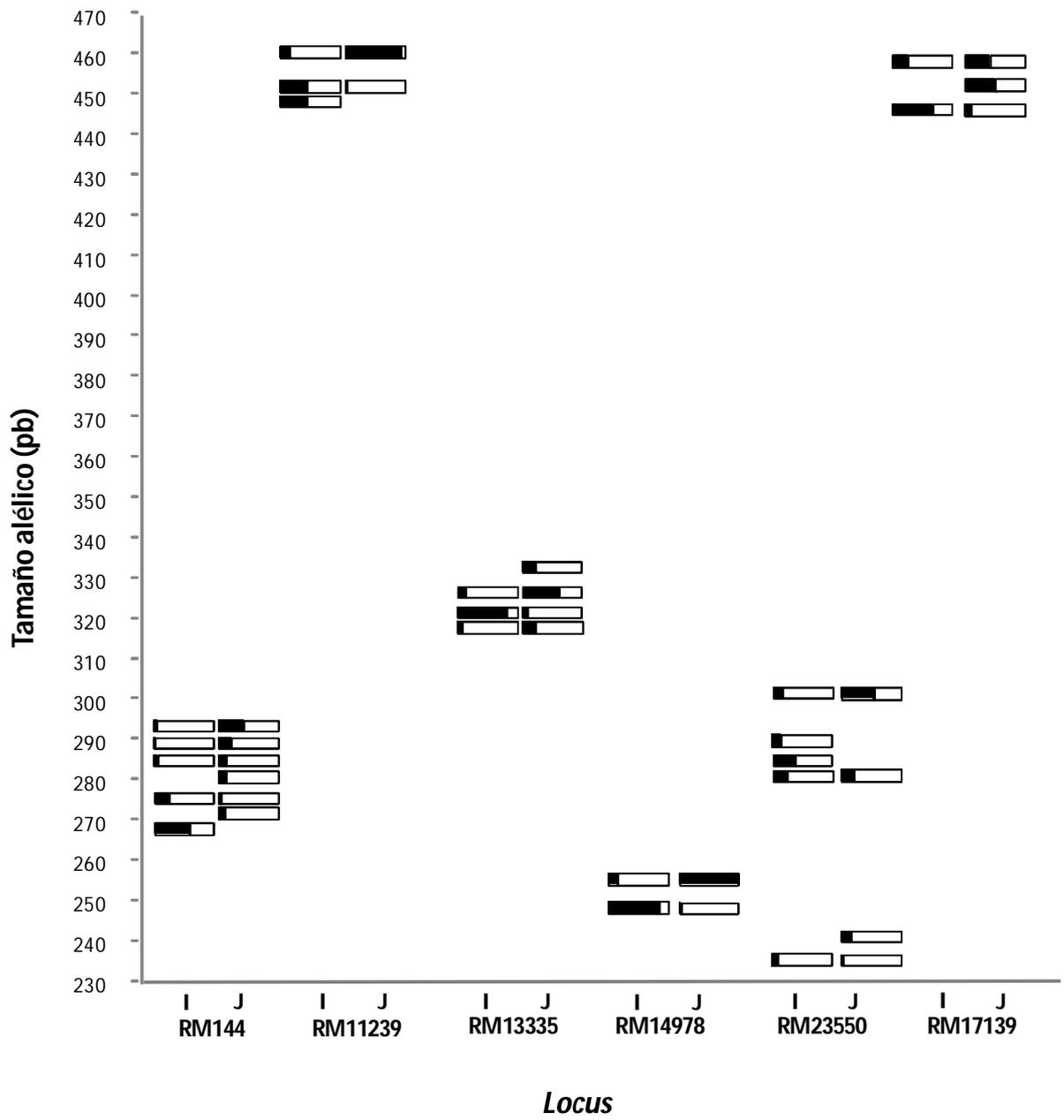
Repetido	Total	%
Di	12	34,2
AT/AT	4	11,4
TA/TA	2	5,7
CT/CT	2	5,7
TC/GA	2	5,7
Otros	2	5,7
Tri	22	62,8
CCG/CGG	6	17
GCC/GGC	3	8,7
GTG/GTG	2	5,7
GCG/CGC	2	5,7
GGA/TCC	2	5,7
Otros	7	20
Tetra	1	3
ATTA	1	3
Total	35	100

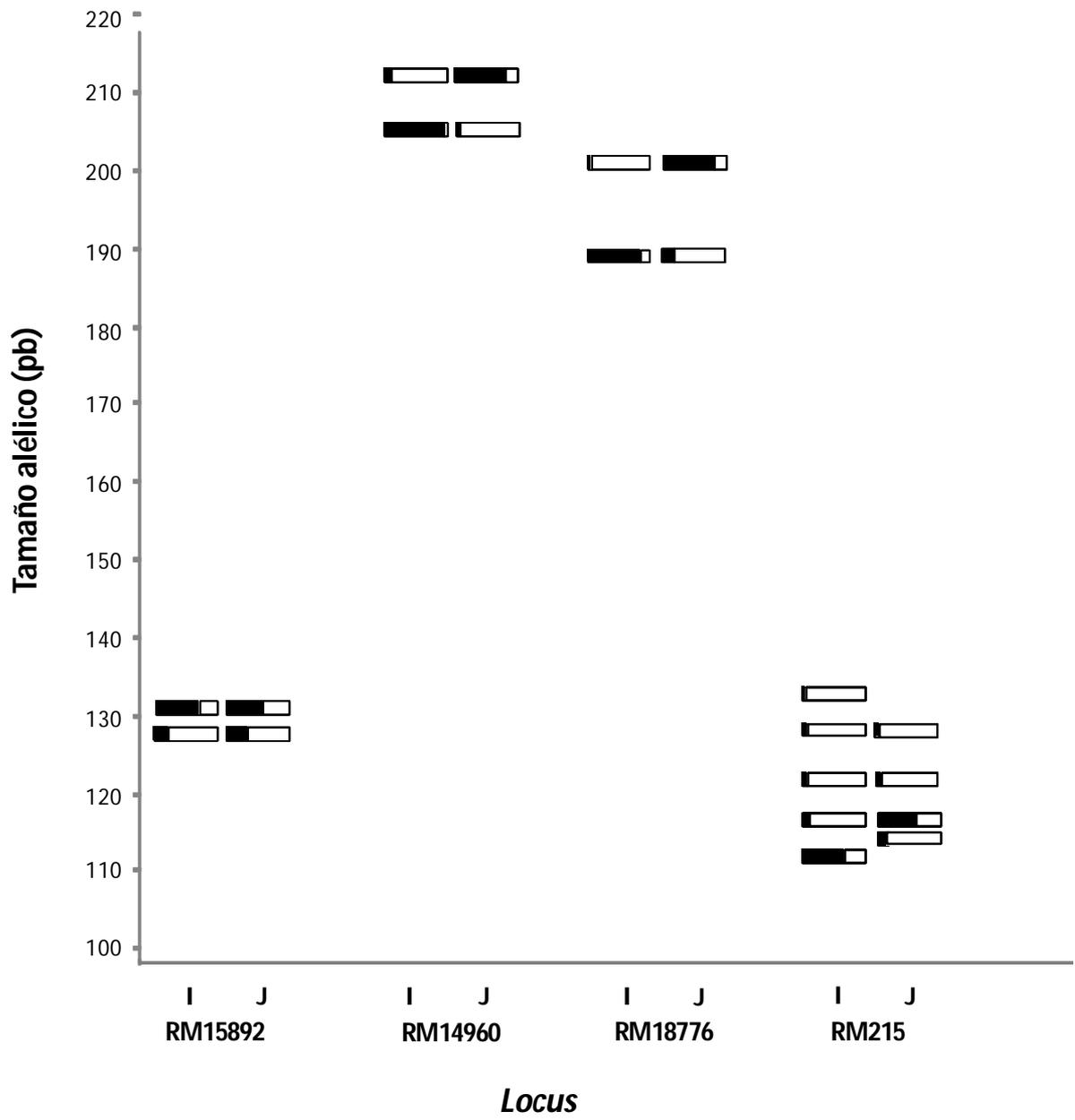
Tabla 4: Frecuencia de codones y aminoácidos codificados en genes de tolerancia a frío.

Codón	Aminoácido codificado	%
GGA/GGG/GGC/GGT	Glicina	13,64
GCA/GCG/GCC/GCT	Alanina	18,18
ATA/ATC/ATT	Isoleucina	4,54
CCA/CCG/CCC/CC	Prolina	22,73
TAC/TCT/TCC/TCT/AGC/AG	Serina	4,54
GAA/GAG	Ácido glutámico	4,54
AAA/AAG	Lisina	4,54
CGA/CGG/CGC/CGT/AGA/AGG	Arginina	13,65
GTT/GTC/GTA/GTG	Valina	9,1
CTT/CTC/CTA/CTG/TTA/TTG	Leucina	4,54

Análisis de variabilidad alélica

Se intentó amplificar los 36 microsatélites seleccionados, para los 131 genotipos del Programa de mejoramiento de arroz de INIA, con los cebadores presentados en la Tabla 2. De los 36 marcadores analizados, 6 no lograron amplificarse, 14 resultaron monomórficos y 16 polimórficos (Tabla 2). De los 14 marcadores localizados en exones, 8 resultaron ser monomórficos, mientras que 3 no amplificaron, y sólo 2 fueron polimórficos (RM14960 y RM12403). El número de alelos por locus varió entre 2 (RM6547) y 15 (RM23550). Si bien existen alelos exclusivos para varios marcadores en una u otra subespecie, para varios de estos marcadores se observa la ocurrencia de dos alelos con frecuencias marcadamente diferentes entre ambas subespecies. Esto último se cumple principalmente para los marcadores RM22034, RM14978, RM6651, RM14960, RM6547, RM18776 y RM215 (Figura 1).





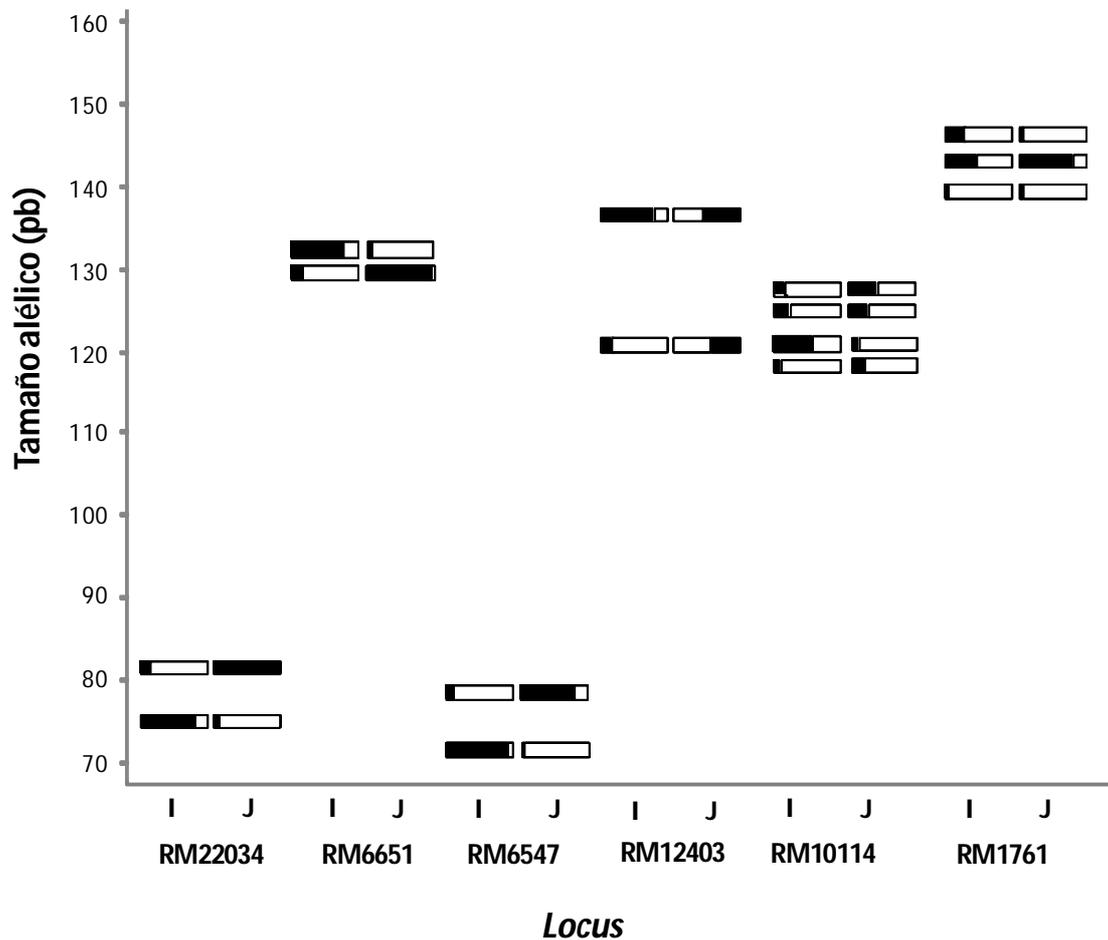


Figura 1: Variación alélica para los 16 marcadores de arroz seleccionados. Se muestran los tamaños de los alelos en pares de bases para cada marcador (barras horizontales) pertenecientes a cada genotipo, *indica* (izquierda) y *japonica* (derecha), y en negro se representa la frecuencia relativa de cada alelo en cada subespecie.

Los análisis de varianza (AMOVA) mostraron que la mayor parte de la diversidad genética (58% en el primero y 54% en el segundo) se encuentra entre ambos ecotipos, mientras que el resto de la diversidad se halla dentro de cada ecotipo en ambos casos (Tabla 5).

Tabla 5: Análisis de la varianza molecular para los 131 individuos del programa de mejoramiento de arroz de INIA.

Por ecotipo: <i>indica/japonica</i>					
Fuente de variabilidad	gl	SC	CM	Est. Var.	Porcentaje de varianza
Interpoblacional	1	488,382	488,382	4,019	58%
Intrapoblacional	240	686,047	2,859	2,859	42%
Total	241	1174,428		6,878	100%
Por programa de mejoramiento: <i>indica-INIA/indica-FLAR/japonica</i>					
Fuente de variabilidad	gl	SC	CM	Est. Var.	Porcentaje de varianza
Interpoblacional	2	504,753	252,377	3,301	54%
Intrapoblacional	239	675,234	2,825	2,825	46%
Total	241	1179,987		6,126	100%

Análisis de estructura

Los valores promedio de LnPD muestran un aumento sostenido hasta alcanzar un máximo en K=6 (Figura 2), a partir del cual los valores promedio de LnPD tienden a disminuir.

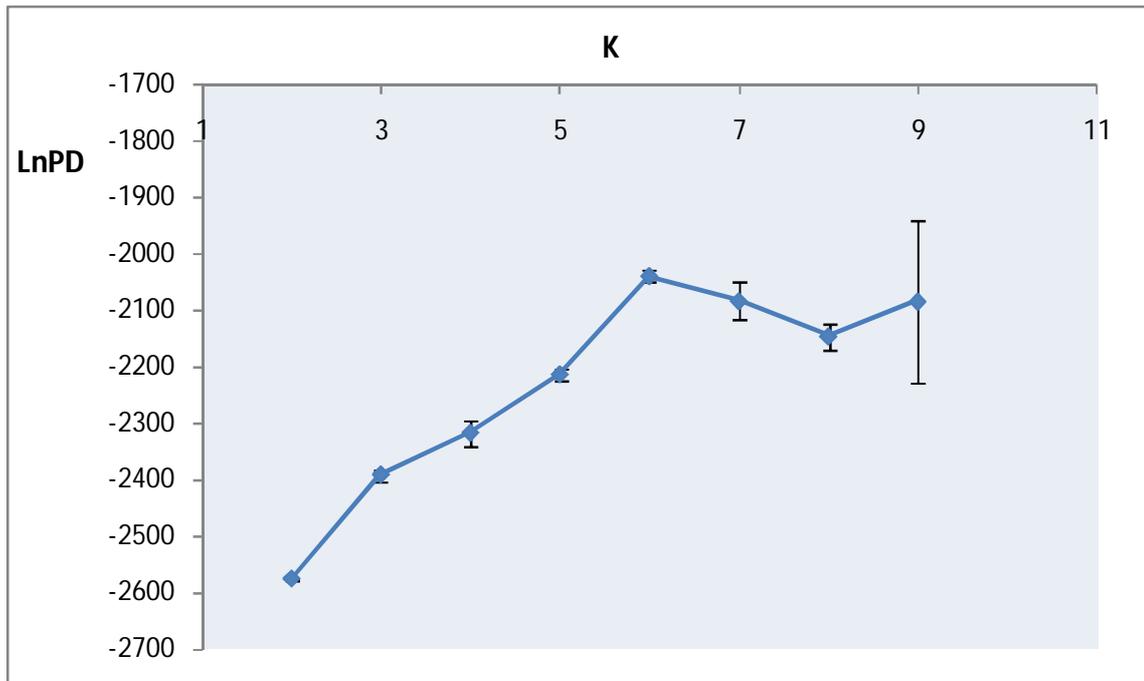


Figura 2: Representación gráfica de los valores promedio de LnPD obtenidos mediante el programa Structure versión 2.3.3 para diferentes valores de K entre 2 y 9.

A continuación se realizaron las simulaciones con información previa sobre poblaciones, asumiendo tres valores de K (Figura 3). Dos de ellos, $K = 2$ y $K = 3$, son asignados *a priori* asumiendo *indica* y *japonica* como poblaciones diferentes en el primero y asumiendo los grupos *indica*-INIA, *indica*-FLAR y *japonica* como poblaciones diferentes en el segundo. La tercera simulación se llevó a cabo asumiendo un valor de $K = 6$, tomando en cuenta los 6 grupos obtenidos a partir del análisis de estructura genética. Cuando el valor de K es igual a 2, se distingue una clara diferenciación entre los genotipos *indica* y *japonica*, aunque se observa de todos modos la presencia de 6 componentes del grupo *indica*-INIA (L3790CA, L4816, L4820, L5850, L5879 y L5855), con un alto porcentaje (llegando incluso al 100%) de alelos compartidos con *japonica*. Por su parte, el grupo *indica*-FLAR contiene solamente 3 genotipos (FL01986-16P-2-5-1, FL04225-CA-5P y FL04542-9M-6P-5M) con una totalidad de alelos compartidos con el grupo *japonica* para estos marcadores (Figura 3).

Cuando K toma el valor de 3, el grupo *indica* se divide en dos poblaciones, las cuales corresponden a los grupos *indica*-INIA e *indica*-FLAR, manteniendo ambos grupos los mismos genotipos con elevada pertenencia al grupo *japonica*. Este grupo a su vez contiene un componente (L6334) con un elevado porcentaje de alelos de los otros grupos *indica*, principalmente de *indica*-FLAR.

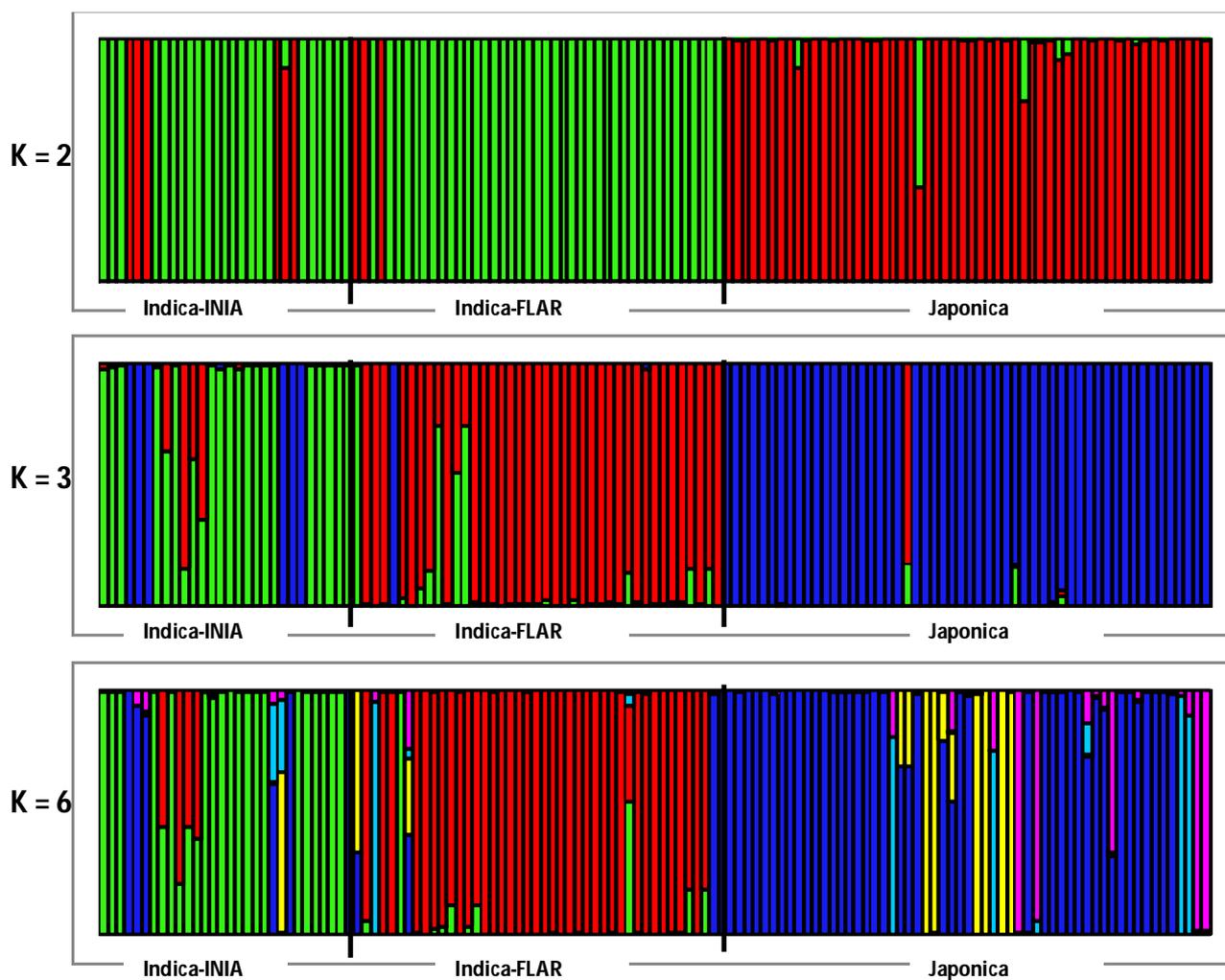


Figura 3: Poblaciones de arroz analizadas mediante el programa Structure asumiendo 3 valores de K (K=2, K=3 y K=6).

Cuando K toma el valor de 6, se mantienen los mismos tres grupos principales, donde ambos grupos *indica* mantienen prácticamente los mismos niveles de estructura interna. Por su parte, el grupo *japonica* muestra una mayor diferenciación en subpoblaciones, apareciendo esta estructura también en aquellos genotipos *indica* que comparten alelos con el grupo *japonica*. En este caso, el genotipo L6334 (*japonica* tropical), que en K =3 portaba alelos asignados al grupo *indica* en su extensa mayoría, pasa a portar alelos asignados a *japonica*. Por otro lado, se mantienen los 9 genotipos con un elevado porcentaje de alelos asignados al grupo *japonica* (6 dentro del grupo *indica*-INIA y 3 individuos dentro del grupo *indica*-FLAR). Cuando se realizó nuevamente el análisis de frecuencias alélicas quitando estos 6 individuos, la frecuencia de los alelos con menor representación dentro de cada grupo en los marcadores RM22034, RM14978, RM6651, RM14960, RM6547, RM18776 y RM215 se

vio drásticamente disminuida. Para algunos marcadores incluso, estas frecuencias parcialmente minoritarias llegaron a igualarse a 0. Tal fue el caso de los marcadores RM14978, RM14960 y RM6547, los cuales se encuentran dentro de los genes que codifican para una glutatión peroxidasa (LOC_Os03g24380), para una proteína de respuesta a estrés por sequía (LOC_Os03g23010) y para un factor de transcripción de tipo NAC (LOC_Os01g60020), respectivamente. Esto indicaría que dichos individuos fueron responsables en gran medida de la distribución sesgada de las frecuencias alélicas observada para estos marcadores, lo cual no hubiera sido posible de detectar sin el análisis con Structure.

Análisis fenotípico

Se obtuvieron los valores de fluorescencia de clorofila a partir de hoja (Fv/Fm) para determinar la máxima eficiencia cuántica del PSII. Dicha medida (Fv/Fm) constituye el mejor indicador de los niveles de fotoinhibición.

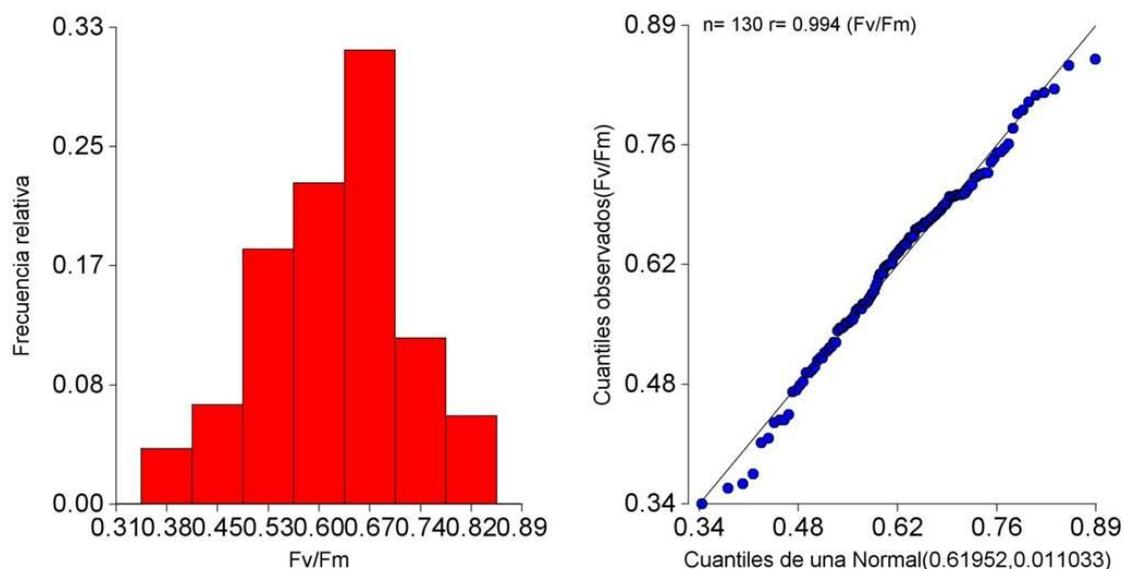


Figura 4: Histograma de frecuencias (izquierda) y gráfico cuantiles de los valores de Fv/Fm para 131 genotipos de arroz.

Los valores de Fv/Fm se hallaron dentro del rango de 0,167 y 0,867, ajustándose a una distribución normal (Figura 4). El valor de Fv/Fm promedio para los genotipos *japonica* (0,121) fue superior a los valores promedio de los genotipos *indica*-INIA (0,09) e *indica*-FLAR (0,08) (Figura 5). Como criterio para la asignación de las clases se calculó el promedio de los valores de fluorescencia y su desvío estándar. A partir de esto se tomaron como susceptibles aquellos genotipos cuyos valores de Fv/Fm se hallaban por debajo del promedio menos un desvío estándar, resistentes los que mostraron un valor de Fv/Fm superior al promedio más un desvío estándar, y de respuesta moderada aquellos cuyos valores de Fv/Fm se hallaban entre ambos desvíos.

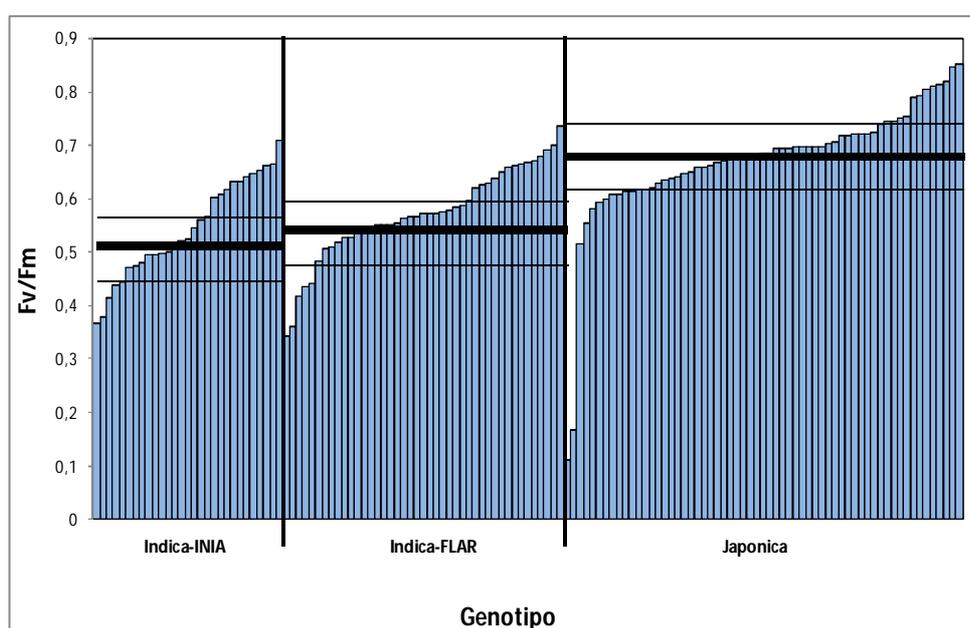


Figura 5: Valores de Fv/Fm tomados para los 131 genotipos de arroz. Se indican los valores promedio para cada grupo (líneas gruesas) y sus respectivos desvíos (líneas finas).

Tomando en cuenta la totalidad de los genotipos, el 72% mostró ser de respuesta moderada, un 14% fueron resistentes y un mismo número susceptibles. Dentro de los genotipos *indica*, el 24% resultó susceptible, mientras que ninguno mostró ser resistente, concordando con lo esperado. Entre los genotipos *japonica* el 31% resultó ser resistente, mientras que un genotipo, el cultivar Caraguatá, resultó susceptible como era de esperarse.

Análisis discriminante

El análisis discriminante se planteó para buscar cuáles marcadores resultaban más adecuados para diferenciar entre clases fenotípicas. Cuando se tomó la totalidad de los individuos ordenados en tres clases fenotípicas (susceptibles, tolerantes y moderados) como grupo de entrenamiento, los resultados del análisis fueron poco satisfactorios. En este caso, 84 de las 131 instancias fueron correctamente clasificadas (64%), con un valor del estadístico Kappa de 0,1299, y una tasa de falsos positivos igual a 0,55.

En base a este resultado, se decidió disminuir el número de dimensiones del problema para lo cual se redujo el número de clases dentro de cada grupo (*indica* y *japonica*) a dos (tolerantes y susceptibles). Dado que dentro de *indica* todos los individuos resultaron ser susceptibles y moderados, se tomaron aquellos individuos moderados como tolerantes. Del mismo modo se procedió con *japonica*.

Se realizaron tres tipos de análisis de acuerdo al criterio de clasificación. En el primero se utilizaron todos los individuos sin considerar estructura de poblaciones. El segundo criterio de clasificación estableció el análisis con dos grupos considerando separación entre ecotipos *indica* y *japonica*, y el tercero consideró el programa de mejoramiento de donde provenían los genotipos *indica* (INIA y FLAR) además del grupo *japonica*.

Dado que en este trabajo se realizaron análisis considerando diferentes grupos, las matrices debieron clonarse de forma de poder estimar el porcentaje de clasificación correcta por validación cruzada, permitiendo de esta forma hacer una evaluación del método más precisa. El clonado de los datos no afecta los resultados finales, ya que cada una de las instancias con igual perfil de marcadores SSR será asignada a la misma clase.

Tabla 6: Porcentajes de clasificación para los diferentes grupos analizados mediante análisis discriminante.

Grupos	% de Clasificación	Kappa
Sin considerar estructura	64,88%	0,2983
<i>indica</i>	97%	0,9286
<i>japonica</i>	98%	0,9543
<i>indica</i> -INIA	98%	0,9619
<i>indica</i> -FLAR	100%	1

El porcentaje de clasificación correcta para todos los genotipos sin considerar la estructura siguió resultando bajo (64,88%, Kappa = 0,2983). Por el contrario, cuando se analizaron los grupos por separado, tomando en cuenta la estructura poblacional, los porcentajes de clasificación así como los valores de Kappa aumentan considerablemente. En el caso del grupo *indica*-FLAR el porcentaje de instancias correctamente clasificadas alcanzó incluso un 100% (Tabla 6).

Tabla 7: Atributos seleccionados a través del método *best- first*. Se muestra también el porcentaje de clasificación de los mismos, así como su valor de Kappa.

Grupo	Atributos seleccionados	% Clasificación	Kappa
Sin considerar estructura	RM144_255 (100%) RM23550_280 (100%) RM23550_286 (85%) RM215_158 (100%)	74%	0,4868
<i>indica</i>	RM6651_131 (100 %) RM23550_286 (100 %) RM215_147 (100%) RM215_158 (90 %) RM1761_143 (80%)	76%	0,1282
<i>japonica</i>	RM144_291 (90 %) RM23550_250 (80 %) RM23550_280 (100 %) RM23550_380 (80 %) RM1761_152 (80 %)	81,4%	0,2938
<i>indica</i> -INIA	RM6651_131 (100 %) RM6651_134 (100 %) RM23550_296 (100 %) RM215_147 (90 %) RM1761_143 (100 %)	85%	0,6747
<i>indica</i> -FLAR	RM23550_286 (100 %) RM10114_125 (100 %) RM215_145 (100 %)	97,6%	0,8959

Cuando se utilizan las variables seleccionadas mediante el algoritmo *best first* se observa que no mejora el desempeño del método, ya que en general los valores de clasificación correcta son menores que los obtenidos con la totalidad de los datos (Tabla 7). Para el único caso en que esto no se cumple es cuando se toman en

cuenta todos los genotipos sin considerar la estructura, ya que el porcentaje de clasificación en este caso aumenta levemente (74%; $K = 0,4968$), aunque sin llegar a valores satisfactorios. Por otro lado, cuando se realiza la separación dentro del grupo *indica*, los porcentajes de clasificación, si bien disminuyen, continúan siendo satisfactorios, sobretodo dentro del grupo *indica*-FLAR. Respecto a los atributos seleccionados puede observarse que, salvo el marcador RM10114, el resto de los marcadores e incluso algunos alelos tienen alto poder discriminante en más de un grupo.

Discusión

Abundancia de microsatélites polimórficos en genes de respuesta a frío

El hecho de que varios de los genes de tolerancia a frío contengan SSRs en su secuencia esta en concordancia con resultados previos a partir de levaduras, donde los SSRs resultaron estar distribuidos preferentemente en los marcos abiertos de lectura (*Open Reading Frames- ORFs*) codificantes para proteínas regulatorias (p.ej. factores de transcripción y quinasas) en lugar de genes codificantes para proteínas estructurales (Hancock and Simon, 2005; Young et al., 2000). La importancia de los factores de transcripción ha sido muy estudiada debido al importante rol que juegan en la tolerancia de varios tipos de estrés incluyendo el estrés por bajas temperaturas (Khong et al., 2008). Esto último concuerda con la hipótesis de los SSR génicos podrían tener un rol importante en la rápida evolución de los fenotipos adaptativos (Young et al., 2000).

Por su parte, los tri-nucleótidos fueron el tipo de repetición más frecuente, lo cual concuerda con estudios previos en arroz (Lawson and Zhang, 2006a), y en otras monocotiledóneas (Morgante et al., 2002). Lawson & Zhang (2006) en su análisis encontraron que los tri-nucleótidos se encontraban más frecuentemente en todas las regiones salvo en 3'UTR, preferentemente en 5'UTR. En este estudio, si bien se encontró un solo tri-nucleótido en 3'UTR, éstos tendieron a ubicarse preferentemente en exones, pero en su composición prevalecieron los nucleótidos G y C al igual que los resultados reportados por Lawson & Zhang (2006). Estas diferencias pueden ser explicadas por la menor cantidad de genes analizados en este estudio frente a las 57,915 secuencias analizadas por Lawson & Zhang (2006). La prevalencia de repetidos de tipo tri-nucleótidos o múltiplos de tres en exones no es casual, ya que la selección natural actuaría en contra de mutaciones capaces de generar cambios en el marco de lectura de los genes (Li, 2004). Por otro lado, se encontró que el repetido más frecuente dentro de los di-nucleótidos fue AT, lo cual no concuerda con otros estudios en arroz y otros cereales donde la frecuencia de GA resultó la más elevada dentro de los di-nucleótidos. Sin embargo, tanto en arroz como en otros cereales, el tri-nucleótido CCG era el más abundante, lo cual concuerda con este estudio (Kantety et al., 2002; Lawson and Zhang, 2006a). Este tri-nucleótido codifica para el aminoácido prolina, siguiéndole en frecuencia la alanina, la arginina y la glicina, lo cual concuerda con lo reportado por Lawson & Zhang (2006).

De los SSR seleccionados para el genotipado de las 131 líneas de arroz, 6 no lograron amplificarse. Si bien es menos frecuente que en microsatélites genómicos, la

ocurrencia de alelos nulos en microsatélites génicos ya ha sido previamente reportada (Cho et al., 2000; Eujayl et al., 2001; Fraser et al., 2004). Dado que la amplificación de los microsatélites utilizados en este estudio ya fue previamente evaluada por Mc Couch et al. (2002), la ocurrencia de alelos nulos en este caso puede explicarse por mutaciones en el sitio de unión de los cebadores (Lehmann et al., 1996). De los 6 microsatélites que no amplificaron, 4 residían en exones. De este modo, es posible que la secuencia flanqueante que reconoce alguno de los cebadores diseñados para amplificarlos fuera un intrón o algún tipo de secuencia reguladora con mayor tasa de variación. Para los dos microsatélites restantes que no amplificaron puede cumplirse lo mismo. Sólo 3 de los 14 marcadores localizados en exones mostraron ser polimórficos en este estudio. Esto concuerda con el hecho de que las tasas de mutación en este tipo de marcadores están fuertemente limitadas por la selección natural, debido a las drásticas consecuencias que podría acarrear el agregado de un aminoácido en la secuencia de una proteína (Li, 2004).

Las frecuencias alélicas fueron marcadamente diferentes para cada subespecie, cumpliéndose para varios marcadores la existencia de dos alelos con una marcada diferencia de distribución entre *indica* y *japonica*. Si bien esta distribución diferencial puede ser el reflejo de una marcada estructura poblacional, no puede descartarse una posible relación funcional entre la presencia de ciertos alelos y la capacidad de adaptación al frío de cada subespecie. Esto no es casual dado que los centros de domesticación de ambos ecotipos difieren en cuanto a sus condiciones climáticas, principalmente la temperatura (Khush, 1997; Oka and Morishima, 1982).

La existencia de una marcada estructura genética en arroz ha sido bastante documentada (Garris et al., 2005; Ni et al., 2002; Zhang et al., 1992). Las accesiones analizadas en este trabajo muestran una diferenciación significativa tanto entre los grupos *indica* y *japonica* como entre subgrupos dentro de *indica*. Esto queda apoyado por el análisis de estructura. Esta estructura entre los grupos *indica* y *japonica*, así como entre subgrupos puede ser atribuida a diferentes factores. En primer lugar, no puede soslayarse un legado de estructura histórico proveniente de las poblaciones de arroz ancestrales, las cuales se estima que divergieron hace más de 100 000 años (Kovach et al., 2007; Sweeney and McCouch, 2007). Como se mencionó previamente, ambas subespecies se originaron a partir de dichas poblaciones pre-diferenciadas de *Oryza rufipogon*, su ancestro salvaje, persistiendo un cierto grado de flujo génico entre ambas, permitiendo la fijación de ciertos alelos de domesticación comunes a ambas (Sang and Ge, 2007).

Por otro lado, como ya fue mencionado previamente, la partición de la diversidad genética observada podría ser el producto de una correspondencia entre la presencia de ciertos alelos de marcadores funcionales con una capacidad de tolerancia a bajas temperaturas. Si bien el análisis de varianza molecular dio como resultado que una

mayor proporción de la diversidad genética se encontraba entre las poblaciones (58% considerando división entre ecotipos, y 54% considerando dos grupos dentro de *indica*), estos porcentajes no fueron lo suficientemente elevados considerando la marcada estructura genética evidenciada en estudios previos (Caicedo et al., 2007; Garris et al., 2005). Esto podría ser explicado por un mayor nivel de diversidad dentro del grupo *indica*, considerando también sus subgrupos, lo cual se ve reflejado en el análisis de estructura. Estos mayores niveles de diversidad no son casuales, ya que *indica* constituye el grupo más diverso de todos los grupos comprendidos dentro de la especie *Oryza sativa*, y no existen evidencias sobre la existencia de cuellos de botella en la historia de dicho grupo (Garris et al., 2005).

Dentro de ambos subgrupos de *indica* se encuentran individuos con un alto porcentaje (sino una totalidad) de alelos asignados al grupo *japonica*. A su vez, estos individuos presentan valores medios de Fv/Fm, colocándolos dentro de la clase de respuesta moderada en el análisis fenotípico. Estos genotipos son: L3790CA, L4816, L4820, L5850, L5879 y L5855 (pertenecientes a *indica*-INIA), y FL01986-16P-2-5-1, FL04225-CA-5P y FL04542-9M-6P-5M (pertenecientes a *indica*-FLAR).

El genotipo L3790CA surge del cruzamiento entre la línea L1919 de origen *indica*, y el cultivar INIA Caraguatá, el cual pertenece al grupo *japonica* tropical. Los genotipos L5850, L5879 y L5855 por su parte, son producto del cruzamiento Cuaró x L1753, donde el primero es producto del cruzamiento entre dos cultivares *indica* (IRGA 419 y El Paso 144), y el segundo del cruzamiento de una línea *indica* (CT6584) con el cultivar L435 (proveniente de Texas y del grupo *japonica* tropical). Por último, los genotipos L4816 y L4820 forman parte de la progenie del cruzamiento Cuaró x Chui, siendo este último una línea *indica* desarrollada en EMBRAPA Pelotas, la cual tendría algún antepasado *japonica* de grano corto (Pedro Blanco, Juan Rosas, com. pers). Por otro lado, no pudo obtenerse información respecto a la genealogía de los genotipos provenientes de *indica*-FLAR.

Entre los marcadores mayoritariamente asignados a *japonica* en el análisis de frecuencias alélicas, 3 disminuyeron sus frecuencias a 0 en el grupo *indica* cuando los 9 individuos eran removidos del análisis. Dichos marcadores se encuentran dentro de genes que codifican para proteínas involucradas en la respuesta a frío. Si efectivamente estos alelos están involucrados en la adaptación al frío de *japonica*, (tomando en cuenta que estos genotipos presentarían un antepasado *japonica*, por lo menos en el caso de los genotipos *indica*-INIA), su presencia en genotipos *indica* seleccionados en ambientes templados podría estar otorgando una ventaja selectiva en la tolerancia a las bajas temperaturas. Dicha presencia podría atribuirse al producto de la selección por bajas temperaturas por parte de los mejoradores del programa. Una manera de poner a prueba el efecto de la selección sería realizar un genotipado

de estas mismas líneas con marcadores no asociados a ninguna característica, y evaluar de qué modo estos genotipos son agrupados luego del análisis de estructura.

Relación entre presencia de microsatélites y tolerancia a nivel fenotípico

El principal objetivo del uso de la técnica de análisis discriminante fue el de identificar la mejor combinación de marcadores moleculares que puedan ser utilizados para discriminar entre clases fenotípicas predefinidas. Los resultados indican que la totalidad de los marcadores analizados tienen un alto poder discriminante cuando se toma en cuenta la estructura poblacional, pero no al analizar el germoplasma completo. Este resultado no se condice con estudios previos en arroz, donde la estructura poblacional pareció tener muy poco impacto en la habilidad de asignar correctamente atributos a clases fenotípicas predefinidas mediante análisis discriminante (Zhang et al., 2005).

Entre los 16 marcadores polimórficos analizados, 6 mostraron mayor poder discriminante entre todos los grupos de entrenamiento analizados. Todos estos marcadores salvo el RM10114 fueron compartidos entre dos o más grupos, lo cual estaría dándole seguridad del método. Por otro lado, cuando se realizó el mismo análisis por validación cruzada con los marcadores que resultaron más informativos, los porcentajes de clasificación se vieron reducidos salvo en los grupos *indica*-INIA e *indica*-FLAR, donde si bien dichos valores resultaron más bajos, permanecieron satisfactorios. Esto se debería en parte a la baja cantidad de marcadores analizados. Por lo tanto, si bien estos 6 marcadores tienen un alto poder discriminante, no serían suficientes por sí mismos para clasificar correctamente los datos fenotípicos.

La posibilidad de determinar asociaciones funcionales con alta probabilidad dependerá del tipo de variable considerada. Cuando se trabaja con características más simples, controladas por pocos genes, y con rutas metabólicas ya caracterizadas, la probabilidad de encontrar asociaciones funcionales con un alto porcentaje de clasificación, y con un número más reducido de marcadores es más alta (Mcharo et al., 2004). Por otro lado, como fue comprobado en los resultados, el número de clases fenotípicas también influye sobre los resultados obtenidos ya que este análisis tiene un mayor poder discriminante cuando se trabaja con dos clases o grupos fenotípicos (Witten and Frank, 2005). Este hecho se vio reflejado cuando se disminuyeron las clases fenotípicas a dos (tolerantes y susceptibles), los porcentajes de clasificación correcta aumentaron. En conclusión, podría inferirse que el máximo de clasificación correcta se obtiene utilizando todos los marcadores, aunque de todos modos se mantiene una cierta constancia respecto a los marcadores con alto poder discriminante en todos los grupos analizados.

En este trabajo se lograron identificar 16 genes candidatos de respuesta a frío en arroz en etapas tempranas del desarrollo portadores de microsatélites polimórficos en su secuencia, de los cuales 7 mostraron una marcada distribución diferencial entre genotipos pertenecientes a las subespecies *indica* y *japonica*. Mediante análisis discriminante se encontró que 6 de los 16 marcadores polimórficos eran capaces de discriminar con un alto porcentaje de clasificación entre clases fenotípicas predefinidas. La asociación putativa de este conjunto de marcadores seleccionados mediante análisis discriminante, con una tolerancia a las bajas temperaturas puede ser puesta a prueba con análisis genéticos de estos mismos marcadores en poblaciones segregantes derivadas de cruzamientos controlados.

Bibliografía

- Andersen, J.R., and Lübberstedt, T. (2003). Functional markers in plants. *Trends in Plant Science* 8, 554-560.
- Ayres, N.M., McClung, A.M., Larkin, P.D., Bligh, H.F.J., Jones, C.A., and Park, W.D. (1997). Microsatellites and a single-nucleotide polymorphism differentiate apparent amylose classes in an extended pedigree of US rice germ plasm. *Theoretical Applied Genetics* 94, 773-781.
- Bonamico, N.C., Balzarini, M.G., Arroyo, A.T., Ibañez, M.A., Díaz, D.G., Salerno, J.C., and Renzo, M.A.D. (2010). Association between microsatellites and resistance to Mal de Río Cuarto in maize by discriminant analysis. *International Journal of Experimental Botany* 79, 31-38.
- Caicedo, A.L., Williamson, S.H., Hernandez, R.D., Boyko, A., Fledel-Alon, A., York, T.L., Polato, N.R., Olsen, K.M., Nielsen, R., McCouch, S.R., *et al.* (2007). Genome-Wide Patterns of Nucleotide Polymorphism in Domesticated Rice. *PLoS Genetics* 3, 1745-1756.
- Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement* 20, 37-46.
- Counce, P.A., Keisling, T.C., and Mitchell, A.J. (2000). A Uniform, Objective, and Adaptive System for Expressing Rice Development. *Crop Science* 40, 436-443.
- Cheng, C., Yun, K.-Y., Rensom, H.W., Mohanty, B., Bajic, V.B., Jia, Y., Yun, S., and de los Reyes, B.G. (2007). An early response regulatory cluster induced by low temperature and hydrogen peroxide in seedlings of chilling-tolerant japonica rice. *BMC Genomics* 8, 175.
- Cho, Y.G., Ishii, T., Temnykh, S., Chen, X., Lipovich, L., McCouch, S.R., Park, W.D., Ayres, N., and Cartinhour, S. (2000). Diversity of microsatellites derived from genomic libraries and GenBank sequences in rice (*Oryza sativa* L.). *TAG Theoretical and Applied Genetics* 100, 713-722.
- de los Reyes, B.G., Morsy, M., Gibbons, J., Varma, T.S.N., Antoine, W., McGrath, J.M., Halgren, R., and Redus, M. (2003). A snapshot of the low temperature stress transcriptome of developing rice seedlings (*Oryza sativa* L.) via ESTs from subtracted cDNA library. *TAG Theoretical and Applied Genetics* 107, 1071-1082.
- Eujayl, I., Sorrells, M., Baum, M., Wolters, P., and Powell, W. (2001). Assessment of genotypic variation among cultivated durum wheat based on EST-SSRS and genomic SSRS. *EUPHYTICA* 119, 39-43.
- Excoffier, L., Smouse, P.E., and Quattro, J.M. (1992). Analysis of Molecular Variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics Society of America* 131, 479-491.
- Fisher, R.A. (1936). The Use of Multiple Measurements in Taxonomic Problems. *Annals of Eugenics* 7, 179-188.
- Fraser, L.G., Harvey, C.F., Crowhurst, R.N., and Silva, H.N. (2004). EST-derived microsatellites from *Actinidia* species and their potential for mapping. *TAG Theoretical and Applied Genetics* 108, 1010-1016.
- Garris, A.J. (2004). Genetic Structure and Diversity in *Oryza sativa* L. *Genetics* 169, 1631-1638.

- Hirotsu, N. (2005). The Photosynthetic Properties of Rice Leaves Treated with Low Temperature and High Irradiance. *Plant and Cell Physiology* 46, 1377-1383.
- Ishitani, M., Xiong, L., Stevenson, B., and Zhu, J.-K. (1997). Genetic Analysis of Osmotic and Cold Stress Signal Transduction in Arabidopsis: Interactions and Convergence of Abscisic Acid-Dependent and Abscisic Acid-Independent Pathways. *The Plant Cell* 9, 1935-1949.
- Jeong, S.W., Choi, S.M., Lee, D.S., Ahn, S.N., Hur, Y., Chow, W.S., and Park, Y.I. (2002). Differential susceptibility of photosynthesis to light chilling stress in rice (*Oryza sativa* L.) depends on the capacity for photochemical dissipation of light. *Molecules and cells* 13, 419-428.
- Kantety, R.V., La Rota, M., Matthews, D.E., and Sorrels, M.E. (2002). Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat. *Plant Molecular Biology* 48, 501-510.
- Khong, G.N., Richaud, F., Coudert, Y., Pati, P.K., Santi, C., Périn, C., Breitler, J.-C., Meynard, D., Vinh, D.N., Guiderdoni, E., *et al.* (2008). Modulating Rice Stress Tolerance by Transcription Factors. *Biotechnology and Genetic Engineering Reviews* 25, 381-403.
- Khush, G.S. (1997). Origin, dispersal, cultivation and variation of rice. *Plant Molecular Biology* 35, 25-34.
- Kovach, M.J., Sweeney, M.T., and McCouch, S.R. (2007). New insights into the history of rice domestication. *Trends in Genetics* 23, 578-587.
- Lawson, M., and Zhang, L. (2006). Distinct patterns of SSR distribution in the Arabidopsis thaliana and rice genomes. In *Genome Biology*, pp. R14.
- Lehmann, T., Hawley, W.A., and Collins, F.H. (1996). An Evaluation of Evolutionary Constraints on Microsatellite Loci Using Null Alleles. *Genetics* 144, 1155-1163
- Li, Y.C. (2004). Microsatellites Within Genes: Structure, Function, and Evolution. *Molecular Biology and Evolution* 21, 991-1007.
- M.G.Murray, and W.F.Thompson (1980). Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Research* 8, 4321-4325.
- Maxwell, K., and Johnson, G.N. (2000). Chlorophyll fluorescence-a practical guide *Journal of Experimental Botany* 51, 659-668.
- McCouch, S.R., Teytelman, L., Xu, Y., Lobos, K.B., Clare, K., Walton, M., Fu, B., Maghirang, R., Li, Z., Xing, Y., *et al.* (2002). Development and Mapping of 2240 New SSR Markers for Rice (*Oryza sativa* L.). *DNA Research* 9, 199-207.
- Mcharo, M., Labonte, D.R., Oard, J.H., Kays, S.J., and McLaurin, W.J. (2004). Linking quantitative traits with AFLP markers in sweetpotatoes using discriminant analysis. *Acta Horticulturae* 637, 285-293.
- Morgante, M., Hanafey, M., and Powell, W. (2002). Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. *Nature Genetics* 30, 194-200.
- Murashige, T., and Skoog, F. (1962). A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures. *Physiologia Plantarum* 15, 473-497.
- Ni, J., Colowit, P.M., and Mackill, D.J. (2002). Evaluation of Genetic Diversity in Rice Subspecies Using Microsatellite Markers. *Crop Science* 42, 601-607.

- Nishiyama, Y., Allakhverdiev, S.I., and Murata, N. (2006). A new paradigm for the action of reactive oxygen species in the photoinhibition of photosystem II. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* 1767, 414-421.
- Oka, H.-I., and Morishima, H. (1982). Phylogenetic differentiation of cultivated rice, XXIII. Potentiality of wild progenitors to evolve the Indica and Japonica types of rice cultivars. *Euphytica* 31, 41-50.
- Pritchard, J.K., Stephens, M., and Donnelly, P. (2000). Inference of Population Structure Using Multilocus Genotype Data. *Genetics* 155, 945-959.
- Rabbani, M.A., Maruyama, K., Abe, H., Khan, M.A., Katsura, K., Ito, Y., Yoshiwara, K., Seki, M., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2003). Monitoring Expression Profiles of Rice Genes under Cold, Drought, and High-Salinity Stresses and Abscisic Acid Application Using cDNA Microarray and RNA Gel-Blot Analyses. *Plant Physiology* 133, 1755-1767.
- Sang, T., and Ge, S. (2007). The Puzzle of Rice Domestication. *Journal of Integrative Plant Biology* 49, 60-768.
- Shinozaki, K., and Yamaguchi-Shinozaki, K. (2000). Molecular responses to dehydration and low temperature. *Current Opinion in Plant Biology* 3, 217-223.
- Sweeney, M., and McCouch, S. (2007). The Complex History of the Domestication of Rice. *Annals of Botany* 100, 951-957.
- Varshney, R.K., Graner, A., and Sorrells, M.E. (2005). Genic microsatellite markers in plants: features and applications. *Trends in Biotechnology* 23, 48-55.
- Verhoeven, A.S., Demmig-Adams, B., and III, W.W.A. (1997). Enhanced Employment of the Xanthophyll Cycle and Thermal Energy Dissipation in Spinach Exposed to High Light and N Stress. *Plant Physiology* 113, 817-824.
- Witten, I.H., and Frank, E. (2005). *Data Mining Practical Machine Learning Tools and Techniques* _2nd edition, Vol 2, 2 edn (San Francisco, CA, Morgan Kaufmann).
- Yoshida, S. (1981). *Fundamentals of rice crop science, Vol 1* (Manila, Philippines, The International Rice Research Institute).
- Young, E.T., Sloan, J.S., and Riper, K.V. (2000). Trinucleotide Repeats Are Clustered in Regulatory Genes. *Genetics* 154, 1053-1068.
- Yun, K.-Y., Park, M.R., Mohanty, B., Herath, V., Xu, F., Mauleon, R., Wijaya, E., Bajic, V.B., Bruskiewich, R., and de los Reyes, B.G. (2010). Transcriptional regulatory network triggered by oxidative signals configures the early response mechanisms of japonica rice to chilling stress. *BMC Plant Biology* 10, 16.
- Zhang, N., Xu, Y., Akash, M., McCouch, S., and Oard, J.H. (2005). Identification of candidate markers associated with agronomic traits in rice using discriminant analysis. *Theoretical and Applied Genetics* 110, 721-729.
- Zhang, Q., Maroof, M.A.S., Lu, T.Y., and Shen, B.Z. (1992). Genetic diversity and differentiation of Indica and japonica rice detected by RFLP analysis TAG *Theoretical and Applied Genetics* 83, 495-499.

CAPÍTULO 2

Análisis de expresión

Introducción

Los microsatélites o SSR (del inglés, *Simple Sequence Repeats*) son secuencias cortas de ADN repetido en tándem. Constituyen uno de los tipos más variables de secuencias en el genoma de los organismos vivos, y su polimorfismo radica en el número de repetidos en vez de su secuencia primaria (Ellegren, 2004). Los microsatélites tienen una amplia distribución en el genoma, lo cual hace posible clasificarlos de acuerdo a su asociación a secuencias transcritas (exones, intrones y regiones reguladoras), o a regiones no codificantes (regiones intergénicas). Este tipo de clasificación estaría relacionado con las fuerzas mutacionales y selectivas que actuarían sobre ellos. Generalmente se asume que aquellos que se localizan en regiones no transcritas tienen un mecanismo de evolución neutral. Este tipo de microsatélite es el generalmente utilizado como marcador molecular, y por tanto su frecuencia y distribución estaría reflejando dicho proceso evolutivo. Por otro lado, en las regiones transcritas (exones, intrones y regiones reguladoras), la expansión de los microsatélites estaría sometida a presiones selectivas, lo cual podría restringir sus mecanismos de evolución (Ellegren, 2004; Hancock and Simon, 2005; Jarne and Lagoda, 1996; Li, 2004). Entre los microsatélites presentes en regiones transcritas, aquellos localizados en exones son los menos frecuentes, y dentro de estos, la mayoría constituyen repetidos de tri-nucleótidos o múltiplos de éstos. Esto se debe a que este tipo de microsatélites son mejor tolerados dado que sus polimorfismos resultan en la inserción o deleción de un aminoácido. Sin embargo, los repetidos de tipo mono- o di-nucleótidos (los cuales son los más frecuentes en el genoma) serían fuertemente seleccionados en contra (Hancock and Simon, 2005).

Numerosos análisis sobre la composición y variación del número de repetidos de los microsatélites en diferentes especies vegetales revelaron que la distribución de este tipo de secuencias no es al azar (Kashi and King, 2006; Li, 2004). Análisis comparativos de secuencias genómicas y ESTs en vegetales mostraron que las frecuencias de microsatélites eran mayores en regiones transcritas del genoma (Morgante et al., 2002). La presencia de microsatélites en regiones transcritas y su composición no aleatoria sugieren que estos podrían tener un rol funcional en la expresión y/o

regulación génica (Li, 2004). De este modo, los polimorfismos en diferentes regiones génicas podrían determinar cambios fenotípicos en un organismo. Esto ha sido estudiado en procariotas, donde la presencia de microsatélites génicos variables está relacionada a la patogenicidad (Belkum et al., 1998). En eucariotas, el efecto fenotípico de los microsatélites en zonas codificantes ha sido principalmente estudiado en humanos, dada su implicancia en enfermedades como cáncer y desórdenes neurológicos (Cummings and Zoghbi, 2000; Gatchel and Zoghbi, 2005).

El posible efecto fenotípico de la expansión o contracción de un microsatélite génico depende de su posición dentro del gen. Los microsatélites en los 3'UTRs estarían involucrados en el silenciamiento génico o en el deslizamiento de la enzima ARN polimerasa durante la transcripción, mientras que los microsatélites localizados en los 5'UTR, tendrían efectos en la transcripción y/o regulación de la expresión. Los microsatélites localizados en intrones pueden tener efectos sobre la transcripción, regulación, el silenciamiento génico y sobre el correcto corte y empalme del ARN mensajero. Por último, los microsatélites situados en regiones codificantes pueden activar o inactivar genes o bien, dar lugar a proteínas truncas (revisado por Li et al., 2004).

Las evidencias sobre la dependencia entre la expresión y otras funciones génicas, y el número de copias de sus microsatélites asociados son cada vez más frecuentes (revisado por Li et al., 2004). Como se discutió en el primer capítulo de esta tesis, las diferentes condiciones de estrés a las que puede estar sometido un organismo, incluidas las bajas temperaturas, serían capaces de proporcionar una ventaja selectiva de determinadas variantes en el número de repetidos de los microsatélites génicos. Estas variaciones podrían afectar o bien la función, o la regulación de la expresión de dichos genes. Esto sería especialmente importante para el caso de los genes de respuesta a frío, ya que la presencia de ciertas variantes alélicas podría otorgar ventajas adaptativas (tolerancia) a aquellos organismos portadores. Por lo tanto, el análisis de la relación entre ciertas variantes alélicas con la expresión de los genes de resistencia a frío sería de gran importancia para lograr un mejor conocimiento sobre los mecanismos de tolerancia a estos tipos de estrés.

Cuantificación de la expresión génica

Entre los métodos de cuantificación de la expresión génica, el PCR en tiempo real es el más utilizado en la actualidad. En comparación con otros métodos clásicos de cuantificación, como Northern-blot o RT-PCR, las principales ventajas del PCR en Tiempo Real son su rapidez, su reproducibilidad, y su alta sensibilidad y especificidad (Pfaffl, 2004).

Dado que para muchas aplicaciones biológicas no es necesario conocer la cantidad absoluta de ARNm en una muestra, la cuantificación relativa constituye la estrategia más utilizada. Esta aproximación permite determinar la cantidad de ARNm de un gen de interés en relación a un gen cuya expresión permanece constante bajo las condiciones de estudio (Livak and Schmittgen, 2001). La normalización la expresión del gen de interés respecto a la expresión de un gen de referencia, permite compensar las posibles diferencias de concentración del cDNA del gen objetivo en muestras individuales generadas por diferencias en la cantidad y calidad del material de partida, o diferencias que pueden tener lugar durante el proceso de retrotranscripción (Bustin, 2002; Čikoš et al., 2007; Radonić et al., 2004). Por este motivo, resulta de gran importancia que la expresión del gen de referencia permanezca constante entre las diferentes células del organismo y bajo diferentes condiciones experimentales, ya que de otro modo puede conducir a resultados erróneos.

Entre los genes normalizadores más utilizados en estudios de expresión se encuentran el ARNr 18S y 28S, la ubiquitina, actina, b-tubulina, y la gliceraldehido-3-fosfato deshidrogenasa (GAPDH). Sin embargo, varios estudios han confirmado que la expresión de estos genes puede variar entre diferentes tipos celulares y bajo diferentes condiciones experimentales (Robinson et al., 2007; Ruan and Lai, 2007; Schmittgen et al., 2000), lo cual los hace inadecuados para ciertos estudios de expresión.

Los genes constitutivos pueden ser variables en cuanto a su expresión, estando sujetos a cambios inducidos por las diferentes condiciones experimentales a las que pueden estar sometidos. Por este motivo surge la necesidad de realizar estudios de validación de genes de referencia para diferentes organismos bajo condiciones experimentales específicas (Brunner et al., 2004; Nicot, 2005; Robinson et al., 2007; Ruan and Lai, 2007; Schmittgen et al., 2000). Un estudio previo sobre genes constitutivos en arroz, encontró que el gen del ARNr 18S era el más estable y apropiado para ser utilizado como normalizador en estudios de expresión (Kim et al., 2003). Una de las principales desventajas del uso de este gen como normalizador son sus muy elevados niveles de expresión, y por ende, el requerimiento de usar muestras de ADNc muy diluidas de modo que entren dentro del rango dinámico del PCR, sobre todo cuando los genes de interés a cuantificar tienen niveles de expresión mucho más débiles. Además, su transcripción es llevada a cabo por la ARN polimerasa I, haciendo que su regulación sea independiente de la regulación de la transcripción de ARN mensajero, la cual es llevada a cabo por la ARN polimerasa II. Esto lo haría inapropiado para el estudio de expresión de la transcripción de los genes de interés, ya que pueden haber desbalances entre la expresión de ARNr y del ARNm entre diferentes muestras (Bustin, 2002). En otro estudio más reciente sobre la expresión de 10 genes constitutivos en arroz, se encontró que los genes eEF-1 α y UBO5 mostraban los niveles

más estables de expresión entre diferentes tejidos, mientras que el gen de ARNr 18S era el más estable bajo diferentes tratamientos experimentales (Jain et al., 2006b).

En arroz se han realizado análisis sobre la expresión de diversos genes comúnmente considerados "*housekeeping*" sobre diferentes cultivares, y bajo diversas condiciones experimentales como radiación UV (Kim et al., 2003), tratamiento con hormonas, estrés hídrico y salino (Jain et al., 2006). Sin embargo, aún no hay disponibles estudios publicados sobre validación de genes de referencia para estrés por frío en arroz.

En esta tesis se realizó una revisión bibliográfica con el fin de encontrar los genes de referencia más utilizados en estudios sobre arroz y se seleccionaron aquellos comúnmente usados como referencia en trabajos previos (Cheng et al., 2007; Jain et al., 2006b; Kim et al., 2003), bajo diferentes condiciones de estudio (Tabla 1). Los genes seleccionados fueron actina, el gen del ARNr 18S, Ubiquitina 10 y el factor de iniciación eucariótico 4a. La expresión de estos 4 genes fue analizada con el objetivo de seleccionar aquél cuya expresión fuera más constante de acuerdo a las condiciones experimentales impuestas en este trabajo.

Genes candidatos asociados a la respuesta a frío en arroz

En el primer capítulo se realizó un análisis de polimorfismos de microsatélites localizados en genes de respuesta a bajas temperaturas en 131 líneas de arroz. En base a esta información se realizó un estudio de clasificación entre fenotipos y alelos mediante análisis discriminante. Este análisis dio como resultado que los marcadores con mayor poder discriminante fueron aquellos localizados en los siguientes genes: Hs70, ROS-bzip1, factor de transcripción de tipo MYB, proteína inducida por auxina (AIP), ferritina y factor de transcripción de tipo ERF.

En este capítulo se realizará un estudio de cuantificación relativa de la expresión de genes de respuesta a bajas temperaturas en tres genotipos de arroz uruguayo, los cuales difieren en cuanto a su grado de tolerancia al frío. Este análisis de cuantificación se realizará mediante PCR en Tiempo Real, la cual será relativa a la expresión de un gen constitutivo. Por otro lado se seleccionaron los genes de interés para este estudio de acuerdo a los resultados obtenidos en el capítulo anterior. El principal criterio de elección de los genes de interés fue la presencia de polimorfismo del microsatélite génico entre los tres genotipos seleccionados. Se seleccionaron los 4 genes que además de cumplir con este requisito presentaron un porcentaje alto de clasificación en el Análisis Discriminante realizado en el Capítulo 1. Estos genes fueron Hs70, ROS-bzip1, ferritina y AIP. El resto de los genes de estudio en este análisis (PP2Ca y AMPds) fueron seleccionados según el criterio principal, además de su relevancia en la

respuesta a frío analizada en trabajos anteriores a pesar de no haber estado incluidos entre los que mostraron mayor poder discriminante en el análisis realizado en el Capítulo 1. (Oh, 2005; Rabbani et al., 2003; Yun et al., 2010).

Objetivos

Objetivo específico 1: Validar genes de referencia para estudios de expresión génica bajo estrés por frío en tres diferentes genotipos de arroz.

Objetivo específico 2: Analizar y correlacionar la expresión de los genes con variabilidad alélica observada (genes candidatos) en tres variedades contrastantes de arroz mediante PCR en tiempo real.

Hipótesis

Hipótesis objetivo específico 1: Los genes tradicionalmente utilizados como normalizadores en estudios de expresión han mostrado expresarse diferencialmente en diferentes tipos celulares y bajo diferentes condiciones experimentales. De todos modos, podría esperarse que al menos uno de los genes seleccionados para la validación presente expresión constante entre los genotipos seleccionados para el ensayo de expresión y bajo las condiciones experimentales impuestas.

Hipótesis objetivo específico 2: Existe una relación entre los diferentes niveles de expresión de los genes de tolerancia a frío, y las diferentes variantes alélicas presentes en genotipos resistentes, susceptibles y de tolerancia intermedia.

Materiales y métodos

Material Vegetal

Para los tratamientos de frío y posteriores ensayos de expresión génica se eligieron tres genotipos de arroz uruguayo en virtud de su grado de tolerancia al frío. El Paso 144 es una variedad de la subespecie *indica* susceptible a las bajas temperaturas, y constituye la variedad más cultivada en el país. Las variedades INIA Tacuarí y L2825CA pertenecen a la subespecie *japonica*, y presentan diferentes estrategias de tolerancia al frío (Bonnecarrère et al., 2011). Las semillas correspondientes a estos tres genotipos fueron obtenidas a partir del programa de mejoramiento genético de arroz de INIA Treinta y Tres.

Las semillas correspondientes a los tres genotipos fueron embebidas en agua destilada estéril por 10 min, y luego esterilizadas superficialmente mediante inversiones consecutivas en etanol 70% por 3 min e hipoclorito de sodio al 50 % 3 veces por 15 min cada vez. Las semillas fueron posteriormente enjuagadas 3 veces con agua destilada estéril, y colocadas en placas de Petri sobre gasa embebida en agua e incubadas a 30 °C por 4 días para favorecer la germinación.

Tratamientos de frío

Una vez germinadas las semillas, se hicieron crecer las plántulas en un sistema de hidroponía optimizado para arroz, en medio de crecimiento Murashige & Skoog sin sacarosa (Bonnecarrère et al., 2011). Las plántulas fueron así cultivadas en cámara de crecimiento con un régimen de alternancia de períodos de luz de 16 hs y períodos de oscuridad de 8 hs, y con $300 \mu\text{E m}^{-2} \text{s}^{-1}$ de irradiación en la superficie foliar, con una temperatura de 28 °C por 13 días. Luego se separaron las plantas control de cada genotipo por triplicado, y el resto fue sometido a tratamiento de baja temperatura (10 °C, en cámara de frío), por ½, 1, 2, 4, 8, y 24 h. El material vegetal correspondiente a cada genotipo y tratamiento fue colectado por triplicado y conservado a -80 °C.

Extracciones de ARN

Las extracciones de ARN total fueron realizadas usando el reactivo Trizol (Invitrogen) de acuerdo a las instrucciones aportadas por el fabricante. Todas las

muestras de ARN fueron resuspendidas en agua libre de RNAsas y cuantificadas mediante NanoDrop ND8000 (Thermo Scientific). Fueron consideradas para el análisis aquellas muestras con una $A_{260/280}$ (indicador de contaminación con proteínas) entre 1,9 y 2,0 y con una $A_{260/230}$ (indicador de contaminación con reactivo) mayor a 1,8. A su vez, la integridad de dichas muestras fue verificada mediante electroforesis en geles de agarosa al 1%.

Las muestras utilizadas para el ensayo de validación de genes de referencia fueron además analizadas mediante Bioanalyzer Agilent 2100 (Agilent Technologies, USA). Esto permite la obtención de medidas precisas de concentración. Además este sistema proporciona la medida de RIN (*RNA Integrity Number*), la cual es una herramienta que le asigna un valor de integridad a una muestra basándose no sólo en la relación entre las bandas ribosomales, sino en toda la corrida electroforética. El valor de integridad asignado es luego clasificado en una escala del 1 al 10, donde un valor de 1 significa totalmente degradado y un valor de 10 significa que la muestra está intacta (Imbeaud et al., 2005; Schroeder et al., 2006). El objetivo de estas mediciones fue asegurar el cargado de una misma cantidad de ARN de partida para las reacciones de retrotranscripción entre todas las muestras, lo cual es importante cuando se realizan estudios de validación de genes de referencia (Huggett et al., 2005).

Diseño de cebadores

Cebadores para genes constitutivos. Se realizó una revisión bibliográfica sobre genes frecuentemente utilizados como referencia en estudios de cuantificación, mediante PCR en tiempo real en arroz, en condiciones de bajas temperaturas y otros tipos de estrés abiótico. Dicha revisión se realizó con el objetivo de identificar un conjunto de genes que serían sometidos a validación, con el objetivo de seleccionar aquél cuya expresión sea más estable bajo las condiciones de estudio impuestas en este trabajo.

Cebadores en genes candidatos. Los cebadores utilizados para amplificar los genes ROSbzip-1, AIP y Ferritina fueron obtenidos a partir de estudios transcriptómicos en arroz sometido a bajas temperaturas (Cheng et al., 2007). Para la amplificación del resto de los genes se utilizaron cebadores diseñados a través del software Primer3Plus (<http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi>). Los parámetros establecidos fueron una temperatura de fusión de 60 °C, un porcentaje de GC cercano al 70%, y un tamaño de amplicón de entre 80 y 200 pares de bases. Los cebadores fueron preferentemente diseñados de modo que se expandan a lo largo de dos exones consecutivos en el ARNm, para evitar de este modo la amplificación de ADN genómico.

Sin embargo cuando esto último no era posible, los cebadores fueron diseñados de modo que ambos flanquearan un intrón, para detectar así posible contaminación con ADN genómico debido a la diferencia de tamaño entre ambos productos de PCR (a partir de genómico y de cDNA), o eventualmente, dentro de un mismo exón (Tabla 2).

La secuencia de los cebadores así seleccionados fue posteriormente analizada a través de la aplicación Oligo Analyzer 3.1, (<http://www.idtdna.com/analyzer/Applications/OligoAnalyzer/>) en busca de posibles estructuras secundarias.

PCR en tiempo Real

Retrotranscripción. El cDNA fue sintetizado a partir de 1 µg de ARN total en un volumen final de 50 µL, mediante el kit *QuantiTect Reverse Transcription kit* (QIAGEN, 2009) de acuerdo a las instrucciones del fabricante. El protocolo incluye un paso de remoción del ADN genómico, y una mezcla de cebadores oligo-dT y de hexámeros al azar, lo cual garantiza la síntesis de cDNA a partir de cualquier secuencia de transcripto, incluyendo desde las regiones 5´.

PCR cuantitativo. Los ensayos de validación y de cuantificación fueron llevados a cabo bajo las mismas condiciones. Se utilizó el kit Rotor-Gene SYBR Green PCR kit (QIAGEN). Las mezclas de PCR contenían 2 µL de una dilución 1:5 del cDNA obtenido a partir de la retrotranscripción, Rotor-Gene SYBR Green PCR Master Mix 2x y 0,3 µM de cada cebador en un volumen final de 10 µL. Para cada cebador se realizó un control de PCR sin cDNA. Los PCR fueron llevados a cabo en la plataforma Corbett Rotor Gene 6000 (QIAGEN Sample & Assay Technologies, Hilden Germany), y posteriormente analizados mediante el software Rotor Gene 6000 Series v.1.7 (QIAGEN Sample & Assay Technologies, Hilden Germany). Todos los PCR fueron llevados a cabo bajo las siguientes condiciones: 5 min a 95 °C y 40 ciclos de 5 s a 95 °C y 10 s a 60 °C. La especificidad de los amplicones fue confirmada por análisis de curva de desnaturalización (60-95 °C) y por electroforesis en geles de agarosa. Las curvas de eficiencia para cada gen fueron realizadas para dos réplicas técnicas.

Las pruebas de validación de genes constitutivos fueron llevadas a cabo para los tres genotipos de arroz (El Paso 144, Tacuarí y L2825CA) con 0 hs y 4 h de tratamiento de frío con dos réplicas técnicas para cada muestra. Por otro lado, para los estudios de expresión bajo condiciones de frío se analizaron los tres genotipos bajo todas las condiciones experimentales (0, ½, 1, 2, 4, 8, y 24 h de frío) con tres réplicas biológicas para cada muestra.

Análisis de resultados

Validación de genes constitutivos. Se realizaron curvas de eficiencia de PCR para los cuatro genes a partir de diluciones seriadas de cDNA (1000, 500, 250, 125 y 62,5 ng). Se promediaron los valores de C_T (*cycle threshold*) de cada réplica técnica para cada dilución y se graficaron. El C_T se define como el número de ciclos requeridos para que la fluorescencia alcance un cierto nivel de umbral de detección escogido arbitrariamente y es inversamente proporcional a la cantidad de ADNc blanco presente en la reacción (Pfaffl, 2004).

La eficiencia de reacción se calculó de acuerdo a la siguiente fórmula:

$$\text{Eficiencia de PCR} = (10^{[-1/\text{pendiente}]} - 1) \times 100$$

El rango de valores aceptables de pendiente es entre -3,1 y -3,6, lo cual equivale a un rango de eficiencias entre 90 y 110% (Stratagene, 2007).

Se graficaron los valores de C_T en gráficos de tipo boxplot mediante el macro de Excel QI Macros versión 2011.06 (KnowWare International Inc.), y se realizaron análisis de varianza de los datos de C_T de muestras control y tratadas mediante el programa Infostat versión 2011 (Universidad Nacional de Córdoba, Argentina. <http://www.infostat.com.ar>). La expresión de estos genes fue representada sin normalizar, como $2^{-\Delta C_T}$, donde ΔC_T es $C_{T\text{tiempo } x} - C_{T\text{tiempo } 0}$. Este valor sería un indicador del número de veces que varía la expresión del gen luego del tratamiento en relación al control. Valores bajos de $2^{-\Delta C_T}$ significan mayor estabilidad de expresión bajo ambas condiciones de estudio.

Cuantificación de la expresión. Para generar el gráfico del logaritmo del aumento de fluorescencia (ΔR_n) vs. el número de ciclo, se tomaron los datos entre los ciclos 3 y 40. Todos los gráficos de amplificación fueron analizados con un valor de umbral de 0,058. Este valor de umbral fue escogido por el programa como el más apropiado, mediante la herramienta *Auto-Find Threshold*. Los niveles de expresión (valores de C_T) fueron exportados del software y analizados en planilla de cálculo (Microsoft Excel). Se graficaron las curvas de eficiencia para cada gen siguiendo el mismo procedimiento que para la validación. Para la cuantificación se utilizó el

método del $2^{-\Delta\Delta C_T}$, el cual se aplica cuando la eficiencia de amplificación del gen de referencia y del gen blanco son similares y cercanas al 100% (Livak and Schmittgen, 2001). La expresión de cada gen fue normalizada respecto al gen de referencia escogido en la validación, restando el valor de C_T del gen de referencia al valor de C_T del gen blanco.

Se realizaron análisis de variancia de un solo factor para los valores de C_T mediante el programa Infostat versión 2011 (Universidad Nacional de Córdoba, Argentina. <http://www.infostat.com.ar>). Como test de comparación de medias se utilizó el test de Tukey.

Resultados y discusión

I. Validación de genes de referencia

Extracciones de ARN y verificación de la calidad

Las muestras cuantificadas con Bioanalyzer mostraron una calidad aceptable (Figura 1), con un valor de RIN entre 5,6 y 6,6 (Tabla 1), el cual cae dentro del rango aceptable para ensayos de cuantificación por PCR en tiempo real (Fleige and Pfaffl, 2006).

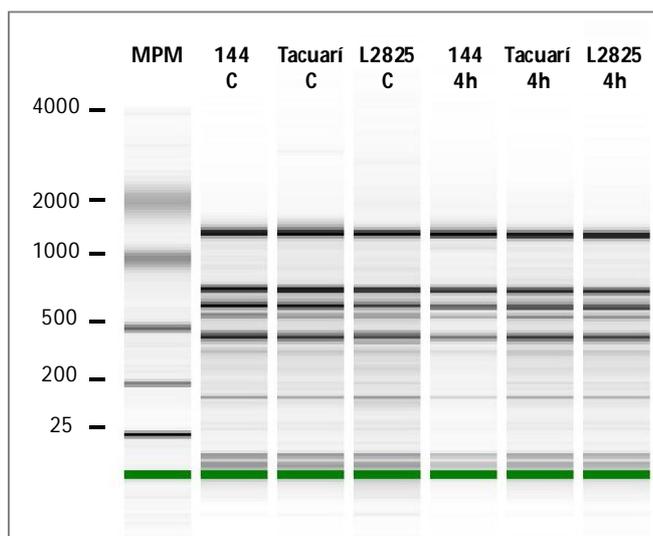


Figura 1: Representación de las muestras de ARN corridas en Bioanalyzer. Las líneas verdes representan el frente de corrida.

Cuando se comparan los valores de concentración obtenidos a partir de ambas plataformas, se observa que los valores obtenidos a partir del Nanodrop son mayores para todas las muestras (Tabla 1). Dado que uno de los criterios excluyentes de evaluación de las muestras era el de una $A_{260/280}$ superior a 1,9, y una $A_{260/230}$ superior a 1,8, una posible contaminación con proteínas o grupos fenólicos capaces de dicha sobreestimación debe ser descartada. De este modo, los valores elevados de

concentración medidos en Nanodrop pueden ser atribuidos a la presencia de ADN genómico en la muestra.

Tabla 1: Concentraciones de las muestras de ARN utilizadas para el ensayo de validación, medidas en Nanodrop y Bioanalyzer.

Muestra	Concentración (ng/ μ l)		RIN ^a
	Nanodrop	Bioanalyzer	
El Paso 144_C	1460	789	6,1
Tacuarí_C	1905	860	6,4
L2825CA_C	2148	922	6,6
El Paso 144_4h	2484	1602	6
Tacuarí_4h	2531	533	5,6
L2825CA_4h	2157	550	5,9

a. RNA Integrity Number

Validación de genes constitutivos

Se puso a prueba la expresión de los cuatro genes (actina, ARNr18S, eEF-1 α y UBQ5) frecuentemente utilizados como genes de referencia en estudios de expresión en arroz, bajo condiciones de bajas temperaturas (Tabla 2). La secuencia de los cebadores para los genes de ubiquitina, ARNr18S y eEF-1 α fue obtenida de un trabajo anterior donde se analiza la expresión de genes frecuentemente utilizados como normalizadores en arroz bajo diferentes condiciones experimentales (Jain et al., 2006b). La secuencia del cebador de actina fue tomada del trabajo de Cheng et al. (2007) donde se mide la expresión relativa de genes de respuesta a frío mediante PCR en Tiempo Real.

Tabla 2: Genes constitutivos analizados en el ensayo de validación, y cebadores utilizados para amplificarlos.

Gen	Accesión ^a	Descripción	Secuencia de los cebadores ^b	Localización de los cebadores ^c	Largo del amplicón (pb)	Referencia
ACT	AK058421	Actina	5'-CAAGCCAATCGTGAGAAG-3' 5'-AGCAATGCCAGGGAACATAGT-3'	M	487	Cheng et al. (2007)
ARNr 18S	AK059783	ARN ribosomal 18S	5'-CTACGTCCTGCCCTTTGTACA-3' 5'-ACACTTCACCGGACCATTCAA-3'	M	65	Jain et al. (2006b)
eEF-1 α	AK073620	Eukaryotic initiation factor 4a	5'-TTGTGCTGGATGAAGCTGATG-3' 5'-GGAAGGAGCTGGAAGATATCATAGA-3'	D	76	Jain et al. (2006b)
UBQ5	AK101547	Ubiquitina 5	5'-TGGTCAGTAATCAGCCAGTTTGG-3' 5'-GCACCACAATACTTGACGAACAG-3'	M	81	Jain et al. (2006b)

- Nº de accesión de la secuencia de ADNc completa.
- Secuencia del Cebador Forward (arriba) y Reverse (abajo).
- Localización de los cebadores en el mismo (M) o en diferentes (D) exones.

En primer lugar, se determinó la eficiencia de amplificación para cada uno de estos genes con el objetivo de asegurar que sus niveles de expresión fueran comparables. Se graficaron los promedios de los valores de C_T de dos réplicas técnicas, en función del logaritmo de la concentración de ADNc (Figura 2).

Los valores de eficiencia estuvieron entre 100 y 102%, lo cual permite la comparación directa de los valores de C_T de los 4 genes a analizar. Por otro lado, los valores de r^2 fueron elevados, confirmando el ajuste lineal, y garantizando así la sensibilidad del ensayo.

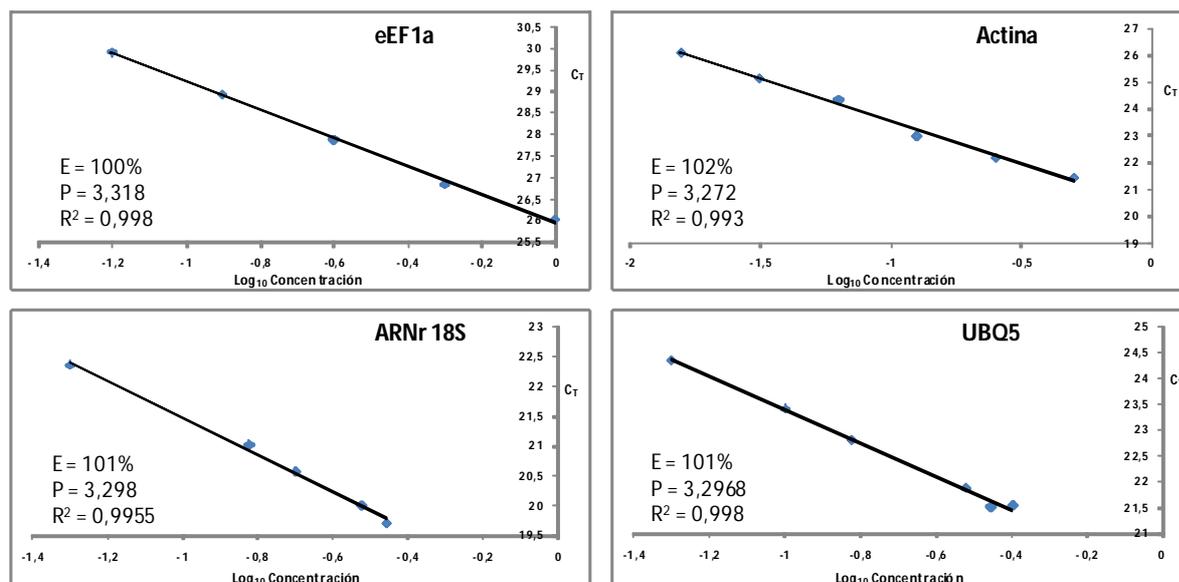


Figura 2: Curvas de eficiencia para los 4 genes del ensayo de validación. Se grafica el C_T en función del logaritmo de la concentración de ADNc. Los valores de eficiencia (E) fueron calculados a partir de la pendiente del gráfico (P).

Para analizar los diferentes niveles de expresión se compararon los valores de C_T directamente (Figura 3). El gen de ARNr 18S muestra los valores de C_T más bajos, lo cual se corresponde con niveles más elevados de transcritos iniciales, dado que ambos parámetros son inversamente proporcionales. Los genes de actina y eEF-1 α muestran valores elevados y similares de C_T , correspondiéndose con bajos niveles de expresión, mientras que el gen de UBO5 posee los valores intermedios de C_T y por tanto de expresión. Por su parte, eEF-1 α y UBO5 presentaron los menores rangos de expresión (5,42 y 5,83 respectivamente), mientras que el gen de actina y el de ARNr 18S presentaron rangos más elevados (7,01 y 12,49 respectivamente). El rango de expresión está definido por los valores máximo y mínimo de C_T , y es un indicador de la estabilidad de expresión.

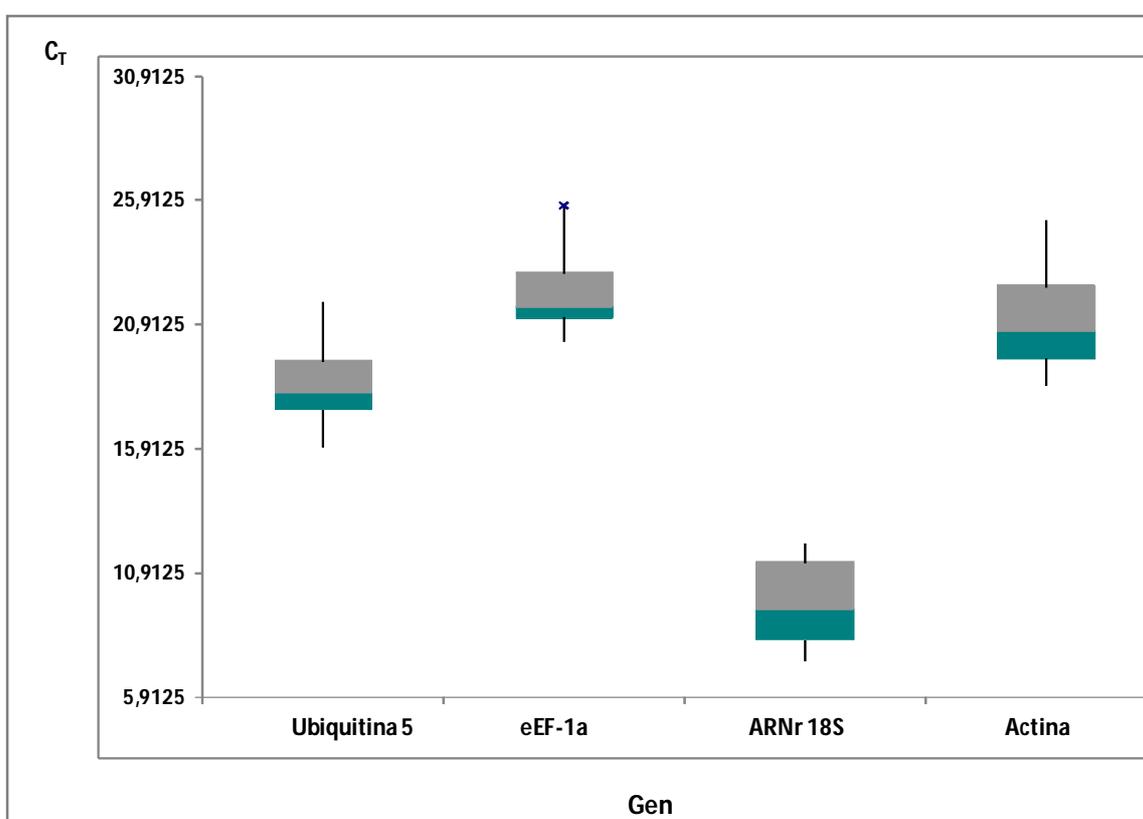


Figura 3: Valores absolutos de C_T para los 4 genes constitutivos analizados para todos los genotipos y tratamientos, expresados como medianas (línea entre ambos colores), percentiles 25 y 75 (cajas de colores), rangos (líneas) y valores *outliers* (cruces).

El análisis de varianza dio diferencias significativas entre 0 y 4 hs para el gen de ARNr 18S a un 95% de significancia (Tabla 3). Dicho gen presenta también el mayor valor de CV (23,84). El resto de los genes no mostraron diferencias significativas de expresión entre los dos tratamientos ni genotipos, presentando eEF-1 α el menor valor de CV (5,99).

Tabla 3: Resultados del análisis de varianza utilizado para analizar la estabilidad de la expresión de cuatro genes constitutivos entre los tres genotipos y tratamientos.

Gen	F	Valor de R	CV (%)
UBQ 5	0,73	0,6157	8,25
eEF-1 α	1,71	0,2062	5,99
Actina	0,5	0,7742	10,61
ARNr 18S	3,92	0,0243	23,84

El gen que codifica para el eEF-1 α , presentó el menor valor promedio y desvío de $2^{-\Delta CT}$. Por otro lado, el gen que codifica para el ARNr 18S fue el que presentó mayor valor de $2^{-\Delta CT}$ promedio y mayor desvío (Figura 4).

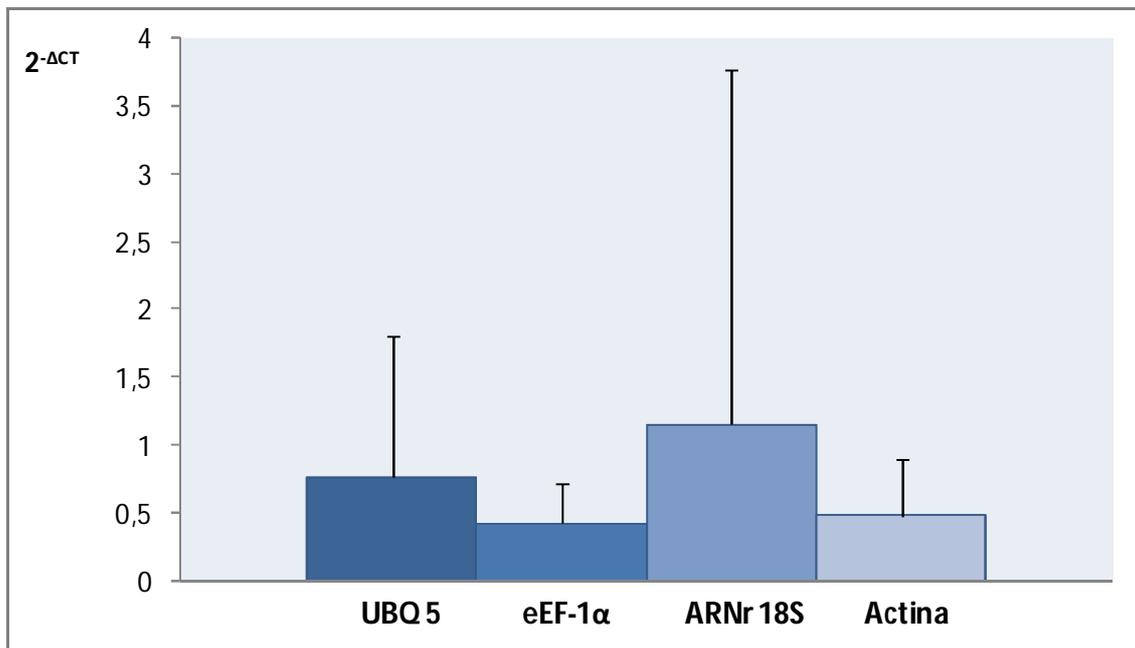


Figura 4: Niveles de expresión sin normalizar de los 4 genes constitutivos bajo tratamiento de 4 horas de frío, expresados como $2^{-\Delta CT}$ promedio y desvío estándar (barras).

En un trabajo anterior sobre análisis de expresión de genes de tolerancia a bajas temperaturas en arroz mediante PCR en tiempo real, se utilizó el gen de Actina como gen de referencia (Cheng et al., 2007). Varios ensayos de validación de genes constitutivos en arroz como en otras especies han establecido que este gen no es el más apropiado para este tipo de ensayo, dado que se ha probado que su expresión puede ser influida por diferentes condiciones experimentales (Jain et al., 2006b; Kim et al., 2003; Nicot, 2005; Ruan and Lai, 2007). En este estudio, el gen de Actina no demostró ser el más apropiado aunque sus niveles de expresión no resultaron ser los más variables entre los genes analizados (Figuras 3 y 4).

Al igual que en trabajos anteriores, el gen que codifica para el ARNr 18S fue el que presentó los mayores niveles de expresión en este estudio (Jain et al., 2006b; Kim et al., 2003). Sin embargo, a diferencia de resultados previos (Goidin, 2001; Kim et al., 2003) mostró ser el más variable (Fig. 3 y 4), presentando incluso diferencias significativas de expresión entre muestras controles y tratadas, como revela el análisis de varianza (Tabla 3).

Las ubiquitinas por su parte, han mostrado una gran estabilidad de expresión en otros trabajos tanto en arroz (Jain et al., 2006b) como en álamo (Brunner et al., 2004). En este trabajo se analizó la expresión de un único gen de ubiquitina (UBQ 5), el cual no mostró estabilidad en sus niveles de expresión entre muestras controles y tratadas.

De los cuatro genes de referencia analizados en este estudio solamente uno, el Factor de Elongación eucariótico 1 α (eEF-1 α), fue el gen cuya expresión resultó ser más estable entre los tres genotipos diferentes de arroz, y bajo condiciones experimentales específicas (4 horas, 10 °C). Esto concuerda con resultados previos en arroz, donde este gen mostró gran estabilidad de expresión durante diferentes etapas del desarrollo (Jain et al., 2006b). En base a esto se eligió al eEF-1 α para seguir con el resto de los estudios de expresión.

II. Cuantificación de la expresión génica

Tratamientos de frío y extracciones de ARN

Se realizaron extracciones de ARN para todos los genotipos sometidos a los 7 tratamientos (control, 0.5, 1, 2, 4, 8 y 24 hs). En la Figura 5 se muestra un ejemplo de una extracción realizada con el reactivo Trizol. Aquellas muestras cuyos valores de $A_{260/280}$ y $A_{260/230}$ no caían dentro de los rangos aceptables (ver Materiales y Métodos) fueron nuevamente extraídas. En general la calidad de las muestras fue buena y los valores de concentración fueron elevados (562-2547 ng/ μ L).

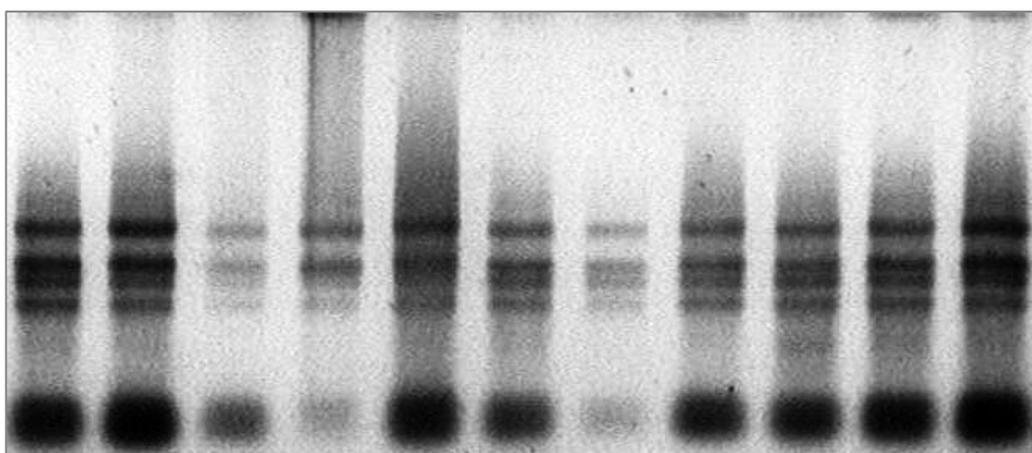


Figura 5: Ejemplo de extracción de ARN a partir de hojas de arroz.

Curvas de eficiencia

El número de ciclos requeridos para alcanzar un cierto nivel de fluorescencia no depende solamente de la cantidad de ADNc en la muestra, sino también de la eficiencia de reacción. Cuando la cantidad de producto de PCR se duplica en cada ciclo, la eficiencia de reacción es de 100%. En este caso, los cebadores fueron diseñados para amplificar fragmentos de 116-198 pb, con el objetivo de maximizar los valores de eficiencia. La secuencia de los cebadores utilizados se muestra en la Tabla 4.

Tabla 4: Secuencia de los cebadores utilizados para amplificar los 6 genes de interés mediante PCR cuantitativo.

Gen	Accesión	Descripción	Cebadores ^a	Largo del amplicón
ROSbZIP-1	LOC_Os08g43090	Factor de transcripción	5'-AGCTAAGAGATGCCCTGAATGAAG-3'* 5'-TCTGCCGATGACTGTTGTTGCTGTAT-3'*	168
AIP	LOC_Os01g13030	Proteína inducida por auxina	5'-CAAGAACACGATGGCAACCAACCA-3'* 5'-TTACCGGTGCTGAAGCCAATGAAC-3'*	198
Fer	LOC_Os11g01530	Ferritina	5'-CCCCCAGGCCAAGGACCAGT-3 ^D 5'-ACGATCAAAGTAGGCGAAAAGGGAGTG-3 ^D	133
PP2Ca	LOC_Os01g40094	Fosfatasa 2C	5'-GAAGTGGGAACAGGCCTTTG-3 ^M 5'-CAACGATGATGTGCGAGGAG-3 ^M	146
Hs70	LOC_Os11g47760	Heat shock protein 70	5'-GAACCAGGTCGCCATGAACC-3 ^D 5'-TCACCTGAAGGGCCAGAGC-3 ^D	116
AMPds	LOC_Os02g32490	Proteína de unión a AMP	5'-GCCTGTCCAGATCAGTTGGT-3'* 5'-CATTGCCCTCCAATACATC-3'*	116

a. Secuencia del Cebador Forward (arriba) y Reverse (abajo). Posición de los cebadores en el gen: (*) límite entre dos exones, (D) diferentes exones, (M) mismo exón.

Las curvas de eficiencia fueron representadas como los valores promedio de C_T entre las réplicas técnicas, en función del logaritmo de la concentración de ADNc (Figura 6A). Los valores de r^2 fueron superiores a 0,99, lo cual es un buen indicador de la sensibilidad del ensayo (Stratagene, 2007). Los valores de eficiencia de reacción fueron calculados a partir de la pendiente del gráfico, cuyos valores se hallan entre 3,2588 y 3,3186, los que se corresponden a su vez a eficiencias de entre 100 y 102%. Dada la semejanza de estos valores entre sí, y con el gen de referencia escogido, el eEF-1 α (Figura 2), se procedió al análisis de expresión mediante el método del $2^{-\Delta\Delta C_T}$.

Por otro lado, se realizaron análisis de curva de disociación de dichas reacciones, los cuales confirmaron la amplificación de un único producto de PCR (Figura 6B).

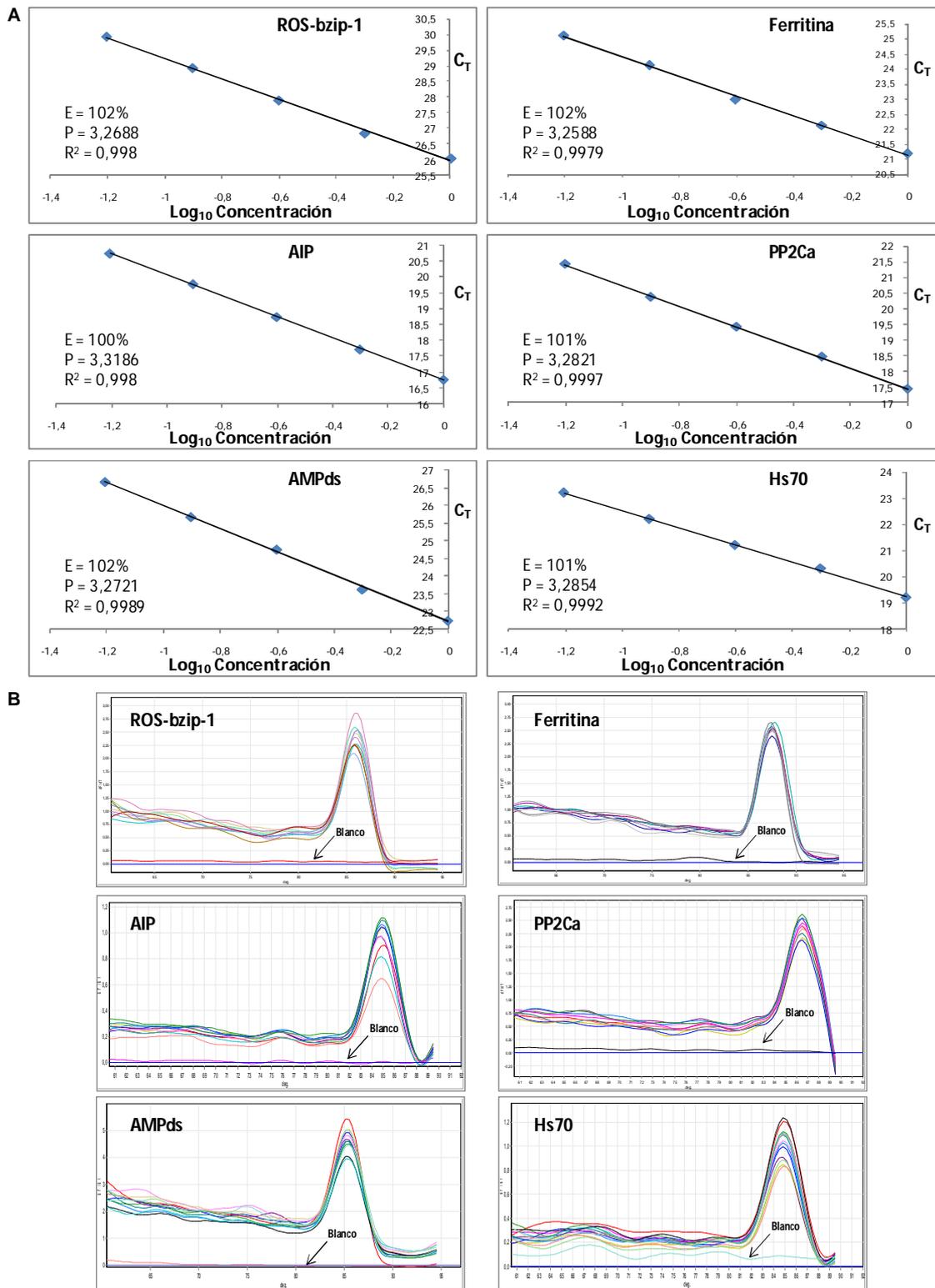


Figura 6: Eficiencia y especificidad de la amplificación por PCR en tiempo real. A) Curvas de eficiencia para los 6 genes de interés, la eficiencia (E) calculada a partir de la pendiente del gráfico (P). B) Curvas de disociación para los 6 genes calculadas a partir del programa RotorGene 6000.

Cuantificación de la expresión

En base los resultados previos, se realizó el análisis de expresión de los 6 genes de interés utilizando el gen de eEF-1 α como gen de referencia. Cada uno de los 6 genes mostró ser expresado diferencialmente en algún genotipo y/o tratamiento.

Tabla 5: Genes seleccionados, número de accesión, y posición del microsatélite dentro del gen

Gen	Accesión	Localización
ROS_bZIP1	LOC_Os08g43090	3' UTR
Auxin induced protein	LOC_Os01g13030	intrón
Ferritina	LOC_Os11g01530	intrón
Heat-shock protein 70	LOC_Os11g47760	intrón
AMP binding protein	LOC_Os02g32490	intrón
PP2Ca	LOC_Os01g40094	intrón

ROS-bZIP1

ROS-bZIP1 es un factor de transcripción perteneciente a la familia *bZIP* cuyos miembros cumplen roles importantes en la respuesta a ABA y en la regulación de la respuesta frente al estrés oxidativo y patógenos (Lu et al., 2009). Este gen ha sido identificado por primera vez mediante estudios transcriptómicos realizados sobre microarreglos de arroz enriquecidos con genes candidatos de respuesta a frío (Cheng et al., 2007; Yun et al., 2010). Estos trabajos revelaron que este gen es inducido tempranamente por bajas temperaturas y por H₂O₂, y que presenta un perfil de expresión diferencial entre genotipos tolerantes y susceptibles de arroz (Cheng et al., 2007).

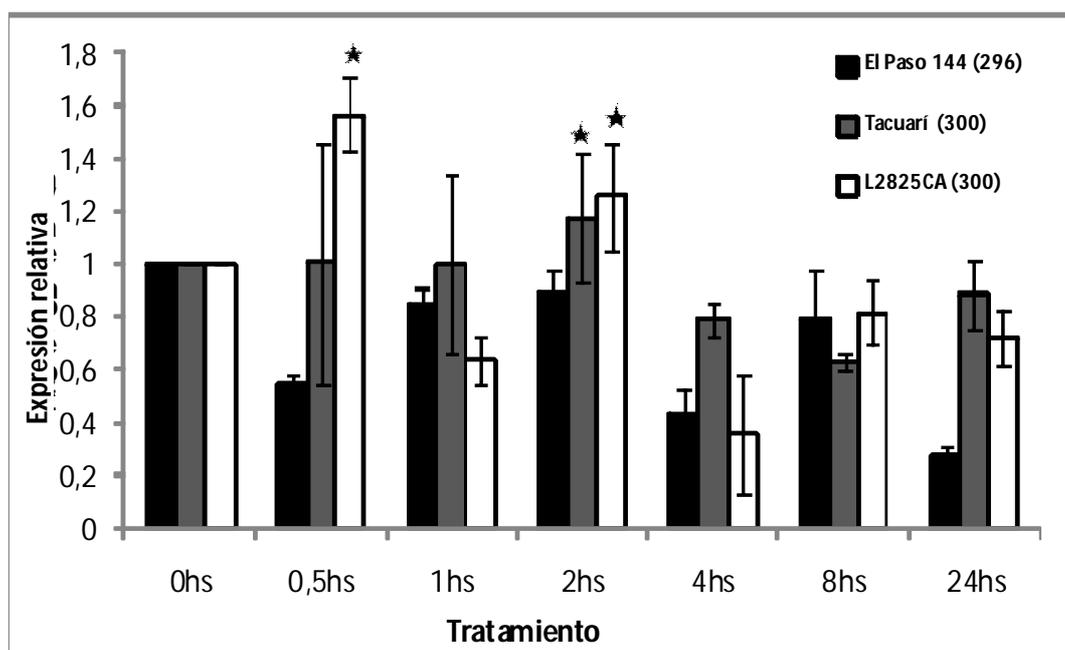


Figura 7: Perfil de expresión temporal del gen ROS-bZIP1 en tres genotipos de arroz uruguayo obtenido por PCR en Tiempo Real. En paréntesis se detalla el tamaño alélico del marcador génico para cada genotipo. Se destacan los valores significativos según test de comparación de Tukey (estrellas). Los niveles de transcrito fueron normalizados contra un gen de referencia eEF-1 α . Los valores de expresión relativa son el resultado del promedio de ΔC_T de réplicas biológicas.

En este caso, los niveles generales de expresión de este gen no fueron elevados, lo cual es esperable al tratarse de un factor de transcripción (Figura 7). Ambos genotipos tolerantes, L2825CA e INIA Tacuarí, fueron los únicos que presentaron diferencias significativas de expresión respecto a los controles. En el caso de L2825CA, se observa un pico de expresión a la media hora de tratamiento, y luego otro de inferior magnitud a las 2 horas de frío. Por su parte, INIA Tacuarí también reveló un pico de inducción a las 2 horas de tratamiento, pero de menor magnitud que el presentado por L2825CA (Figura 7). Estos resultados no concuerdan con lo publicado por Cheng et al. (2007), quienes encontraron que en la línea resistente a frío CT748-8-CA-17 los niveles de expresión de este gen permanecían significativamente superiores a los basales durante las 24 horas de exposición a 10 °C.

Sin embargo, Cheng et al (2007) plantean que este gen es inducido por H₂O₂. Por su parte, estudios de producción de H₂O₂ por INIA Tacuarí, L2825CA y El Paso 144 muestran que la aparición del mismo coincide con un aumento en los niveles de expresión de este factor de transcripción, pudiendo tener una relación causal (Bonnetcarrère et al. 2011 y Bonnetcarrère com. pers.). El genotipo El Paso 144 no presentó producción de H₂O₂ ni tampoco inducción de la expresión del gen ROS-bzip1, mientras que en los genotipos tolerantes existió un aumento tanto en la inducción de

H₂O₂ como en la expresión del gen (Bonnecarrère et al. 2011 y com. pers. Bonnecarrère, 2011).

El transcripto de este gen presenta un microsatélite de tipo di-nucleótido en su región 3'UTR (Tabla 5). Los repetidos en esta región podrían afectar la expresión del gen que la porta provocando un deslizamiento de la ARN polimerasa dando lugar así a la transcripción de un ARNm más largo que lo esperado, como ya fue estudiado en levaduras (Fabre et al., 2002). En humanos, se ha visto que estos transcriptos más largos tienden a acumularse formando focos en el núcleo, interrumpiendo de este modo procesos celulares como el corte y empalme, como en el caso de la distrofia muscular (Mankodi et al., 2002). En este caso, ambos genotipos tolerantes, INIA Tacuarí y L2825CA comparten el mismo alelo de microsatélite, y como se mencionó previamente, ambos muestran niveles significativos de expresión de ROS-bzip1 aunque INIA Tacuarí lo hace en menor magnitud (Figura 7). Por tanto, no es viable establecer una posible influencia entre el tamaño de microsatélite y una posible inhibición de la expresión génica.

Ferritina

Las ferritinas vegetales se localizan en cloroplastos y mitocondrias, y se organizan en familias génicas pequeñas cuyos miembros son regulados diferencialmente (Stein et al., 2009). Las ferritinas intervienen en varios procesos fisiológicos durante el desarrollo de la planta, gracias a su capacidad de secuestrar y almacenar hierro, el cual es útil para la síntesis de otras proteínas (Briat et al., 1999). Su expresión también es inducida bajo condiciones de estrés oxidativo, debido a su capacidad de atenuar la reacción de Fenton, secuestrando los iones de hierro e impidiendo así la formación de especies reactivas del oxígeno (Hegedüs et al., 2002).

Los niveles de expresión globales de ferritina fueron mucho más elevados, lo cual concuerda con los resultados presentados por Cheng et al. (2007). Nuevamente, L2825CA e INIA Tacuarí presentaron los niveles más elevados de expresión, presentando diferencias significativamente superiores a los niveles basales a las 4 y 8 horas de tratamiento a 10 °C en el caso de INIA Tacuarí, y a las 8 horas en el caso de L2825CA. En ambos casos esta inducción decae a las 24 horas (Figura 8). Temporalmente se observa una inducción de la expresión de este gen posterior a la inducción del gen de ROS-bZIP1, lo cual coincide con el supuesto de que el producto de este gen actúa como activador de la expresión de un regulón de genes de respuesta a frío, del cual la ferritina forma parte (Cheng et al., 2007).

Este gen presenta un microsatélite localizado en un intrón (Tabla 5). Los microsatélites localizados en estas regiones pueden afectar el proceso de corte y

empalme dando lugar a nuevas variantes proteicas, las cuales pueden o no ser funcionales (Li, 2004). En este caso, INIA Tacuarí y El Paso 144 comparten un mismo alelo largo, sin embargo ambos muestran perfiles de expresión muy disímiles (Figura 8), no pudiéndose establecer una relación directa entre número de repetidos y expresión.

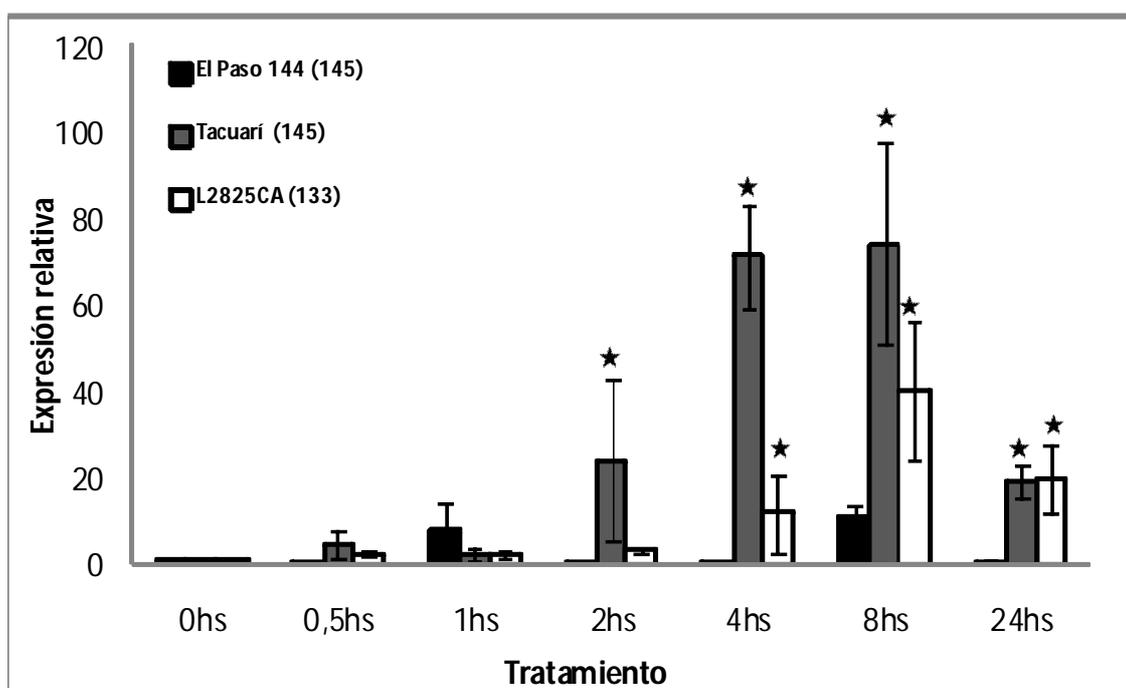


Figura 8: Perfil de expresión temporal del gen de ferritina para tres genotipos de arroz uruguayo obtenido por PCR en Tiempo Real. En paréntesis se detalla el tamaño alélico del marcador génico para cada genotipo. Se destacan los valores significativos según test de comparación de Tukey (estrellas). Los niveles de transcripto fueron normalizados contra un gen de referencia eEF-1 α . Los valores de expresión relativa son el resultado del promedio de ΔC_T de réplicas biológicas.

AIP

La proteína inducida por auxina (AIP) pertenece a la familia de genes de respuesta a auxina Aux/IAA (Jain et al., 2006a). Estudios transcriptómicos han revelado que varios de los componentes de esta familia son inducidos por bajas temperaturas entre otros tipos de estreses abióticos (Cheng et al., 2007; Jain and Khurana, 2009).

Los niveles generales de expresión de este gen fueron menores que los presentados por el gen de la ferritina (Figura 9). Los genotipos INIA Tacuarí y L2825CA fueron quienes presentaron los mayores niveles de inducción mostrando un aumento significativo a las 4 horas, lo cual no concuerda con trabajos previos donde la inducción

de este gen tiene lugar a la media hora luego del tratamiento con bajas temperaturas en genotipos tolerantes (Cheng et al., 2007).

Ambos genotipos tolerantes mostraron niveles significativos de expresión además de compartir el mismo alelo de microsatélite (Figura 9). Este microsatélite es un trinucleótido y se encuentra en un intrón (Tabla 5) al igual que en el caso del gen de ferritina. En este caso, al igual que con ROS-bZIP1, nuevamente se cumple que ambos genotipos que presentan un aumento en la expresión del gen comparten un mismo alelo de microsatélite.

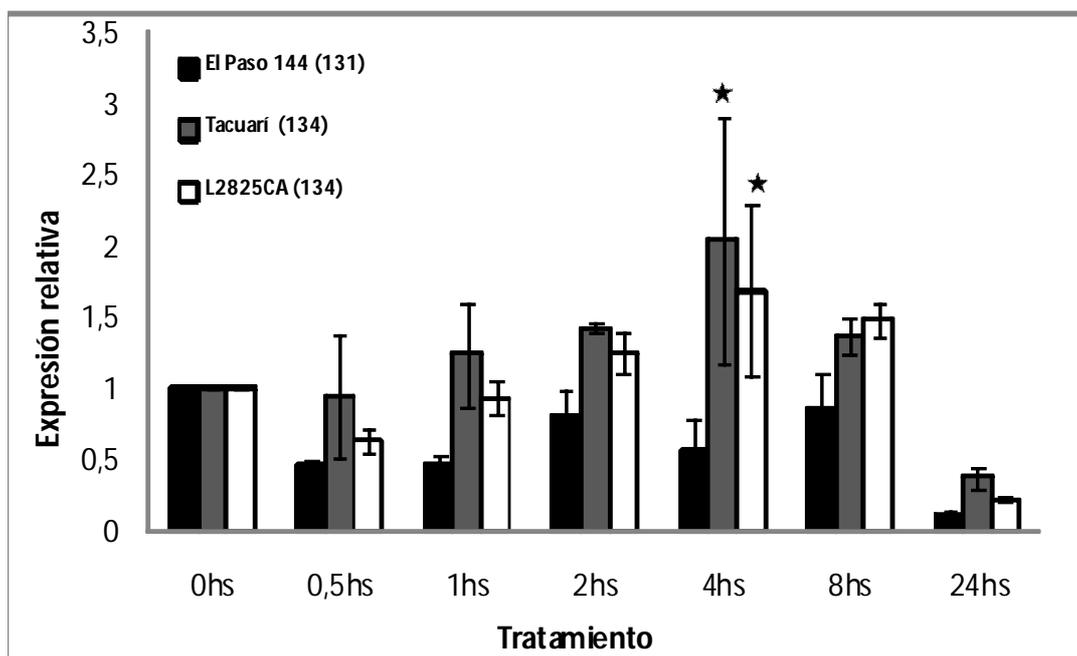


Figura 9: Perfil de expresión temporal del gen de la proteína inducida por auxina (*AIP*) para tres genotipos de arroz uruguayo obtenidos por PCR en Tiempo Real. En paréntesis se detalla el tamaño alélico del marcador génico para cada genotipo. Se destacan los valores significativos según test de comparación de Tukey (estrellas). Los niveles de transcritpo fueron normalizados contra un gen de referencia *eEF-1 α* . Los valores de expresión relativa son el resultado del promedio de ΔC_T de réplicas biológicas.

PP2Ca

La fosfatasa de proteína 2Ca (*PP2Ca*) pertenece a la familia de las fosfatasas de proteínas 2C que actúan revirtiendo la acción de las quinasas modulando así la regulación por fosforilación de las proteínas, que ocurren en la transducción de señales desencadenada por los estreses ambientales (Xue et al., 2008). La inducción de la expresión de este gen por bajas temperaturas fue previamente estudiada en plantas transgénicas capaces de expresar constitutivamente los genes *CBF3* y *ABF3* de *Arabidopsis* (Oh, 2005), y también por estudios transcriptómicos (Rabbani et al., 2003).

A diferencia del resto de los genes analizados, PP2Ca no mostró diferencias significativas de expresión en los genotipos tolerantes, sino que por el contrario, el único genotipo que mostró inducir su expresión fue El Paso 144 a las 4 horas de tratamiento (Figura 10). Este gen posee un microsatélite localizado en un intrón (Tabla 5), del cual Tacuarí y L2825CA comparten el mismo alelo. Nuevamente, se cumple que los genotipos que comparten el mismo alelo, comparten el mismo perfil de expresión del gen en cuestión.

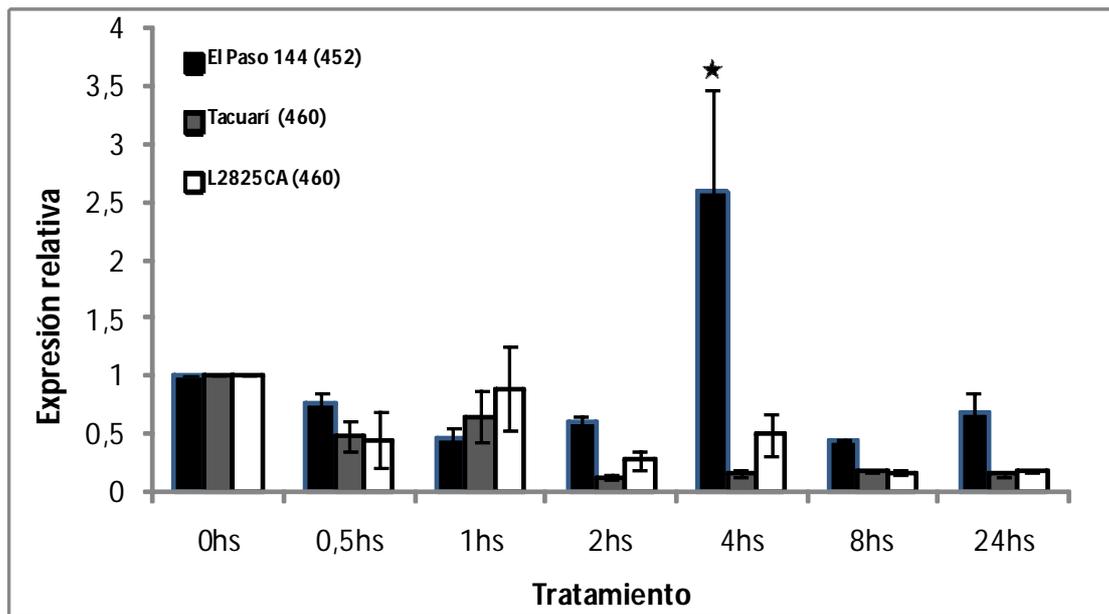


Figura 10: Perfil de expresión temporal del gen fosfatasa de proteína 2Ca (PP2Ca) para tres genotipos de arroz uruguayo obtenido por PCR en Tiempo Real. En paréntesis se detalla el tamaño alélico del marcador génico para cada genotipo. Se destacan los valores significativos según test de comparación de Tukey (estrellas). Los niveles de transcrito fueron normalizados contra un gen de referencia eEF-1 α . Los valores de expresión relativa son el resultado del promedio de ΔC_T de réplicas biológicas.

AMPds

AMPds es una Acetil CoA sintetasa (ACS), las cuales se ha probado tienen el rol de sintetizar y modificar lípidos (Shockey et al., 2000). Recientemente se ha probado que estas enzimas participan en procesos de defensa a estrés biótico (Turner et al., 2005; Zhang et al., 2009), pero hasta ahora no existen datos sobre su posible implicancia en procesos de respuesta a estrés abiótico. Este gen fue elegido no por revisiones bibliográficas sino a través de búsquedas bioinformáticas sobre bancos de ESTs de respuesta a frío en arroz.

En este caso, se observa un incremento tardío en la expresión, presentando El Paso 144 y L2825CA niveles significativos de expresión a las 24 horas de tratamiento (Figura 11). Ambos genotipos comparten el mismo alelo del microsatélite, el cual se encuentra localizado en un intrón de dicho gen (Tabla 5).

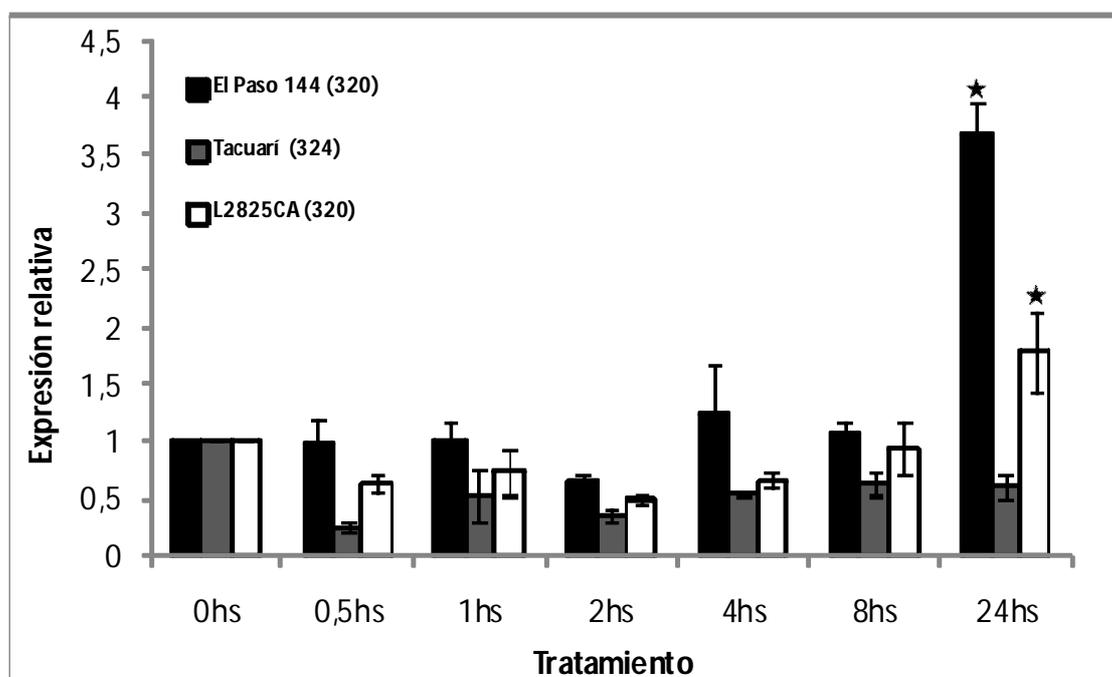


Figura 11: Perfil de expresión temporal del gen de unión a AMP (AMPd) para tres genotipos de arroz uruguayo obtenido por PCR en Tiempo Real. En paréntesis se detalla el tamaño alélico del marcador génico para cada genotipo. Se destacan los valores significativos según test de comparación de Tukey (estrellas). Los niveles de transcritos fueron normalizados contra un gen de referencia eEF-1 α . Los valores de expresión relativa son el resultado del promedio de ΔC_T de réplicas biológicas.

Hs70

Este gen codifica para un tipo de chaperona, la cual se activa bajo condiciones de estrés, encargándose de prevenir el agregado entre proteínas y asistiendo el correcto plegamiento de las proteínas mal plegadas (Bukau and Horwich, 1998). Varios estudios han revelado que este tipo de proteínas son expresadas bajo condiciones de estrés abiótico incluyendo bajas temperaturas (Wang et al., 2003; Wang et al., 2004).

Los niveles de expresión significativamente elevados se presentan en los tres genotipos pero a diferentes horas de tratamiento (Figura 12). Mientras que El Paso 144 presenta un pico de expresión a las 4 horas, el cual luego disminuye, Tacuarí y L2825CA, lo hacen más tardíamente, a las 24 horas. Los tres genotipos presentan diferentes alelos (Tabla 5).

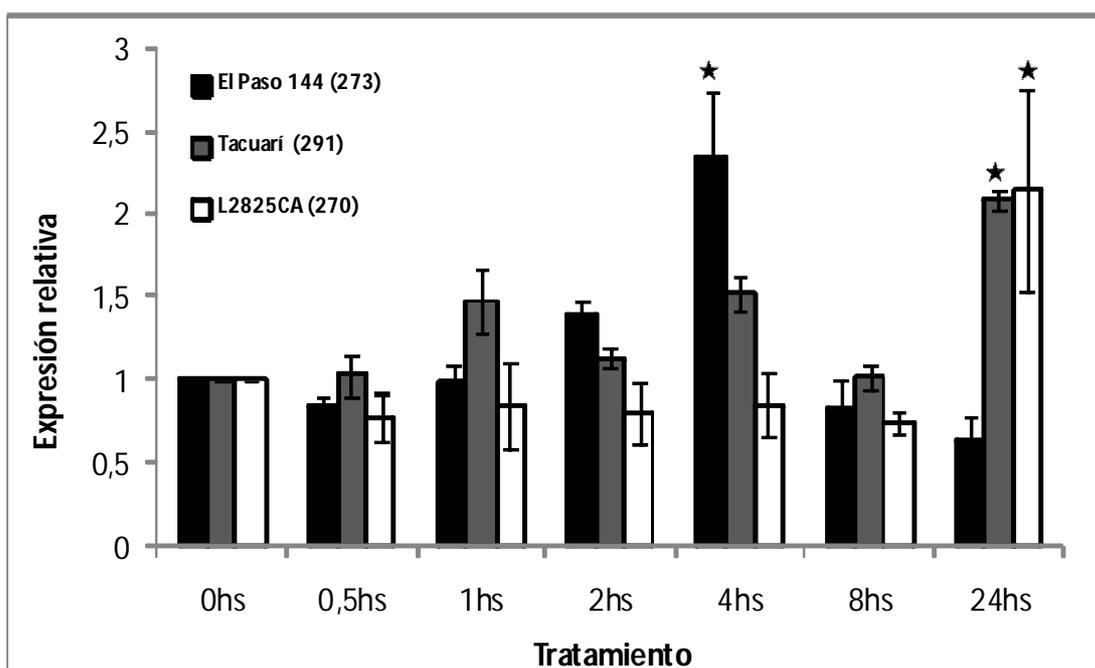


Figura 12: Perfil de expresión temporal del gen Hs70 para tres genotipos de arroz uruguayo obtenido por PCR en Tiempo Real. En paréntesis se detalla el tamaño alélico del marcador génico para cada genotipo. Se destacan los valores significativos según test de comparación de Tukey (estrellas). Los niveles de transcritos fueron normalizados contra un gen de referencia eEF-1 α . Los valores de expresión relativa son el resultado del promedio de ΔC_T de réplicas biológicas.

Hasta la fecha, pocos estudios han analizado la expresión de genes que contengan microsatélites en su secuencia en vegetales. En un estudio en pimienta, se ha relacionado la presencia de microsatélites en varios genes con perfiles diferenciales de expresión entre diferentes tejidos y etapas del desarrollo (Ince et al., 2010). En arroz, si bien existen estudios sobre asociación de alelos de microsatélites génicos con ciertas características fenotípicas, como es el caso del gen *waxy* (Ayres et al., 1997; Bao et al., 2002), aún no se han realizado análisis vinculando la presencia de un microsatélite génico con la expresión del gen en cuestión.

En este trabajo se realizó un análisis de expresión por PCR en tiempo real sobre 6 genes de tolerancia a frío en arroz, los cuales contienen microsatélites polimórficos en su secuencia. Este análisis de expresión se realizó sobre tres genotipos de arroz, los cuales muestran una respuesta diferencial a las bajas temperaturas. El cultivar El Paso 144 constituye la variedad más cultivada en Uruguay y pertenece al ecotipo *indica*, y por tanto susceptible a las bajas temperaturas. El cultivar INIA Tacuarí y el genotipo L2825CA pertenecen al ecotipo *japonica*, y ambos si bien han demostrado tolerancia a

las bajas temperaturas, sus mecanismos de respuesta involucran diferentes estrategias. Mientras que INIA Tacuarí presenta un mecanismo de disipación de energía a través del ciclo de las xantofilas, L2825CA posee niveles basales de actividad enzimática antioxidante elevados (Bonnecarrère et al., 2011). En este trabajo, ambos genotipos tolerantes mostraron patrones similares de inducción de la expresión para los cuatro genes seleccionados (Hs70, ROS-bzip1, ferritina y AIP) a partir de los resultados obtenidos en el Capítulo 1 con el análisis discriminante. Igualmente, estos resultados no son suficientes para afirmar una asociación entre el nivel de expresión de estos genes y la tolerancia a frío.

Todos los genes mostraron variaciones significativas en los niveles de expresión para alguno de los 3 genotipos en el transcurso del tratamiento a 10 °C (Figuras 7-12). Concordando con resultados previos, ROS-bzip1, ferritina y AIP mostraron niveles elevados de expresión en alguno de ambos genotipos tolerantes (Cheng et al., 2007). Entre los 6 genes analizados en este estudio, se cumple que para 4 (ROS-bzip1, AIP, PP2Ca y AMPds) existe una coincidencia entre el perfil de expresión y el tamaño alélico compartido entre genotipos. Sin embargo, en algunos de estos casos o bien los niveles de expresión entre los genotipos sobreexpresantes no son similares entre sí, o bien no existe una coincidencia en los tiempos de inducción. Dado que los genes analizados no fueron secuenciados para ninguno de los genotipos, no es posible determinar que la única diferencia en la secuencia de estos genes entre genotipos está determinada por el SSR, o si existen otros tipos de mutaciones (por ejemplo en secuencias reguladoras), capaces de ser responsables de las diferencias de expresión.

En este capítulo se analizó la expresión de 4 genes constitutivos utilizados previamente como referencia para análisis de expresión en tiempo real en arroz, frente a condiciones experimentales específicas. De estos 4 genes pudo validarse uno, el eEF-1 α , cuya expresión permaneció constante entre los tres genotipos analizados durante las diferentes condiciones experimentales impuestas. Por otro lado, se realizó un estudio exploratorio con el objetivo de buscar una relación entre la expresión de 6 genes identificados previamente mediante revisiones bibliográficas y estudios de asociación por análisis discriminante, como vinculados a la tolerancia a bajas temperaturas en arroz durante etapas tempranas del desarrollo los cuales a su vez contenían secuencias de microsatélites asociadas. Los niveles de expresión de estos genes fueron analizados entre tres genotipos de arroz uruguayo contrastantes en cuanto a su respuesta a las bajas temperaturas portadores de alelos diferentes de dichos microsatélites génicos. Para 4 de los 6 genes analizados se encontró una congruencia entre el alelo identificado y su nivel de expresión en distintos genotipos, lo cual constituye una primera aproximación directa al estudio sobre el efecto de la variación de los microsatélites génicos en la expresión. La secuenciación completa de los alelos analizados, así como la realización de estos mismos análisis con un mayor

número de genes permitirá investigar el efecto de los microsatélites génicos sobre la expresión.

Bibliografía

- Ayres, N.M., McClung, A.M., Larkin, P.D., Bligh, H.F.J., Jones, C.A., and Park, W.D. (1997). Microsatellites and a single-nucleotide polymorphism differentiate apparent amylose classes in an extended pedigree of US rice germ plasm. *Theoretical Applied Genetics* *94*, 773-781.
- Bao, J., Corke, H., and Sun, M. (2002). Microsatellites in starch-synthesizing genes in relation to starch physicochemical properties in waxy rice (*Oryza sativa* L.). *TAG Theoretical and Applied Genetics* *105*, 898-905.
- Belkum, A.V., Scherer, S., Alphen, L.V., and Verbrugh, H. (1998). Short-Sequence DNA Repeats in Prokaryotic Genomes. *Microbiology and Molecular Biology Reviews* *62*, 275-293.
- Bonnecarrère, V., Borsani, O., Díaz, P., Capdevielle, F., Blanco, P., and Monza, J. (2011). Response to photooxidative stress induced by cold in japonica rice is genotype dependent. *Plant Science* *180*, 726-732.
- Briat, J.-F., Lobréaux, S., Grignon, N., and Vansuyt, G. (1999). Regulation of plant ferritin synthesis: how and why. *Cellular and molecular life sciences* *56*, 155-166.
- Brunner, A.M., Yakovlev, I.A., and Strauss, S.H. (2004). Validating internal controls for quantitative plant gene expression studies. *BMC Plant Biology* *4*, 14.
- Bukau, B., and Horwich, A.L. (1998). The Hsp70 and Hsp60 Chaperone Machines. *Cell* *92*, 351-366.
- Bustin, S.A. (2002). Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems. *Journal of Molecular Endocrinology* *29*, 23-39.
- Čikoš, Š., Bukovská, A., and Koppel, J. (2007). Relative quantification of mRNA: comparison of methods currently used for real-time PCR data analysis. *BMC Molecular Biology* *8*, 113.
- Cummings, C.J., and Zoghbi, H.Y. (2000). Fourteen and counting: unraveling trinucleotide repeat diseases. *Human Molecular Genetics* *9*, 909-916.
- Cheng, C., Yun, K.-Y., Ransom, H.W., Mohanty, B., Bajic, V.B., Jia, Y., Yun, S., and de los Reyes, B.G. (2007). An early response regulatory cluster induced by low temperature and hydrogen peroxide in seedlings of chilling-tolerant japonica rice. *BMC Genomics* *8*, 175.
- Ellegren, H. (2004). Microsatellites: simple sequences with complex evolution. *Nature Reviews Genetics* *5*, 435-445.
- Fabre, E., Dujon, B., and Richard, G.-F. (2002). Transcription and nuclear transport of CAG/CTG trinucleotide repeats in yeast. *Nucleic Acids Research* *30*, 3540-3547.
- Fleige, S., and Pfaffl, M.W. (2006). RNA integrity and the effect on the real-time qRT-PCR performance. *Molecular Aspects of Medicine* *27*, 126-139.
- Gatchel, J.R., and Zoghbi, H.Y. (2005). Diseases of Unstable Repeat Expansion: Mechanisms and Common Principles. *Nature Reviews Genetics* *6*, 743-755.
- Goidin, D. (2001). Ribosomal 18S RNA Prevails over Glyceraldehyde-3-Phosphate Dehydrogenase and β -Actin Genes as Internal Standard for Quantitative Comparison of mRNA Levels in Invasive and Noninvasive Human Melanoma Cell Subpopulations. *Analytical Biochemistry* *295*, 17-21.

- Hegedűs, A., Erdei, S., Janda, T., Szalai, J., Dudits, D., and Horváth, G. (2002). Effects of low temperature stress on ferritin or aldose reductase overexpressing transgenic tobacco plants. *Proceedings of the 7th Hungarian Congress on Plant Physiology* 46, 97-98.
- Huggett, J., Dheda, K., Bustin, S., and Zumla, A. (2005). Real-time RT-PCR normalisation; strategies and considerations. *Genes and Immunity* 6, 279-284.
- Imbeaud, S., Graudens, E., Boulanger, V., Barlet, X., Zaborski, P., Eveno, E., Mueller, O., Schroeder, A., and Auffray, C. (2005). Towards standardization of RNA quality assessment using user-independent classifiers of microcapillary electrophoresis traces. *Nucleic Acids Research* 33, e56-e56.
- Ince, A.G., Karaca, M., and Naci Onus, A. (2010). Differential expression patterns of genes containing microsatellites in *Capsicum annum* L. *Molecular Breeding* 25, 645-658.
- Jain, M., Kaur, N., Garg, R., Thakur, J.K., Tyagi, A.K., and Khurana, J.P. (2006a). Structure and expression analysis of early auxin-responsive Aux/IAA gene family in rice (*Oryza sativa*). *Functional & Integrative Genomics* 6, 47-59.
- Jain, M., and Khurana, J.P. (2009). Transcript profiling reveals diverse roles of auxin-responsive genes during reproductive development and abiotic stress in rice. *FEBS Journal* 276, 3148-3162.
- Jain, M., Nijhawan, A., Tyagi, A.K., and Khurana, J.P. (2006b). Validation of housekeeping genes as internal control for studying gene expression in rice by quantitative real-time PCR. *Biochemical and Biophysical Research Communications* 345, 646-651.
- Jarne, P., and Lagoda, P.J.L. (1996). Microsatellites from molecules to populations and back. *Tree* 11, 424-429.
- Kashi, Y., and King, D. (2006). Simple sequence repeats as advantageous mutators in evolution. *Trends in Genetics* 22, 253-259.
- Kim, B.-R., Nam, H.-Y., Kim, S.-U., Kim, S.-I., and Chang, Y.-J. (2003). Normalization of reverse transcription quantitative-PCR with housekeeping genes in rice. *Biotechnology Letters* 25, 1869-1872.
- Li, Y.C. (2004). Microsatellites Within Genes: Structure, Function, and Evolution. *Molecular Biology and Evolution* 21, 991-1007.
- Livak, K.J., and Schmittgen, T.D. (2001). Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2- $\Delta\Delta$ CT Method. *Methods* 25, 402-408.
- Lu, G., Gao, C., Zheng, X., and Han, B. (2009). Identification of OsZIP72 as a positive regulator of ABA response and drought tolerance in rice. *Planta* 229, 605-615.
- Mankodi, A., Takahashi, M.P., Jiang, H., Beck, C.L., Bowers, W.J., Moxley, R.T., Cannon, S.C., and Thornton, C.A. (2002). Expanded CUG Repeats Trigger Aberrant Splicing of ClC-1 Chloride Channel Pre-mRNA and Hyperexcitability of Skeletal Muscle in Myotonic Dystrophy. *Molecular Cell* 10, 35-44.
- Morgante, M., Hanafey, M., and Powell, W. (2002). Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. *Nature Genetics* 30, 194-200.
- Nicot, N. (2005). Housekeeping gene selection for real-time RT-PCR normalization in potato during biotic and abiotic stress. *Journal of Experimental Botany* 56, 2907-2914.
- Oh, S.J. (2005). Arabidopsis CBF3/DREB1A and ABF3 in Transgenic Rice Increased Tolerance to Abiotic Stress without Stunting Growth. *Plant Physiology* 138, 341-351.

- Pfaffl, M.W. (2004). Quantification strategies in real-time PCR. In A-Z of quantitative PCR, S.A. Bustin, ed. (La Joya, CA, International University Line), pp. 87-112.
- Qiagen. (2009). QuantiTect_Reverse_Transcription_Handbook.
- Rabbani, M.A., Maruyama, K., Abe, H., Khan, M.A., Katsura, K., Ito, Y., Yoshiwara, K., Seki, M., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2003). Monitoring Expression Profiles of Rice Genes under Cold, Drought, and High-Salinity Stresses and Abscisic Acid Application Using cDNA Microarray and RNA Gel-Blot Analyses. *Plant Physiology* 133, 1755-1767.
- Radonić, A., Thulke, S., Mackay, I.M., Landt, O., Siebert, W., and Nitsche, A. (2004). Guideline to reference gene selection for quantitative real-time PCR. *Biochemical and Biophysical Research Communications* 313, 856-862.
- Robinson, T.L., Sutherland, I.A., and Sutherland, J. (2007). Validation of candidate bovine reference genes for use with real-time PCR. *Veterinary Immunology and Immunopathology* 115, 160-165.
- Ruan, W., and Lai, M. (2007). Actin, a reliable marker of internal control? *Clinica Chimica Acta* 385, 1-5.
- Schmittgen, T.D., Zakrajsek, B.A., Mills, A.G., Gorn, V., Singer, M.J., and Reed, M.W. (2000). Quantitative Reverse Transcription–Polymerase Chain Reaction to Study mRNA Decay: Comparison of Endpoint and Real-Time Methods. *Analytical Biochemistry* 285, 194-204.
- Schroeder, A., Mueller, O., Stocker, S., Salowsky, R., Leiber, M., Gassmann, M., Lightfoot, S., Menzel, W., Granzow, M., and Ragg, T. (2006). The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC Molecular Biology* 7, 3.
- Shockey, J., Schnurr, J., and Browse, J. (2000). Characterization of the AMP-binding protein gene family in *Arabidopsis thaliana*: will the real acyl-CoA synthetases please stand up? *Biochemical Society Transactions* 28, 955-957.
- Stein, R.J., Ricachenevsky, F.K., and Fett, J.P. (2009). Differential regulation of the two rice ferritin genes (OsFER1 and OsFER2). *Plant Science* 177, 563-569.
- Stratagene (2007). Introduction to Quantitative PCR. In *Methods and applications guide* (Stratagene).
- Turner, J., Greville, K., Murphy, E., and Hooks, M. (2005). Characterization of *Arabidopsis* fluoroacetate-resistant mutants reveals the principal mechanism of acetate activation for entry into the glyoxylate cycle. *Journal of Biological Chemistry* 280, 2780-2787.
- Wang, W., Vinocur, B., and Altman, A. (2003). Plant responses to drought, salinity and extreme temperatures: towards genetic engineering for stress tolerance. *Planta* 218, 1-14.
- Wang, W., Vinocur, B., Shoseyov, O., and Altman, A. (2004). Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. *Trends in Plant Science* 9, 244-252.
- Xue, T., Wang, D., Zhang, S., Ehling, J., Ni, F., Jakab, S., Zheng, C., and Zhong, Y. (2008). Genome-wide and expression analysis of protein phosphatase 2C in rice and *Arabidopsis*. *BMC Genomics* 9, 550.
- Yun, K.-Y., Park, M.R., Mohanty, B., Herath, V., Xu, F., Mauleon, R., Wijaya, E., Bajic, V.B., Bruskiwich, R., and de los Reyes, B.G. (2010). Transcriptional regulatory network triggered by oxidative signals configures the early response mechanisms of japonica rice to chilling stress. *BMC Plant Biology* 10, 16.

Zhang, X.-c., Yu, X., Zhang, H.-j., and Song, F.-m. (2009). Molecular characterization of a defense-related AMP-binding protein gene, OsBIABP1, from rice. *Journal of Zhejiang University SCIENCE B* 10, 731-739.

Conclusiones finales

Las bajas temperaturas constituyen uno de los mayores problemas en el cultivo de arroz en Uruguay. Por este motivo, es importante el mejoramiento del germoplasma local, especialmente mediante el desarrollo de nuevas variedades resistentes.

La caracterización de la diversidad genética entre líneas de mejoramiento es crucial para el correcto aprovechamiento de los recursos genéticos en un programa de mejoramiento. En este sentido, los marcadores moleculares son de gran utilidad. Debido a su completo ligamiento a los genes de interés, el uso de microsatélites génicos en la caracterización de la diversidad y construcción de mapas genéticos, constituye una gran ventaja.

Por otro lado, la presencia de microsatélites en regiones transcritas sugiere un posible rol ya sea en la regulación de la expresión génica, o en la función de dicho gen. El efecto de los microsatélites ha sido bastamente estudiado en humanos debido a su relación con enfermedades (Cummings and Zoghbi, 2000; Gatchel and Zoghbi, 2005). Sin embargo en vegetales, salvo unos pocos ejemplos (Ayres et al., 1997; Ince et al., 2010) dicho efecto ha sido poco estudiado, y esos pocos trabajos se han centralizado en caracteres simples, controlados por uno o pocos genes.

En la primer parte de este trabajo se intentó identificar marcadores funcionales (SSR génicos) asociados a respuesta a bajas temperaturas en plántulas de arroz. Se identificaron 36 marcadores de los cuales 16 resultaron polimórficos en 131 líneas de arroz *indica* y *japonica*. Esto último revela que la estrategia de búsqueda de microsatélites génicos utilizada en el primer capítulo fue efectiva. La mayoría de estos mostraron una clara distribución diferencial de las frecuencias alélicas entre genotipos *indica* y *japonica*. Los estudios de clasificación mediante análisis discriminante resultaron en un alto porcentaje de clasificación entre clases fenotípicas cuando se consideraba la totalidad de los marcadores. Sin embargo, cuando este análisis se realizaba utilizando aquellos marcadores con un mayor poder discriminante, los porcentajes de clasificación resultaban considerablemente menores. Esto puede ser explicado en parte por la baja cantidad de marcadores polimórficos utilizados (solamente 16).

En el Capítulo 2 se realizó el análisis de expresión para 6 de los 16 genes cuyos SSR asociados resultaron polimórficos. Cuatro de estos genes mostraron alto poder discriminante en el primer capítulo (ROS-bzip1, AIP, Hs70 y ferritina), y perfiles similares de expresión en ambos genotipos resistentes. Esto último valida de cierta

forma los resultados obtenidos para estos marcadores a partir del análisis discriminante.

Para 4 de los 6 genes analizados mediante PCR en tiempo real analizados en la segunda parte de este trabajo (ROS-bzip1, AIP, PP2Ca y AMPds), se encontró una congruencia entre el perfil de expresión y el alelo compartido entre genotipos. De todos modos, no es posible aún afirmar una relación causal entre presencia de alelo y nivel de expresión. Esto se debe principalmente a que si bien se encontró que los genotipos presentaban el mismo alelo, los genes analizados no fueron completamente secuenciados en esta etapa, por lo que no se puede afirmar que la única diferencia de secuencia entre genotipos esté constituida por el tamaño del microsatélite. Por otro lado, estos análisis fueron realizados entre genotipos diferentes, y por tanto, otros factores genéticos podrían estar afectando la expresión.

Una prueba directa del efecto de la variación del número de repetidos sobre la expresión génica sería la comparación entre genotipos idénticos que varíen únicamente en el número de repetidos del microsatélite génico a estudiar. Un posible modelo sería la utilización de líneas recombinantes isogénicas. Otro mecanismo posible sería la generación de nuevos mutantes a través de métodos transgénicos. Sin embargo, ambos métodos tienen la desventaja de requerir una gran inversión de tiempo y recursos económicos.

Los resultados obtenidos en este trabajo permitirán direccionar futuros abordajes donde mediante análisis de polimorfismos de un mayor número de marcadores y la aplicación de técnicas de asociación más robustas, como es el mapeo asociativo, en un mayor número de genotipos, puedan identificarse un mayor número de marcadores asociados a frío en arroz. Por otro lado, el hecho de que 5 de los 6 genes analizados mostraran patrones similares de expresión entre genotipos tolerantes a frío, los convierte en un conjunto de genes candidatos interesante para futuros estudios de asociación.

Bibliografía General

- <AmEconRev 62 1-2 177-186 Cleaver 1972.pdf>.
<Propuesta de Doctorado_Victoria Bonnecarrere.pdf>.
- Andersen, J.R., and Lübberstedt, T. (2003). Functional markers in plants. *Trends in Plant Science* 8, 554-560.
- Ayres, N.M., McClung, A.M., Larkin, P.D., Bligh, H.F.J., Jones, C.A., and Park, W.D. (1997). Microsatellites and a single-nucleotide polymorphism differentiate apparent amylose classes in an extended pedigree of US rice germ plasm. *Theoretical Applied Genetics* 94, 773-781.
- Bao, J., Corke, H., and Sun, M. (2002). Microsatellites in starch-synthesizing genes in relation to starch physicochemical properties in waxy rice (*Oryza sativa* L.). *TAG Theoretical and Applied Genetics* 105, 898-905.
- Belkum, A.V., Scherer, S., Alphen, L.V., and Verbrugh, H. (1998). Short-Sequence DNA Repeats in Prokaryotic Genomes. *Microbiology and Molecular Biology Reviews* 62, 275-293.
- Bonamico, N.C., Balzarini, M.G., Arroyo, A.T., Ibañez, M.A., Díaz, D.G., Salerno, J.C., and Renzo, M.A.D. (2010). Association between microsatellites and resistance to Mal de Río Cuarto in maize by discriminant analysis. *International Journal of Experimental Botany* 79, 31-38.
- Bonnecarrère, V., Borsani, O., Díaz, P., Capdevielle, F., Blanco, P., and Monza, J. (2011). Response to photooxidative stress induced by cold in japonica rice is genotype dependent. *Plant Science* 180, 726-732.
- Briat, J.-F., Lobréaux, S., Grignon, N., and Vansuyt, G. (1999). Regulation of plant ferritin synthesis: how and why. *Cellular and molecular life sciences* 56, 155-166.
- Brunner, A.M., Yakovlev, I.A., and Strauss, S.H. (2004). Validating internal controls for quantitative plant gene expression studies. *BMC Plant Biology* 4, 14.
- Bukau, B., and Horwich, A.L. (1998). The Hsp70 and Hsp60 Chaperone Machines. *Cell* 92, 351-366.
- Bustin, S.A. (2002). Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems. *Journal of Molecular Endocrinology* 29, 23-39.
- Caicedo, A.L., Williamson, S.H., Hernandez, R.D., Boyko, A., Fledel-Alon, A., York, T.L., Polato, N.R., Olsen, K.M., Nielsen, R., McCouch, S.R., *et al.* (2007). Genome-Wide Patterns of Nucleotide Polymorphism in Domesticated Rice. *PLoS Genetics* 3, 1745-1756.
- Čikoš, Š., Bukovská, A., and Koppel, J. (2007). Relative quantification of mRNA: comparison of methods currently used for real-time PCR data analysis. *BMC Molecular Biology* 8, 113.
- Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement* 20, 37-46.
- Cook, D. (2004). From The Cover: A prominent role for the CBF cold response pathway in configuring the low-temperature metabolome of Arabidopsis. *Proceedings of the National Academy of Sciences* 101, 15243-15248.
- Counce, P.A., Keisling, T.C., and Mitchell, A.J. (2000). A Uniform, Objective, and Adaptive System for Expressing Rice Development. *Crop Science* 40, 436-443.
- Cruz, M., Pulver, E., Jennings, P.R., Torres, E., and Berrío, L.E. (2001). Identificación de materiales genéticos de arroz para tolerancia a temperaturas bajas. In *Congresso Brasileiro de Arroz Irrigado, Reunião da cultura de arroz irrigado*, B. Porto Alegre-RS, IRGA, ed. (Porto Alegre-RS), pp. 15-17.
- Cummings, C.J., and Zoghbi, H.Y. (2000). Fourteen and counting: unraveling trinucleotide repeat diseases. *Human Molecular Genetics* 9, 909-916.
- Cheng, C., Yun, K.-Y., Ressom, H.W., Mohanty, B., Bajic, V.B., Jia, Y., Yun, S., and de los Reyes, B.G. (2007). An early response regulatory cluster induced by low temperature and hydrogen peroxide in seedlings of chilling-tolerant japonica rice. *BMC Genomics* 8, 175.

Cho, Y.G., Ishii, T., Temnykh, S., Chen, X., Lipovich, L., McCouch, S.R., Park, W.D., Ayres, N., and Cartinhour, S. (2000). Diversity of microsatellites derived from genomic libraries and GenBank sequences in rice (*Oryza sativa* L.). *TAG Theoretical and Applied Genetics* *100*, 713-722.

de los Reyes, B.G., Morsy, M., Gibbons, J., Varma, T.S.N., Antoine, W., McGrath, J.M., Halgren, R., and Redus, M. (2003). A snapshot of the low temperature stress transcriptome of developing rice seedlings (*Oryza sativa* L.) via ESTs from subtracted cDNA library. *TAG Theoretical and Applied Genetics* *107*, 1071-1082.

Deambrosi, E., Méndez, R., and Roel, A. (1997). Estrategia en la producción de arroz: para un mejor aprovechamiento de las principales variables climáticas. Serie Técnica INIA *89*.

Dubouzet, J.G., Sakuma, Y., Ito, Y., Kasuga, M., Dubouzet, E.G., Miura, S., Seki, M., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2003). OsDREB genes in rice, *Oryza sativa* L., encode transcription activators that function in drought-, high-salt- and cold-responsive gene expression. *The Plant Journal* *33*, 751-763.

Ellegren, H. (2004). Microsatellites: simple sequences with complex evolution. *Nature Reviews Genetics* *5*, 435-445.

Eujayl, I., Sorrells, M., Baum, M., Wolters, P., and Powell, W. (2001). Assessment of genotypic variation among cultivated durum wheat based on EST-SSRS and genomic SSRS. *EUPHYTICA* *119*, 39-43.

Excoffier, L., Smouse, P.E., and Quattro, J.M. (1992). Analysis of Molecular Variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics Society of America* *131*, 479-491.

Fabre, E., Dujon, B., and Richard, G.-F. (2002). Transcription and nuclear transport of CAG/CTG trinucleotide repeats in yeast. *Nucleic Acids Research* *30*, 3540-3547.

Fisher, R.A. (1936). The Use of Multiple Measurements in Taxonomic Problems. *Annals of Eugenics* *7*, 179-188.

Fleige, S., and Pfaffl, M.W. (2006). RNA integrity and the effect on the real-time qRT-PCR performance. *Molecular Aspects of Medicine* *27*, 126-139.

Fondon, J.W., and Garner, H.R. (2004). Molecular origins of rapid and continuous morphological evolution. *Proceedings of the National Academy of Sciences* *101*, 18058-18063.

Fowler, S. (2002). Arabidopsis Transcriptome Profiling Indicates That Multiple Regulatory Pathways Are Activated during Cold Acclimation in Addition to the CBF Cold Response Pathway. *The Plant Cell Online* *14*, 1675-1690.

Fraser, L.G., Harvey, C.F., Crowhurst, R.N., and Silva, H.N. (2004). EST-derived microsatellites from *Actinidia* species and their potential for mapping. *TAG Theoretical and Applied Genetics* *108*, 1010-1016.

Garris, A.J., Tai, T.H., Coburn, J., Kresovich, S., and McCouch, S. (2005). Genetic Structure and Diversity in *Oryza sativa* L. *Genetics* *169*, 1631-1638.

Gatchel, J.R., and Zoghbi, H.Y. (2005). Diseases of Unstable Repeat Expansion: Mechanisms and Common Principles. *Nature Reviews Genetics* *6*, 743-755.

Goidin, D. (2001). Ribosomal 18S RNA Prevails over Glyceraldehyde-3-Phosphate Dehydrogenase and β -Actin Genes as Internal Standard for Quantitative Comparison of mRNA Levels in Invasive and Noninvasive Human Melanoma Cell Subpopulations. *Analytical Biochemistry* *295*, 17-21.

Gur-Arie, R., Cohen, C.J., and Eitan, Y. (2000). Simple Sequence Repeats in *Escherichia coli*: Abundance, Distribution, Composition, and Polymorphism. *Genome Research* *10*, 62-71.

Hancock, J.M., and Simon, M. (2005). Simple sequence repeats in proteins and their significance for network evolution. *Gene* *345*, 113-118.

Hegedűs, A., Erdei, S., Janda, T., Szalai, J., Dudits, D., and Horváth, G. (2002). Effects of low temperature stress on ferritin or aldose reductase overexpressing transgenic tobacco plants. *Proceedings of the 7th Hungarian Congress on Plant Physiology* *46*, 97-98.

Hirotsu, N. (2005). The Photosynthetic Properties of Rice Leaves Treated with Low Temperature and High Irradiance. *Plant and Cell Physiology* *46*, 1377-1383.

Huggett, J., Dheda, K., Bustin, S., and Zumla, A. (2005). Real-time RT-PCR normalisation; strategies and considerations. *Genes and Immunity* 6, 279-284.

Imbeaud, S., Graudens, E., Boulanger, V., Barlet, X., Zaborski, P., Eveno, E., Mueller, O., Schroeder, A., and Auffray, C. (2005). Towards standardization of RNA quality assessment using user-independent classifiers of microcapillary electrophoresis traces. *Nucleic Acids Research* 33, e56-e56.

Ince, A.G., Karaca, M., and Naci Onus, A. (2010). Differential expression patterns of genes containing microsatellites in *Capsicum annuum* L. *Molecular Breeding* 25, 645-658.

Ishitani, M., Xiong, L., Stevenson, B., and Zhu, J.-K. (1997). Genetic Analysis of Osmotic and Cold Stress Signal Transduction in *Arabidopsis*: Interactions and Convergence of Abscisic Acid-Dependent and Abscisic Acid-Independent Pathways. *The Plant Cell* 9, 1935-1949.

Jackson, A.L., Chen, R., and Loeb, L.A. (1998). Induction of microsatellite instability by oxidative DNA damage. *Proceedings of the National Academy of Sciences* 95, 12468-12473.

Jain, M., Kaur, N., Garg, R., Thakur, J.K., Tyagi, A.K., and Khurana, J.P. (2006a). Structure and expression analysis of early auxin-responsive Aux/IAA gene family in rice (*Oryza sativa*). *Functional & Integrative Genomics* 6, 47-59.

Jain, M., and Khurana, J.P. (2009). Transcript profiling reveals diverse roles of auxin-responsive genes during reproductive development and abiotic stress in rice. *FEBS Journal* 276, 3148-3162.

Jain, M., Nijhawan, A., Tyagi, A.K., and Khurana, J.P. (2006b). Validation of housekeeping genes as internal control for studying gene expression in rice by quantitative real-time PCR. *Biochemical and Biophysical Research Communications* 345, 646-651.

Jarne, P., and Lagoda, P.J.L. (1996). Microsatellites from molecules to populations and back. *Tree* 11, 424-429.

Jeong, S.W., Choi, S.M., Lee, D.S., Ahn, S.N., Hur, Y., Chow, W.S., and Park, Y.I. (2002). Differential susceptibility of photosynthesis to light chilling stress in rice (*Oryza sativa* L.) depends on the capacity for photochemical dissipation of light. *Molecules and cells* 13, 419-428.

Kantety, R.V., La Rota, M., Matthews, D.E., and Sorrels, M.E. (2002). Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat. *Plant Molecular Biology* 48, 501-510.

Kariya, K. (2003). Chilling injuries in reproductive phase of rice plants. In 3ra Conferencia Internacional de arroz de Clima Templado (Uruguay).

Kashi, Y., and King, D. (2006). Simple sequence repeats as advantageous mutators in evolution. *Trends in Genetics* 22, 253-259.

Katti, M.V., Ranjekar, P.K., and Gupta, V.S. (2001). Differential Distribution of Simple Sequence Repeats in Eukaryotic Genome Sequences. *Molecular Biology and Evolution* 18, 1161-1167.

Khong, G.N., Richaud, F., Coudert, Y., Pati, P.K., Santi, C., Périn, C., Breidler, J.-C., Meynard, D., Vinh, D.N., Guiderdoni, E., *et al.* (2008). Modulating Rice Stress Tolerance by Transcription Factors. *Biotechnology and Genetic Engineering Reviews* 25, 381-403.

Khush, G.S. (1997). Origin, dispersal, cultivation and variation of rice. *Plant Molecular Biology* 35, 25-34.

Kim, B.-R., Nam, H.-Y., Kim, S.-U., Kim, S.-I., and Chang, Y.-J. (2003). Normalization of reverse transcription quantitative-PCR with housekeeping genes in rice. *Biotechnology Letters* 25, 1869-1872.

Kovach, M.J., Sweeney, M.T., and McCouch, S.R. (2007). New insights into the history of rice domestication. *Trends in Genetics* 23, 578-587.

Kruglyak, S., Durrett, R.T., Schug, M.D., and Aquadro, C.F. (1998). Equilibrium distributions of microsatellite repeat length resulting from a balance between slippage events and point mutations. *Proceedings of the National Academy of Sciences* 95, 10774-10778.

- Lagercrantz, U., Ellegren, H., and Andersson, L. (1993). The abundance of various polymorphic microsatellite motifs differs between plants and vertebrates. *Nucleic Acids Research* *21*, 1111-1115.
- Lawson, M., and Zhang, L. (2006a). Distinct patterns of SSR distribution in the *Arabidopsis thaliana* and rice genomes. In *Genome Biology*, pp. R14.
- Lawson, M., and Zhang, L. (2006b). Distinct patterns of SSR distribution in the *Arabidopsis thaliana* and rice genomes. *Genome Biology* *7*, R14.
- Lehmann, T., Hawley, W.A., and Collins, F.H. (1996). An Evaluation of Evolutionary Constraints on Microsatellite Loci Using Null Alleles. *Genetics* *144*, 1155-1163
- Li, Y.C. (2004). Microsatellites Within Genes: Structure, Function, and Evolution. *Molecular Biology and Evolution* *21*, 991-1007.
- Livak, K.J., and Schmittgen, T.D. (2001). Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2^{-\Delta\Delta CT}$ Method. *Methods* *25*, 402-408.
- Londo, J.P. (2006). From the Cover: Phylogeography of Asian wild rice, *Oryza rufipogon*, reveals multiple independent domestications of cultivated rice, *Oryza sativa*. *Proceedings of the National Academy of Sciences* *103*, 9578-9583.
- Lu, G., Gao, C., Zheng, X., and Han, B. (2009). Identification of OsbZIP72 as a positive regulator of ABA response and drought tolerance in rice. *Planta* *229*, 605-615.
- M.G.Murray, and W.F.Thompson (1980). Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Research* *8*, 4321-4325.
- Mankodi, A., Takahashi, M.P., Jiang, H., Beck, C.L., Bowers, W.J., Moxley, R.T., Cannon, S.C., and Thornton, C.A. (2002). Expanded CUG Repeats Trigger Aberrant Splicing of CIC-1 Chloride Channel Pre-mRNA and Hyperexcitability of Skeletal Muscle in Myotonic Dystrophy. *Molecular Cell* *10*, 35-44.
- Marcotte, E.M., Pellegrini, M., Yeates, T.O., and Eisenberg, D. (1998). A Census of Protein Repeats. *Journal of Molecular Biology* *293*, 151-160.
- Maxwell, K., and Johnson, G.N. (2000). Chlorophyll fluorescence-a practical guide *Journal of Experimental Botany* *51*, 659-668.
- McCouch, S.R., Teytelman, L., Xu, Y., Lobos, K.B., Clare, K., Walton, M., Fu, B., Maghirang, R., Li, Z., Xing, Y., *et al.* (2002). Development and Mapping of 2240 New SSR Markers for Rice (*Oryza sativa* L.). *DNA Research* *9*, 199-207.
- Mcharo, M., Labonte, D.R., Oard, J.H., Kays, S.J., and McLaurin, W.J. (2004). Linking quantitative traits with AFLP markers in sweetpotatoes using discriminant analysis. *Acta Horticulturae* *637*, 285-293.
- Morgante, M., Hanafey, M., and Powell, W. (2002). Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. *Nature Genetics* *30*, 194-200.
- Murashige, T., and Skoog, F. (1962). A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures. *Physiologia Plantarum* *15*, 473-497.
- Murata, N., and Los, D.A. (1997). Membrane Fluidity and Temperature Perception. *Plant Physiology* *115*, 875-879.
- Nakashima, K., Ito, Y., and Yamaguchi-Shinozaki, K. (2009). Transcriptional Regulatory Networks in Response to Abiotic Stresses in *Arabidopsis* and Grasses. *Plant Physiology* *149*, 88-95.
- Ni, J., Colowit, P.M., and Mackill, D.J. (2002). Evaluation of Genetic Diversity in Rice Subspecies Using Microsatellite Markers. *Crop Science* *42*, 601-607.
- Nicot, N. (2005). Housekeeping gene selection for real-time RT-PCR normalization in potato during biotic and abiotic stress. *Journal of Experimental Botany* *56*, 2907-2914.
- Nishiyama, Y., Allakhverdiev, S.I., and Murata, N. (2006). A new paradigm for the action of reactive oxygen species in the photoinhibition of photosystem II. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* *1767*, 414-421.
- Oh, S.J. (2005). *Arabidopsis* CBF3/DREB1A and ABF3 in Transgenic Rice Increased Tolerance to Abiotic Stress without Stunting Growth. *Plant Physiology* *138*, 341-351.

- Oka, H.-I., and Morishima, H. (1982). Phylogenetic differentiation of cultivated rice, XXIII. Potentiality of wild progenitors to evolve the Indica and Japonica types of rice cultivars. *Euphytica* 31, 41-50.
- Pfaffl, M.W. (2004). Quantification strategies in real-time PCR. In A-Z of quantitative PCR, S.A. Bustin, ed. (La Jolla, CA, International University Line), pp. 87-112.
- Plieth, C., Hansen, U.-P., Knight, H., and Knight, M.R. (1999). Temperature sensing by plants: the primary characteristics of signal perception and calcium response. *Plant Science* 18, 491-497.
- Powell, W., Machray, G.C., and Provan, J. (1996). Polymorphism revealed by simple sequence repeats. *Trends in Plant Science* 1, 215-222.
- Pritchard, J.K., Stephens, M., and Donnelly, P. (2000). Inference of Population Structure Using Multilocus Genotype Data. *Genetics* 155, 945-959.
- Pupko, T., and Graur, D. (1999). Evolution of Microsatellites in the Yeast *Saccharomyces cerevisiae*: Role of Length and Number of Repeated Units. *Journal of Molecular Evolution* 48, 313-316.
- QIAGEN (2009). <QuantiTect_Reverse_Transcription_Handbook.pdf>.
- Rabbani, M.A., Maruyama, K., Abe, H., Khan, M.A., Katsura, K., Ito, Y., Yoshiwara, K., Seki, M., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2003). Monitoring Expression Profiles of Rice Genes under Cold, Drought, and High-Salinity Stresses and Abscisic Acid Application Using cDNA Microarray and RNA Gel-Blot Analyses. *Plant Physiology* 133, 1755-1767.
- Radonić, A., Thulke, S., Mackay, I.M., Landt, O., Siegert, W., and Nitsche, A. (2004). Guideline to reference gene selection for quantitative real-time PCR. *Biochemical and Biophysical Research Communications* 313, 856-862.
- Robinson, T.L., Sutherland, I.A., and Sutherland, J. (2007). Validation of candidate bovine reference genes for use with real-time PCR. *Veterinary Immunology and Immunopathology* 115, 160-165.
- Ruan, W., and Lai, M. (2007). Actin, a reliable marker of internal control? *Clinica Chimica Acta* 385, 1-5.
- Sang, T., and Ge, S. (2007). The Puzzle of Rice Domestication. *Journal of Integrative Plant Biology* 49, 60-768.
- Schmidt, A.L., and Mitter, V. (2004). Microsatellite mutation directed by an external stimulus. *Mutation Research* 568, 233-243.
- Schmittgen, T.D., Zakrajsek, B.A., Mills, A.G., Gorn, V., Singer, M.J., and Reed, M.W. (2000). Quantitative Reverse Transcription-Polymerase Chain Reaction to Study mRNA Decay: Comparison of Endpoint and Real-Time Methods. *Analytical Biochemistry* 285, 194-204.
- Schroeder, A., Mueller, O., Stocker, S., Salowsky, R., Leiber, M., Gassmann, M., Lightfoot, S., Menzel, W., Granzow, M., and Ragg, T. (2006). The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC Molecular Biology* 7, 3.
- Shinozaki, K., and Yamaguchi-Shinozaki, K. (2000). Molecular responses to dehydration and low temperature. *Current Opinion in Plant Biology* 3, 217-223.
- Shockey, J., Schnurr, J., and Browse, J. (2000). Characterization of the AMP-binding protein gene family in *Arabidopsis thaliana*: will the real acyl-CoA synthetases please stand up? *Biochemical Society Transactions* 28, 955-957.
- Stansel, J. (1975). Effective utilization of sunlight. In *Six Decades of Rice Research in Texas*, T.A.M. University, ed. (Texas, USA).
- Stein, R.J., Ricachenevsky, F.K., and Fett, J.P. (2009). Differential regulation of the two rice ferritin genes (*OsFER1* and *OsFER2*). *Plant Science* 177, 563-569.
- Stratagene (2007). Introduction to Quantitative PCR. In *Methods and applications guide* (Stratagene).
- Sweeney, M., and McCouch, S. (2007). The Complex History of the Domestication of Rice. *Annals of Botany* 100, 951-957.
- Toth, G. (2000). Microsatellites in Different Eukaryotic Genomes: Survey and Analysis. *Genome Research* 10, 967-981.

- Trifonov, E.N. (2003). Tuning function of tandemly repeating sequences: a molecular device for fast adaptation. In *Evolutionary theory and processes: modern horizons, papers in honor of Eviatar Nevo, S.P. Wasser, ed.* (Amsterdam, The Netherlands, Kluwer Academic Publishers), pp. 1-24.
- Turner, J., Greville, K., Murphy, E., and Hooks, M. (2005). Characterization of Arabidopsis fluoroacetate-resistant mutants reveals the principal mechanism of acetate activation for entry into the glyoxylate cycle. *Journal of Biological Chemistry* *280*, 2780-2787.
- Varshney, R.K., Graner, A., and Sorrells, M.E. (2005). Genic microsatellite markers in plants: features and applications. *Trends in Biotechnology* *23*, 48-55.
- Verhoeven, A.S., Demmig-Adams, B., and III, W.W.A. (1997). Enhanced Employment of the Xanthophyll Cycle and Thermal Energy Dissipation in Spinach Exposed to High Light and N Stress. *Plant Physiology* *113*, 817-824.
- Wang, W., Vinocur, B., and Altman, A. (2003). Plant responses to drought, salinity and extreme temperatures: towards genetic engineering for stress tolerance. *Planta* *218*, 1-14.
- Wang, W., Vinocur, B., Shoseyov, O., and Altman, A. (2004). Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. *Trends in Plant Science* *9*, 244-252.
- Witten, I.H., and Frank, E. (2005). *Data Mining Practical Machine Learning Tools and Techniques* _2nd edition, Vol 2, 2 edn (San Francisco, CA, Morgan Kaufmann).
- Wren, J.D., Forgacs, E., III, J.W.F., Pertsemliadis, A., Cheng, S.Y., Gallardo, T., Williams, R.S., Shohet, R.V., Minna, J.D., and Garner, H.R. (2000). Repeat Polymorphisms within Gene Regions: Phenotypic and Evolutionary Implications. *Journal of Human Genetics* *67*, 345-356.
- Xiong, L., Schumaker, K.S., and Zhu, J.-K. (2002). Cell Signaling during Cold, Drought, and Salt Stress. *The Plant Cell* *14*, S165-S183.
- Xue, T., Wang, D., Zhang, S., Ehling, J., Ni, F., Jakab, S., Zheng, C., and Zhong, Y. (2008). Genome-wide and expression analysis of protein phosphatase 2C in rice and Arabidopsis. *BMC Genomics* *9*, 550.
- Yamaguchi-Shinozaki, K., and Shinozaki, K. (1992). Characterization of the expression of a desiccation-responsive rd29 gene of Arabidopsis thaliana and analysis of its promoter in transgenic plants. *Molecular and General Genetics* *236*, 331-340.
- Yazaki, J. (2004). Transcriptional profiling of genes responsive to abscisic acid and gibberellin in rice: phenotyping and comparative analysis between rice and Arabidopsis. *Physiological Genomics* *17*, 87-100.
- Yoshida, S. (1981). *Fundamentals of rice crop science, Vol 1* (Manila, Philippines, The International Rice Research Institute).
- Young, E.T., Sloan, J.S., and Riper, K.V. (2000). Trinucleotide Repeats Are Clustered in Regulatory Genes. *Genetics* *154*, 1053-1068.
- Yun, K.-Y., Park, M.R., Mohanty, B., Herath, V., Xu, F., Mauleon, R., Wijaya, E., Bajic, V.B., Bruskiwich, R., and de los Reyes, B.G. (2010). Transcriptional regulatory network triggered by oxidative signals configures the early response mechanisms of japonica rice to chilling stress. *BMC Plant Biology* *10*, 16.
- Zhang, N., Xu, Y., Akash, M., McCouch, S., and Oard, J.H. (2005). Identification of candidate markers associated with agronomic traits in rice using discriminant analysis. *Theoretical and Applied Genetics* *110*, 721-729.
- Zhang, Q., Maroof, M.A.S., Lu, T.Y., and Shen, B.Z. (1992). Genetic diversity and differentiation of indica and japonica rice detected by RFLP analysis TAG *Theoretical and Applied Genetics* *83*, 495-499.
- Zhang, X.-c., Yu, X., Zhang, H.-j., and Song, F.-m. (2009). Molecular characterization of a defense-related AMP-binding protein gene, OsBIABP1, from rice. *Journal of Zhejiang University SCIENCE B* *10*, 731-739.
- Zhang, X., Fowler, S.G., Cheng, H., Lou, Y., Rhee, S.Y., Stockinger, E.J., and Thomashow, M.F. (2004). Freezing-sensitive tomato has a functional CBF cold response pathway, but a CBF regulon that differs from that of freezing-tolerant Arabidopsis. *The Plant Journal* *39*, 905-919.

Zhu, J., Dong, C.-H., and Zhu, J.-K. (2007). Interplay between cold-responsive gene regulation, metabolism and RNA processing during plant cold acclimation. *Current Opinion in Plant Biology* 10, 290-295.