

MÉTODO PARA LA IDENTIFICACIÓN Y CARACTERIZACIÓN FUNCIONAL DE GENES DE INTERÉS PARA EL MEJORAMIENTO AGRONÓMICO DE ESPECIES

5 CAMPO DE LA INVENCION

La presente invención se relaciona con el campo técnico de los métodos y procesos para el mejoramiento agronómico de germoplasma vegetal. En particular, la presente invención se relaciona con un método integrativo para la identificación y caracterización funcional de genes con un elevado potencial de implementación exitosa en procesos de mejoramiento agronómico de especies.

ANTECEDENTES DE LA INVENCION

El aumento poblacional, el cambio climático y el surgimiento de nuevos patógenos virales o bacterianos son algunos de los problemas fundamentales que deberá enfrentar la humanidad en los años venideros. Estos tres problemas afectan directamente a la agricultura mundial retando la producción agrícola en un contexto en el cual se deberá aumentar la producción en condiciones cada vez más restrictivas, desde paradigmas complejos como el medioambiental, hasta el agroeconómico y de innovación técnico-científica.

Modelos conservadores, prevén que el crecimiento de la población mundial se duplicará durante los primeros 45 años del siglo XXI, como consecuencia, la producción de alimento tendrá que triplicarse para garantizar una seguridad alimentaria, obligando a un incremento en la productividad agrícola y a un manejo sostenible de la biomasa residual. No solamente es necesario producir más, es decisivo producir mejor, esto lleva a conectar características de rendimiento con factores complementarios como la calidad nutricional de lo que se cultiva y el uso innovador del material residual para la generación de productos con valor agregado. La integración pertinente de estos factores solo es posible a través del uso de tecnologías, metodologías y estrategias sistémicas e integrativas.

La revisión a la política agrícola colombiana realizada por el PNUD, la FAO y la OECD revela limitaciones críticas del sector frente a producción, innovación e implementación de tecnologías. A pesar de ser la principal actividad económica de los territorios rurales, la agricultura nacional ha experimentado múltiples crisis estructurales y una disminución en su contribución al PIB, del 16.5% al 6.2% entre 1990 y 2015, respectivamente.

El Plan Estratégico Colombiano para la Ciencia, la Tecnología y la Innovación del Sector Agropecuario (PECTIA) enfatiza la necesidad de estructurar y proponer estrategias holísticas y convergentes desde la investigación, integradas al desarrollo económico y tecnológico de cultivos claves para el país.

- 5 En este contexto, la generación de nuevas variedades con mayor rendimiento, aunadas al mejoramiento de características complementarias como la calidad nutricional o la tolerancia a estreses abióticos y resistencia a estreses bióticos, son productos tecnológicos de implementación inmediata y de alto impacto para el sector. Desde esta perspectiva, existe un llamado al desarrollo nacional de una
- 10 nueva revolución verde fundamentada en la generación tanto de nuevas variedades como de nuevas tecnologías y metodologías holísticas interdependientes que incrementen la productividad y competitividad de los productos agrícolas nacionales, garanticen la seguridad alimentaria y promuevan el desarrollo de sistemas ambientalmente sostenibles.
- 15 Distintos desarrollos biotecnológicos han contribuido a la generación de nuevas variedades agronómicas que buscan ser una opción pertinente en respuesta a las limitaciones, cada vez más drásticas, del entorno agrícola mundial. No obstante, aún existe un desbalance entre las nuevas variedades generadas y las requeridas para sostener la demanda mundial de alimentos en un contexto complejo debido a
- 20 factores sociales, ecológicos, climáticos, edáficos y agronómicos altamente restrictivos.

En las últimas décadas se han desarrollado tecnologías que permiten, o bien transferir información genética entre organismos de la misma especie (*cisgénesis*), entre individuos de distintas especies (*transgénesis*), o editar fragmentos del

25 genoma en una especie o variedad que se desea mejorar agronómicamente, por medio de las diferentes tecnologías de edición genética disponibles. Sin embargo, para poder aplicar estas tecnologías es necesario tener un conocimiento funcional, o del gen que se desea transferir o de la parte del genoma que se desea editar, generando un obstáculo, debido a que las tecnologías de vanguardia, utilizadas

30 para la generación de nuevas variedades agronómicas, no encuentran genes o circuitos génicos lo suficientemente caracterizados funcionalmente, que permitan su implementación.

Con relación a la modificación genética para la mejora y producción de nuevas variedades, se ha estimado que de 2016 a 2019, los cultivos transgénicos se han

35 cultivado en alrededor de 186 millones de hectáreas de tierra en todo el mundo, representando un mercado de 186 mil millones de dólares. De la misma manera, se estima que sólo en el año 2022, los cultivos genéticamente modificados representaron un mercado de cerca de 65 mil millones de dólares, donde más de la mitad (~52%) de dichos beneficios económicos, corresponden a los cultivos

mejorados genéticamente, presentes en los países en vía de desarrollo (Camille D. Ryan, et al, 2020).

En consecuencia, es posible concluir que, si bien en el estado de la técnica se encuentran reportadas herramientas bastante elaboradas que permiten la modificación o edición genética, no obstante, se carece de técnicas que faciliten la
5 detección y caracterización de dichas secuencias génicas, sobre las cuales debe operar la modificación genética.

En el campo de la literatura de patentes, se observan desarrollos como el descrito por la solicitud EP1442141A4, el cual revela métodos para identificar un subconjunto
10 de genes que se expresan diferencialmente en un estado biológico dado o en un lugar biológico de interés, determinando una distancia de probabilidad ventajosamente grande entre los vectores aleatorios que representan los niveles de expresión de los genes de interés, mediante análisis estadístico de datos de micromatrices a partir del reconocimiento de patrones y la clasificación
15 multidimensional. La invención también proporciona un mecanismo de búsqueda aleatoria con validación cruzada para un subconjunto de genes basado en varias distancias de probabilidad entre dos vectores aleatorios.

La solicitud WO2017106663A1 describe métodos y composiciones para la identificación de genes novedosos útiles para modular características agronómicas
20 deseadas en plantas. Este invención divulga métodos para identificar genes específicos de línea o de grupos de plantas que muestran alteración de la expresión génica en respuesta a cambios en la expresión de un gen modulador del clúster génico, confiriendo alteraciones en las características agronómicas en la planta.

La patente US5437697A revela un método para identificar marcadores genéticos
25 vinculados a alelos que mejoran el potencial de rendimiento de una especie de cultivo. Al realizar un análisis de marcadores genéticos sobre un conjunto de líneas élite actuales y derivadas de poblaciones ancestrales con décadas de fitomejoramiento, se pudieron determinar y comparar las frecuencias alélicas esperadas y observadas dentro de las poblaciones élite en numerosos loci
30 polimórficos. Dado que el esfuerzo tradicional de fitomejoramiento ha utilizado consistentemente el rendimiento como criterio de selección, se utilizaron desviaciones de la frecuencia alélica esperada en ciertos loci para identificar los alelos que confieren mejoría en el potencial de rendimiento, pudiéndose seleccionar una progenie agronómicamente superior utilizando marcadores genéticos.

35 La patente EP4066248A1 divulga un método (100) para caracterizar el impacto funcional de un grupo de variantes génicas, que comprende: obtener (110) información que relacione al menos una de las variantes con información de expresión génica, variación del número de copias o efectos epigenéticos; determinar (120) un estado específico para la variante; determinar (130) una

- correlación de la expresión con las variantes identificadas, para determinar si la variante tiene un efecto sobre la expresión génica; determinar (140) si la variante tiene un impacto funcional en un gen diana; determinar (150) si la variante identificada tiene un efecto sobre el número de copias (CNV) basándose en el
- 5 impacto epigenético, sobre la expresión de un gen; ajustar (160), en relación al impacto epigenético y al CNV, el estado de regulación de la expresión basado en variantes y/o en genes; y reportar (170), al menos, el estado de regulación de la expresión ajustado, basado en variantes y/o genes para cada una de las l variantes y/o genes de la muestra genómica.
- 10 Si bien en el estado de la técnica se reportan distintas técnicas moleculares para la identificación y caracterización de genes promisorios para el mejoramiento agronómico de variedades ante condiciones de estrés bióticos y abióticos, incluyendo métodos experimentales como QTLs (Quantitative Trait Loci o GWAS (Genome Wide Association Studies) que determinan una zona genómica potencial,
- 15 o un marcador genómico particular con efectos fenotípicos, no obstante, estos métodos no logran, per sé, la identificación del gen asociado con el fenotipo estudiado, es decir, no existe una asociación entre la región genómica caracterizada o el marcador detectado y una causa que justifique, biológicamente, la aparición del fenotipo.
- 20 Tomando como base los elementos del estado de la técnica expuestos con anterioridad, se observa que resulta necesario desarrollar e implementar métodos que permitan la adecuada identificación y caracterización funcional de secuencias génicas que presenten un alto potencial de implementación en procesos de mejoramiento agronómico para la producción o fitomejoramiento de nuevas
- 25 variedades vegetales.

SUMARIO DE LA INVENCION.

- En un único objeto, la presente invención revela un método que incorpora la combinación de distintas herramientas de caracterización ómica y de análisis
- 30 bioinformáticos que permite, a partir de miles de genes candidatos y la implementación causal de distintos filtros sucesivos, identificar y caracterizar aquellos genes con un mayor potencial de implementación exitosa en procesos de mejoramiento agronómico de especies de interés.

- El método divulgado por la presente invención es susceptible de ser aplicado en
- 35 distintos modelos biológicos, tanto en especies modelo como en especies agronómicamente relevantes, implementándose para la identificación y caracterización de genes asociados a distintos procesos y respuestas biológicas, tales como respuestas a estreses tanto bióticos como abióticos, respuestas

fisiológicas a estímulos ambientales e incluso para detectar genes vinculados a procesos morfogénéticos asociados a etapas puntuales del desarrollo de una planta.

El método objeto de divulgación en la presente descripción incorpora la concatenación de etapas de obtención de datos transcriptómicos y epigenómicos generados sobre los modelos que se pretenden evaluar y sobre las condiciones de las cuales se requiera la obtención de una respuesta fenotípica.

En una modalidad particular de implementación del método objeto de la presente invención, a partir de datos derivados de un estrés abiótico particular, a saber, las altas concentraciones de Aluminio (Al), que en combinación con un pH inferior a 5, resultan perjudiciales para el desarrollo de las plantas y en especies agrónomicamente relevantes, limita su producción o impide su uso comercial, se analiza el efecto del Al sobre variedades de respuesta contrastante al estrés por Al en una especie de arroz silvestre (*Oriza glumaepatula*) y una especie de arroz comercial (*Oriza sativa*).

En este contexto, la presente invención proporciona un procedimiento que acopla resultados derivados de procesos de genómica estructural y funcional, en donde se logra relacionar, directamente, la expresión del gen con el fenotipo de interés, asociando los datos de expresión con mecanismos de regulación epigenética, configurando una relación de causalidad en donde mecanismos epigenéticos controlan la expresión de los genes postulados como candidatos.

Los objetos anteriormente descritos, al igual que los objetos adicionales a que hubiere lugar, serán expuestos al detalle y con la suficiencia necesaria en el capítulo descriptivo que se divulga a continuación, el cual constituirá el fundamento del capítulo reivindicatorio.

BREVE DESCRIPCIÓN DE LAS FIGURAS.

FIGURA 1. Muestra un esquema de las etapas del método propuesto en la presente descripción y de las distintas fases de análisis, en donde se ilustra las etapas del procedimiento objeto de la presente invención y su respectiva concatenación para la obtención y caracterización de genes candidatos de interés.

FIGURA 2. Muestra un esquema de las estrategias ómicas a utilizar, necesarias para obtener la información base que permitirá la implementación del método propuesto. Para el ejemplo de realización particular expuesto en esta figura se ilustra la estrategia metodológica a ejecutar sobre variedades de arroz comercial (*Oryza sativa*) con respuesta contrastante ante condiciones de estrés por Al.

El mismo procedimiento es ejecutado para el análisis de la variedad silvestre de arroz (*Oryza glumaepatula*).

FIGURA 3. Muestra un diagrama de flujo que indica las etapas de una modalidad de realización del método propuesto y los resultados obtenidos luego de la implementación de cada fase del procedimiento sobre variedades de respuesta contrastante ante condiciones de estrés por AI en una especie de arroz comercial (*Oryza sativa*).

FIGURA 4. Muestra un diagrama de flujo que indica las etapas de una modalidad de realización del método propuesto y los resultados obtenidos luego de la implementación de cada fase del procedimiento sobre variedades de respuesta contrastante ante condiciones de estrés por AI en una especie de arroz silvestre (*Oryza glumaepatula*).

DESCRIPCIÓN DETALLADA DE LA INVENCION

La presente invención surge como respuesta a la necesidad de proveer soluciones que permitan la adecuada identificación y caracterización funcional de genes con un alto potencial de ser utilizados en procesos de mejoramiento genético de especies, a partir de la generación, evaluación, concatenación y contrastación de datos epigenómicos y transcriptómicos.

La descripción de la realización de la presente invención no pretende limitar su alcance, sino servir como un ejemplo particular de la misma. Se espera que una persona versada en la materia comprenda que las modalidades equivalentes no se apartan del espíritu y alcance de la presente invención en su forma más amplia.

La primera etapa del método consiste en la obtención de datos transcriptómicos derivados de experimentos sistémicos de análisis transcripcional como microarreglos o secuenciación masiva de ARNs (RNA-Seq) que, según el nivel de significancia estadística utilizado, valores *p* y aumentos logarítmicos en los cambios de expresión, pueden derivar en la obtención de miles o cientos de genes diferencialmente expresados.

De manera concurrente al desarrollo de la etapa de evaluación transcriptómica, el modelo experimental debe ser evaluado, utilizando también técnicas epigenómicas globales como la secuenciación de ADN con conversión por bisulfito, que permita la obtención de genes o promotores de genes diferencialmente metilados a una escala epigenómica, esto también deriva en la obtención de listas compuestas por miles de genes diferencialmente metilados.

En la tercera fase del proceso utilizando, opcionalmente, herramientas de análisis de datos disponibles en el estado de la técnica, se combinan ambas listas obtenidas (genes diferencialmente expresados y genes diferencialmente metilados, esto es, genes diferencialmente expresados asociados a diferencias epigenómicas basadas en citocinas metiladas) para obtener un set de genes caracterizados por tener patrones de metilación en donde la hiper o hipo metilación diferencial de promotores o de los cuerpos de los genes en las variedades evaluadas, concuerda con variaciones en sus niveles de expresión, estando éstos o reprimidos o sobre-expresados transcripcionalmente.

Este filtro, es decir, la intersección entre las listas de los genes regulados epigenómicamente con un efecto directo sobre sus niveles de expresión, reduce el grupo de genes convergentes a algunos cientos de genes. Posteriormente, se realizan análisis de correlación, opcionalmente, a través de programas de comparación de archivos, textos y listas disponibles en el estado de la técnica, en donde las diferencias en la expresión transcripcional del gen a retener, debe coincidir con un estado de metilación específico del promotor evaluado: promotor hiper-metilado que se asocia a una baja expresión transcripcional del gen, o viceversa, promotor hipo-metilado que coincide con una alta expresión transcripcional del gen. Esta convergencia reduce la lista a decenas de genes en donde la epigenética está regulando los niveles de expresión de los genes hasta acá filtrados.

La etapa subsiguiente del método se desarrolla a partir de la evaluación de la lista de genes obtenida en la etapa anterior, en donde se han retenido genes con un patrón de metilación específico que repercute en la expresión transcripcional del mismo, correlacionando genes sobrerregulados cuyo promotor se encuentre hipometilado o genes bajoregulados cuyo promotor se encuentre hipermetilado.

Esta correlación se realiza, opcionalmente, vía qPCR en donde se confirma la expresión diferencial de los genes retenidos hasta el momento por el proceso de selección. Este proceso de validación reduce la lista de genes a unos cuantos genes candidatos (1 a 5).

En la última etapa del procedimiento, esta lista final de genes obtenida es evaluada mediante análisis de coexpresión, opcionalmente, mediante la implementación de herramientas web como RiceFRIEND o COXPRESdb acoplado a análisis de enriquecimiento funcional (GO) opcionalmente, vía bases de datos como AMIGO, Panther o ShinyGO, confirmando si los genes retenidos, están asociados con los procesos biológicos ligados a la función objetivo evaluada.

Los ejemplos que se exponen a continuación son presentados con el objetivo de evidenciar modalidades particulares de realización del método objeto de la presente invención que sustentan su aplicación para la selección puntual de genes con un

mayor potencial de implementación exitosa en procesos de mejoramiento agronómico.

En este sentido, a partir de la obtención de los datos tanto transcriptómicos como epigenómicos, el desarrollo de las etapas comprendidas dentro del método objeto de la presente invención genera la evaluación y selección de genes potencialmente implicados en tolerancia a estreses abióticos, resistencia a estreses bióticos, respuestas fisiológicas diferenciales ante estímulos ambientales específicos, entre otros estímulos.

En todo caso, se entiende que la exposición de los siguientes ejemplos no configura limitación del alcance del método objeto de descripción.

EJEMPLO 1.

Tolerancia al estrés por Al.

Se evaluó durante un periodo de exposición de 10 días al Aluminio, variedades contrastantes ante el estrés, tanto en una especie comercial de arroz (*Oryza sativa*), como en una especie silvestre (*Oryza glumaepatula*), obteniendo para ambas especies, luego de la implementación del método objeto de la presente invención, genes potencialmente candidatos para el mejoramiento agronómico de arroz ante el estrés estudiado.

La Tabla 1 detalla el germoplasma evaluado mediante el desarrollo de las etapas del método mencionadas en la Figura 2.

Tabla 1. Germoplasma de arroz evaluado mediante el método de la presente invención.

Especie	Variedad	Nivel de tolerancia
<i>Oryza sativa</i> (especie comercial)	Azucena (AZU)	Tolerante
	BGI9311 (BGI)	Susceptible
<i>Oryza glumaepatula</i> (especie silvestre)	OG97	Tolerante
	OG131	Susceptible

En primera instancia, para la especie comercial de arroz analizada (*Oryza sativa*), se evaluó la respuesta ante el estrés por Al, utilizando dos variedades con comportamientos contrastantes de respuesta ante el estrés; una variedad tolerante (AZU) y una susceptible (BGI).

En el caso de los tratamientos control se implementaron cultivos hidropónicos sin contenido de Al, mientras que, para los tratamientos de exposición, las plantas

crecieron en cultivos hidropónicos complementado con el metal a una concentración de 100 μ M.

Luego de diez días de evaluación, raíces pertenecientes a las variedades evaluadas y crecidas tanto en cultivos control como tratamiento fueron colectadas y maceradas
5 manualmente en nitrógeno líquido para preservar las características de los ácidos nucleicos. El ARN fue extraído mediante kits específicos de extracción comercialmente disponibles siguiendo el protocolo sugerido por la compañía de manufactura, y cuantificado por espectrofotometría para garantizar una cantidad y calidad adecuada de los ácidos nucleicos extraídos. El ARN extraído fue la materia
10 prima para realizar los análisis transcriptómicos. Complementariamente, el ADN fue extraído utilizando kits específicos de extracción, comercialmente disponibles, siguiendo las instrucciones sugeridas por la compañía de manufactura del kit y cuantificado por espectrofotometría para para garantizar una cantidad y calidad adecuada de los ácidos nucleicos extraídos. El ADN extraído fue la materia prima
15 para ejecutar los análisis epigenómicos.

Ambos sets de datos constituyen la materia prima para implementar el método propuesto en la presente invención para la detección de genes con un potencial elevado de implementación exitosa en procesos de mejoramiento agronómico, (en la modalidad particular objeto de exposición, generación de variedades con alta
20 tolerancia al estrés por AI).

Las variedades Azucena y BGI de respuesta contrastante ante el estrés por AI, fueron evaluadas siguiendo las etapas descritas en la Figura 2. El análisis transcriptómico ejecutado, basado en un análisis global de transcritos por secuenciación (RNA-Seq) que incluye filtrado de las lecturas de alta calidad, mapeo
25 de las lecturas a un genoma de referencia para la obtención de los genes expresados y conteos de dichos genes, el cual reveló 3336 genes diferencialmente expresados.

Para el análisis epigenómico, las mismas variedades fueron evaluadas por medio de la técnica de secuenciación bisulfito, la cual por comparación con un genoma de referencia, determina el número y ubicación sobre el genoma de las citosinas metiladas, detectando 8419 genes diferencialmente metilados.
30

Al interceptar ambas listas obtenidas, se obtuvieron 156 genes. Estos genes representan secuencias transcripcionalmente activas que también se caracterizan por estar diferencialmente metiladas.

Posteriormente, se realizó el análisis de correlación, el cual reveló 6 genes en
35 donde, las diferencias en la expresión transcripcional del gen coincidieron con un estado de metilación específico del promotor evaluado. Finalmente, los análisis de qPCR, los cuales cuantifican por técnicas de fluorescencia el nivel de expresión

genes particulares, confirmaron al gen FDH2 como un gen con expresión reducida, cuyo bajo nivel de expresión coincide con un patrón de hiper-metilación presente en su promotor.

Además, análisis complementarios de coexpresión y de ontologías, los cuales permiten determinar los distintos procesos biológicos en los que se enmarca la acción de un gen de interés, confirmaron la relación funcional del gen FDH2 con procesos asociados a respuestas a estreses abióticos.

En relación con la Figura 3, el diagrama de flujo muestra el número de genes retenido en cada etapa del método objeto de la presente invención y describe como la implementación de los distintos filtros sucesivamente concatenados, resulta útil para reducir un amplio espectro de genes candidatos, a un único gen, pero de gran potencial para ser empleado en eventos de transformación agronómica.

En consecuencia, el método de identificación, caracterización y selección propuesto en la presente invención presenta la capacidad de filtrar genes novedosos y con potencial de implementación exitosa en procesos de mejoramiento agronómico ante condiciones de estrés por AI en *Oryza sativa*.

EJEMPLO 2.

Las variedades de arroz silvestre (*Oryza glumaepatula*) OG97 y OG131 de respuesta contrastante ante el estrés por AI fueron evaluadas siguiendo las etapas identificadas en la Figura 2. El análisis transcriptómico ejecutado, basado en un análisis global de transcritos (RNA-Seq), que incluye filtrado de las lecturas de alta calidad, mapeo de las lecturas a un genoma de referencia para la obtención de los genes expresados y conteos de dichos genes, reveló 1959 genes diferencialmente expresados. Las mismas variedades y bajo el mismo modelo experimental descrito, fueron evaluadas para su análisis epigenómico, por medio de la técnica de secuenciación bisulfito, la cual detectó 20157 genes diferencialmente metilados. Al interceptar ambas listas obtenidas, se obtuvieron 340 genes. Estos genes representan secuencias transcripcionalmente activas que también se caracterizan por estar diferencialmente metiladas.

Posteriormente, el análisis de correlación, contrastando e interceptando las dos listas generadas usando herramientas de análisis de datos como Microsoft Access, reveló 8 genes en donde, las diferencias en la expresión transcripcional del gen, coincidieron con un estado de metilación específico del promotor evaluado. Finalmente, los análisis de qPCR, los cuales cuantifican por técnicas de fluorescencia el nivel de expresión genes particulares, confirmaron al gen Os12g0137700 como un gen con expresión incrementada, cuyo alto nivel de expresión coincide con un patrón de hipo-metilación presente en su promotor.

Análisis complementarios de coexpresión y de ontologías, confirmaron la relación funcional del gen Os12g0137700 con procesos asociados a respuestas a estreses abióticos.

La figura 4, muestra el flujo metodológico seguido y los resultados obtenidos en cada fase del proceso al evaluar variedades contrastantes del arroz silvestre (*Oryza glumaepatula*). El diagrama de flujo muestra el número de genes retenido en cada fase de la estrategia de selección diseñada y describe como la implementación de los distintos filtros sucesivamente concatenados son útiles para reducir un amplio espectro de genes candidatos, a un único gen, pero de gran potencial para ser implementado en eventos de transformación agronómica.

En consecuencia, el método de identificación, caracterización y selección propuesto en la presente invención presenta la capacidad de filtrar genes novedosos y con potencial de implementación exitosa en procesos de mejoramiento agronómico ante condiciones de estrés por Al en *Oryza glumaepatula*.

EJEMPLO 3.

Aunque existe un amplio conocimiento de diferentes genes involucrados en la tolerancia al Al, aún existe gran incertidumbre sobre todos los elementos de la respuesta funcional que se activan en las plantas frente a las condiciones de toxicidad por el metal pesado.

El método objeto de divulgación en la presente descripción, a su vez, detectó al gen Os06g0486900 (*FDH2*) como un elemento genético relevante ante la respuesta a estrés por Al en la variedad comercial de arroz analizada. En contraste, ningún estudio previo ha determinado su potencial en la respuesta específica al estrés señalado, lo que determina la capacidad del método de selección de la invención para filtrar nuevos genes relevantes para la respuesta diferencial ante una respuesta fisiológica de interés.

El gen *FDH2* en *Oriza sativa* se encuentra en el cromosoma 6 y tiene una longitud de 1872 pares de bases. Este gen está compuesto por 7 exones y 6 intrones, codificando una proteína de 978 aminoácidos clasificada como “Formate dehydrogenase 2, mitochondrial”, la cual está involucrada en la oxidación del ion formato a dióxido de carbono, acoplado con la reducción de NAD^+ a NADH y ha sido clasificada como una proteína de respuesta a estrés en plantas (Alekseeva et al., 2011).

Los patrones de coexpresión pueden revelar redes de genes relacionados funcionalmente y proporcionar una comprensión más profunda de las funciones de los genes.

A partir de este concepto, se realizó una búsqueda del vecindario de coexpresión del gen *FDH2*, utilizando la base de datos de coexpresión RiceFRIEND (Sato et al., 2012), una plataforma para recuperar redes de genes coexpresados específicamente en arroz. Como resultado de este análisis, en la Tabla 2 se reportan los 10 genes que tienen un mayor nivel de coexpresión, evaluado por medio del coeficiente de correlación de Pearson (PCC), con el gen *FDH2*, entre los cuales se encuentran genes previamente involucrados en la respuesta a estrés en plantas.

Por ejemplo, el gen *Os06g0627500* se encontró dentro de un grupo de 14 QTL que están relacionados con la resistencia al hongo *Gibberella fujikuroi* en arroz, el cual genera la enfermedad conocida como enfermedad de Bakanae. Esta enfermedad genera que el metabolismo de la planta aumente, produciendo un excedente de ácido giberélico. En la planta, esto actúa como una hormona de crecimiento, causando hipertrofia. Las plantas afectadas, que están visiblemente marchitas y cloróticas. Plantas con una resistencia media al hongo son infértiles con panículas vacías y no producen granos comestibles; mientras que plantas susceptibles son incapaces de soportar su propio peso, se vuelcan y mueren (Chen et al., 2019).

Tabla 2. Genes coexpresados con el gen *FDH2*.

Locus	Descripción	Ranking Mutuo	PCC Ponderado
Os06g0486900	Similar to Formate dehydrogenase, mitochondrial precursor (EC 1.2.1.2) (NAD- dependent formate dehydrogenase) (FDH).	1	1.000.000
Os05g0570200	Protein of unknown function DUF295 family protein.	4.899	0.517146
Os10g0499200	Conserved hypothetical protein.	11.662	0.508048
Os05g0468400	Galactose oxidase, central domain containing protein.	19.442	0.539461
Os08g0112900	Lipolytic enzyme, G-D-S-L family protein.	21.166	0.477365
Os10g0499500	Similar to Fertility restorer.	31.289	0.469257
Os05g0140000	Conserved hypothetical protein.	43.955	0.480168
Os12g0625200	Conserved hypothetical protein.	45.365	0.488820
Os06g0627500	Leucine-rich repeat, plant specific containing protein.	47.497	0.439053
Os08g0419100	Conserved hypothetical protein.	50.596	0.386920
*Vecindario de coexpresión con los 10 genes con mayor puntuación identificados por la base de datos de coexpresión Ricefrend.			

El vecindario de coexpresión calculado mediante RiceFRIEND, (Sato et al., 2012), esto es, una plataforma para recuperar redes de genes coexpresados específicamente en arroz se utilizó como insumo para la realización de un análisis de enriquecimiento funcional en GeneOntology (GO), herramienta que permite asociar procesos biológicos, funciones moleculares y compartimentos celulares a un gen particular de interés. Dicha búsqueda evidenció que el gen *FDH2* está altamente coexpresado con genes que pertenecen a categorías funcionales involucradas con la respuesta a estrés en plantas (Tabla 3).

Por ejemplo, se reporta la categoría de “metabolismo de ácidos orgánicos”, los cuales desempeñan un papel importante en procesos de estrés en plantas, como el estrés por metales pesados, deficiencia de nutrientes e interacciones planta-microorganismo (Panchal et al., 2021; López-Bucio et al., 2000).

Así mismo, se encontraron términos asociados con biosíntesis de aminoácidos de la familia de las serinas, biosíntesis de compuestos nitrogenados y metabolismo de aminoácidos. En este sentido, las serinas pueden estar relacionadas con procesos de estrés en las plantas. Por ejemplo, se encontró que la actividad de la serina acetil transferasa (SAT) puede verse afectada por tratamientos con Al. En particular, las plantas sometidas al tratamiento con aluminio mostraron un aumento en los valores de actividad SAT en comparación con el control, lo que sugiere que el aluminio puede aumentar la actividad de SAT en una pequeña proporción. Además, el Al tiene una gran afinidad por el grupo por fosfolípidos de membrana tipo fosfatidilserina, lo que aumenta la rigidez de la membrana, la fusión y el empaquetamiento de la cadena de ácidos grasos, lo que respalda la hipótesis del reordenamiento de la membrana inducido por el Al (Matsumoto, H., 2000).

Existen evidencias adicionales que sugiere una posible asociación entre la serina y el estrés por Al. En un estudio realizado por Navascues et al. (2012) en *Lotus corniculatus*, planta que, al ser sometida a estrés por Al, tuvo una disminución en el contenido de glicina en la raíz y un aumento en el contenido de serina, ácido aspártico y ácido glutámico en las hojas. Esto sugiere que el Al puede afectar tanto el metabolismo de los aminoácidos como el recambio de proteínas en el brote de la planta.

Tabla 3. Análisis de enriquecimiento funcional GO. Procesos potencialmente asociados con la expresión de *FDH2*-Os06g0627500.

Identificador GO	Término GO	VALOR P	FDR		Positivo	Negativo	% de positivos
GO:0006564	L-serine biosynthesis	0.0006617161	0.008878024	Control	17	27184	0.1 %
				Caso	1	0	100%

<u>GO:0009070</u>	serine family amino acid biosynthesis	0.001213146	0.008878024	Control	32	27169	0.1 %
				Caso	1	0	100%
<u>GO:0006563</u>	L-serine metabolism	0.001396956	0.008878024	Control	37	27164	0.1 %
				Caso	1	0	100%
<u>GO:0009069</u>	serine family amino acid metabolism	0.001691052	0.008878024	Control	45	27156	0.2 %
				Caso	1	0	100%
<u>GO:0008652</u>	amino acid biosynthesis	0.004484964	0.01466804	Control	121	27080	0.4 %
				Caso	1	0	100%
<u>GO:0009309</u>	amine biosynthesis	0.004889346	0.01466804	Control	132	27069	0.5 %
				Caso	1	0	100%
<u>GO:0044271</u>	nitrogen compound biosynthesis	0.004889346	0.01466804	Control	132	27069	0.5 %
				Caso	1	0	100%
<u>GO:0006520</u>	amino acid metabolism	0.009594883	0.02167544	Control	260	26941	1%
				Caso	1	0	100%
<u>GO:0006519</u>	amino acid and derivative metabolism	0.01080803	0.02167544	Control	293	26908	1.1 %
				Caso	1	0	100%
<u>GO:0009308</u>	amine metabolism	0.01106536	0.02167544	Control	300	26901	1.1 %
				Caso	1	0	100%
<u>GO:0006807</u>	nitrogen compound metabolism	0.01216822	0.02167544	Control	330	26871	1.2 %
				Caso	1	0	100%
<u>GO:0019752</u>	carboxylic acid metabolism	0.01338137	0.02167544	Control	363	26838	1.3 %
				Caso	1	0	100%
<u>GO:0006082</u>	organic acid metabolism	0.01341813	0.02167544	Control	364	26837	1.3 %
				Caso	1	0	100%

Finalmente, existen varias evidencias de la relación del gen *FDH2* con la respuesta a estrés en plantas. Alekseeva y colaboradores (2011), reportaron que la transcripción de *FDH2* aumenta bajo condiciones de estrés por sequía, temperatura, irradiación con luz ultravioleta, químicos, hipoxia y estrés biótico. Así mismo, este gen ha sido asociado con la resistencia al “cuello de vaina”, una enfermedad producida por el patógeno *Rhizoctonia solani*, ya que, tras infectar la planta con el patógeno, hubo una sobreexpresión del gen, lo que sugiere que puede tener un valor para el mejoramiento genético del arroz en relación a estreses bióticos (Shi et al., 2020). Por lo tanto, la relación del gen *FDH2* con respuesta a otros tipos de estrés, refuerza la posible participación de este gen en procesos de regulación frente a condiciones de estrés por AI en variedades comerciales de arroz.

En consecuencia, el método de identificación, caracterización y selección propuesto en la presente invención presenta la capacidad de detectar al gen *FDH2* en *Oryza sativa*, como un gen relevante en los mecanismos de tolerancia al estrés por AI, siendo avalado por análisis de coexpresión y de enriquecimiento funcional

EJEMPLO 4.

Aunque existe un amplio conocimiento de diferentes genes involucrados en la tolerancia al Al, aún existe gran incertidumbre sobre todos los elementos de la respuesta funcional que se activan en las plantas frente a las condiciones de toxicidad por el metal pesado.

El método objeto de divulgación en la presente descripción, a su vez, detectó al gen *Os12g0137700* como un elemento genético relevante ante la respuesta a estrés por Al en la variedad silvestre analizada. En contraste, ningún estudio previo determina su potencial en la respuesta específica al estrés señalado, lo que determina la capacidad del método de selección para filtrar nuevos genes relevantes para la respuesta diferencial ante una respuesta fisiológica de interés.

El gen *Os12g0137700*, que se localiza en el cromosoma 12 de la especie *Oryza glumaepatula* y tiene una longitud de 1256 pb. Este gen está constituido por un solo exón y codifica una proteína de 337 aminoácidos que pertenece a la familia de las proteínas sulfotransferasas. Estas proteínas pueden desempeñar un papel crucial en el reconocimiento intermolecular y en los procesos de señalización en plantas (Varin et al., en 1997).

De la misma manera como se hizo para el gen *FDH2*, se realizó una búsqueda del vecindario de coexpresión del gen *Os12g0137700*, utilizando la base de datos RiceFRIEND (Sato et al., 2012). Como resultado de este análisis, en la Tabla 4 se reportan los 10 genes con mayor nivel de coexpresión según el PCC calculado para el gen *Os12g0137700*.

Dentro del vecindario de coexpresión de *Os12g0137700*, se reporta el gen *Os04g0604900* que, según el trabajo de Wang et al., (2020) se encuentra subexpresado en experimentos de deficiencia de nitrógeno. Así mismo, se encontró el gen *Os12g0259800* perteneciente al grupo funcional de lacasa (LACs), que suelen ser inducidos por estrés asociado a respuestas a hormonas, sal y metales pesados (Liu et al., 2017). Finalmente, se reporta el gen *Os05g0213900* que presentó subexpresión en una especie silvestre de arroz que presenta tolerancia al cadmio, siendo este un gen relacionado con el clivaje de pectinas, que, al estar presentes en la pared celular, pueden influir en la acumulación de cadmio (Zhang et al., s.f.). La existencia de genes de respuesta a estrés en el vecindario de coexpresión de *Os12g0137700* refuerzan la posible participación de este gen en procesos de regulación frente a condiciones de estrés por Al en plantas de arroz.

Tabla 4. Genes coexpresados con el gen *Os12g0137700*.

Locus	Descripción	Ranking Mutuo	PCC Ponderado
Os12g0137700	Sulfotransferase family protein.	1	1.000.000
Os04g0604900	Similar to Xyloglucan endotransglycosylase (Fragment).	6.928	0.597739
Os01g0934700	Esterase/lipase/thioesterase domain containing protein.	10.909	0.596904
Os11g0557800	Hypothetical protein.	13.856	0.673373
Os01g0957600	Similar to Elicitor-inducible cytochrome P450.	17.321	0.642288
Os11g0501000	Conserved hypothetical protein.	17.664	0.619349
Os12g0259800	Cupredoxin domain containing protein.	18.868	0.644280
Os11g0637200	Similar to Sorbitol transporter.	20.125	0.601399
Os05g0213900	Virulence factor, pectin lyase fold family protein.	23.431	0.651571
Os11g0138900	Esterase/lipase/thioesterase domain containing protein.	23.812	0.622132
*Vecindario de coexpresión con los 10 genes con mayor puntuación identificados por la base de datos de coexpresión RiceFRIEND.			

El vecindario de coexpresión calculado mediante RiceFRIEND, se utilizó como insumo para la realización de un análisis de enriquecimiento funcional en GeneOntology (GO). Dicha búsqueda evidenció que el gen *Os12g0137700* está altamente coexpresado con genes que se asocian funcionalmente a las categorías de “Respuesta a estímulo”, “Respuesta al estímulo químico”, “Respuesta al estímulo abiótico” y “Respuesta al estrés” (tabla 5). Este resultado demuestra una posible relación del gen *Os12g0137700* en procesos de regulación frente a condiciones de estrés por AI en plantas de arroz, ya que, concuerda con los resultados obtenidos por Cohen & Leach (2019), quienes realizaron un metaanálisis de datos transcriptómicos provenientes de diversos tipos de estrés tanto biótico como

abiótico y reportaron categorías funcionales enriquecidas asociadas a estímulos extracelulares, externos y endógenos.

Adicionalmente, se encontraron categorías funcionales asociadas con la pared celular, lo cual es esperado, pues su biosíntesis y organización, hacen parte de procesos que permiten que las células mantengan su forma y se protejan de la lisis osmótica ocasionada por factores abióticos externos (Attwood et al., 2006).

Además, se ha reportado previamente que la interacción del Al con la pared celular genera modificaciones en propiedades como plasticidad o elasticidad para modular la afinidad de los iones Al^{3+} e impactar los niveles de toxicidad (Kochian et al., 2015).

Por otra parte, el gen *Os12g0137700* está altamente coexpresado con genes asociados a la categoría funcional de “Respuesta a estrés oxidativo”, refiriéndose a la habilidad de las células a resistir de una mejor manera frente a agentes tóxicos, como por ejemplo los iones Al^{3+} que induce la producción de especies reactivas de oxígeno lo que, a su vez, genera una cascada de efectos deletéreos en las plantas, tales como disfuncionalidad mitocondrial y daños irreparables en el ADN (Yamamoto, 2002; Ali et al., 2019). La respuesta a estrés oxidativo generada por exposición a Al ha sido reportada previamente en plantas de té (Li et al., 2017), tabaco (Panda et al., 2012), maíz (Maron et al., 2008) y arroz (Tsutsui et al., 2012). La existencia de genes de respuesta a estrés en el vecindario de coexpresión del gen *Os12g0137700* refuerza su potencial participación en procesos de regulación frente a condiciones de estrés por Al en plantas de arroz silvestre.

Tabla 5. Análisis de enriquecimiento funcional GO. Procesos potencialmente asociados con la expresión de *Os12g0137700*.

Identificador GO	Término GO		Valor P	FDR		Positivo	Negativo	% de positivos
GO:0006979	response to oxidative stress		3,00E-07	4,62E-05	Control	146	27055	0.5 %
					Caso	18	282	6%
GO:0006800	oxygen and reactive oxygen species metabolism		9,23E-07	7,11E-05	Control	157	27044	0.6 %
					Caso	18	282	6%
GO:0042221	response to chemical stimulus		9,88E-06	5,07E-05	Control	210	26991	0.8 %
					Caso	19	281	6.3 %
GO:0009628			1,09E-04	4,19E-03	Control	244	26957	0.9 %

	response to abiotic stimulus			Caso	19	281	6.3 %
GO:0050896	response to stimulus	1,34E-03	4,13E-02	Control	637	26564	2.3 %
				Caso	28	272	9.3 %
GO:0006950	response to stress	9,84E-03	2,53E-01	Control	290	26911	1.1 %
				Caso	18	282	6%
GO:0045229	external encapsulating structure organization and biogenesis	3,79E+01	0.000729 0882	Control	77	27124	0.3 %
				Caso	7	293	2.3 %
GO:0007047	cell wall organization and biogenesis	3,79E+01	0.000729 0882	Control	77	27124	0.3 %
				Caso	7	293	2.3 %
GO:0007582	physiological process	0.000296 6766	0.005076 465	Control	8706	18495	32%
				Caso	125	175	41.7 %
GO:0050875	cellular physiological process	0.001333 005	0.020528 28	Control	7403	19798	27.2 %
				Caso	106	194	35.3 %
GO:0042546	cell wall biosynthesis	0.001477 619	0.020686 66	Control	18	27183	0.1 %
				Caso	3	297	1%
GO:0012501	programmed cell death	0.002180 042	0.025724 68	Control	298	26903	1.1 %
				Caso	10	290	3.3 %
GO:0006915	apoptosis	0.002180 042	0.025724 68	Control	298	26903	1.1 %
				Caso	10	290	3.3 %
GO:0009987	cellular process	0.002387 511	0.025724 68	Control	7528	19673	27.7 %
				Caso	106	194	35.3 %
GO:0006952	defense response	0.002862 36	0.025724 68	Control	259	26942	1%
				Caso	9	291	3%
GO:0016265	death	0.002999 074	0.025724 68	Control	312	26889	1.1 %
				Caso	10	290	3.3 %
GO:0008219	cell death			Control	312	26889	1.1 %

		0.00299 9074	0.02572 468	Caso	10	290	3.3 %
GO:0009607	response to biotic stimulus	0.00300 678	0.02572 468	Control	261	26940	1%
				Caso	9	291	3%
GO:0009664	cell wall organization and biogenesis (sensu Magnolioph yta)	0.00344 0348	0.02788 493	Control	25	27176	0.1 %
				Caso	3	297	1%

En consecuencia, el método de identificación, caracterización y selección propuesto en la presente invención presenta la capacidad de detectar al gen *Os12g0137700* en *Oryza glumaepatula*, como un gen relevante en los mecanismos de tolerancia al estrés por AI, siendo avalado por análisis de coexpresión y de enriquecimiento funcional.

Los resultados obtenidos sobre variedades tolerantes y susceptibles, tanto del arroz cultivado (*Oryza sativa*), como del arroz silvestre (*Oriza glumaepatula*) demuestran la capacidad del método propuesto para obtener genes candidatos, incluso, al analizar especies distintas. Lo anterior, evidencia el potencial del método objeto de la presente invención para ser implementado en una gran variedad de especies y en distintos procesos biológicos y fisiológicos de interés para la investigación y la industria.

Aunque la presente invención ha quedado descrita con las realizaciones preferentes mostradas, queda entendido que las modificaciones y variaciones que conserven el espíritu y el alcance de esta invención se entienden dentro del alcance de las reivindicaciones adjuntas.

Referencias.

- Alekseeva, A., Savin, S., & Tishkov, V. (2011). NAD⁺-dependent Formate Dehydrogenase from plants. *Acta Naturae*, 3(4), 38-54.
- 5 Ali, M. A., Fahad, S., Haider, I., Ahmed, N., Ahmad, S., Hussain, S., & Arshad, M. (2019). Oxidative Stress and Antioxidant Defense in Plants Exposed to Metal/Metalloid Toxicity. *Reactive Oxygen, Nitrogen and Sulfur Species in Plants*, 353-370. <https://doi.org/10.1002/9781119468677.ch15>
- 10 Attwood, T., Campbell, P., Parish, J., Smith, A., Stirling, J., & Vella, F. (2006). OXFORD DICTIONARY OF Biochemistry and Molecular Biology. In *Trends in Biochemical Sciences* (Vol. 23, Issue 12). [https://doi.org/10.1016/S0968-0004\(98\)01321-8](https://doi.org/10.1016/S0968-0004(98)01321-8)
- 15 Camille D. Ryan, Andrew J. Schaul, Ryan Butner, John T. Swarthout, 2020. Monetizing disinformation in the attention economy: The case of genetically modified organisms (GMOs). *European Management Journal* (38)1. ISSN 0263-237. <https://doi.org/10.1016/j.emj.2019.11.002>.
- Chen, S. Y., Lai, M. H., Tung, C. W., Wu, D. H., Chang, F. Y., Lin, T. C., & Chung, C. L. (2019). Genome-wide association mapping of gene loci affecting disease resistance in the rice-Fusarium fujikuroi pathosystem. *Rice*, 12, 1-12.
- 20 Cohen, S. P., & Leach, J. E. (2019). Abiotic and biotic stresses induce a core transcriptome response in rice. *Scientific Reports*, 9(1), 6273. <https://doi.org/10.1038/s41598-019-42731-8>
- 25 Li, Y., Huang, J., Song, X., Zhang, Z., Jiang, Y., Zhu, Y., Zhao, H., & Ni, D. (2017). An RNA-Seq transcriptome analysis revealing novel insights into aluminum tolerance and accumulation in tea plant. *Planta*, 246(1), 91-103. <https://doi.org/10.1007/s00425-017-2688-6>
- Liu, J., Zhou, M., Delhaize, E., & Ryan, P. R. (2017). Altered expression of a malate-permeable anion channel, OsALMT4, disrupts mineral nutrition. *Plant Physiology*, 175(4), 1745–1759. <https://doi.org/10.1104/pp.17.01142>
- 30 López-Bucio, J., Nieto-Jacobo, M., Ramírez-Rodríguez, V., & Herrera-Estrella, L. (2000). Organic acid metabolism in plants: from adaptive physiology to transgenic varieties for cultivation in extreme soils. *Plant Science*, 160(1), 1-19. [https://doi.org/10.1016/S0168-9452\(00\)00347-2](https://doi.org/10.1016/S0168-9452(00)00347-2)
- 35 Maron, L. G., Kirst, M., Mao, C., Milner, M. J., Menossi, M., & Kochian, L. V. (2008). Transcriptional profiling of aluminum toxicity and tolerance responses in maize roots. *New Phytologist*, 179(1), 116–128. <https://doi.org/10.1111/j.1469-8137.2008.02440.x>
- 40 Navascués, J., Pérez-Rontomé, C., Sánchez, D. H., Staudinger, C., Wienkoop, S., Rellán-Álvarez, R., & Becana, M. (2011). Oxidative stress is a consequence, not a cause, of aluminum toxicity in the forage legume *Lotus corniculatus*. *New Phytologist*, 193(3), 625-636. <https://doi.org/10.1111/j.1469-8137.2011.03978.x>
- Panchal, P., Miller, A. J., & Giri, J. (2021). Organic acids: versatile stress-response roles in plants. *Journal of Experimental Botany*, 72(11), 4038-4052. <https://doi.org/10.1093/jxb/erab019>
- 45 Panda, S. K., Sahoo, L., Katsuhara, M., & Matsumoto, H. (2012). Overexpression of Alternative Oxidase Gene Confers Aluminum Tolerance by Altering the Respiratory

Capacity and the Response to Oxidative Stress in Tobacco Cells. *Molecular Biotechnology*, 54(2), 551-563. <https://doi.org/10.1007/s12033-012-9595-7>

Sato, Y., Namiki, N., Takehisa, H., Kamatsuki, K., Minami, H., Ikawa, H., Ohyanagi, H., Sugimoto, K., Itoh, J., Antonio, B. A., & Nagamura, Y. (2013). RiceFRIEND: a platform for retrieving coexpressed gene networks in rice. *Nucleic Acids Research*, 41(1), 1214–1221. <https://doi.org/10.1093/nar/gks1122>

Shi, W., Zhao, S. L., Liu, K., Sun, Y. B., Ni, Z. B., Zhang, G. Y., Tang, H. S., Zhu, J. W., Wan, B. J., Sun, H. Q., Dai, J. Y., Sun, M. F., Yan, G. H., Wang, A. M., & Zhu, G. Y. (2020). Comparison of leaf transcriptome in response to *Rhizoctonia solani* infection between resistant and susceptible rice cultivars. *BMC Genomics*, 21(1), 245. <https://doi.org/10.1186/s12864-020-6645-6>

Tsutsui, T., Yamaji, N., Huang, C. F., Motoyama, R., Nagamura, Y., & Ma, J. F. (2012). Comparative Genome-Wide Transcriptional Analysis of Al-Responsive Genes Reveals Novel Al Tolerance Mechanisms in Rice. *PLoS ONE*, 7(10), e48197. <https://doi.org/10.1371/journal.pone.0048197>

Varin, L., Marsolais, F., & Rouleau, M. (1997). Biochemistry and molecular biology of plant sulfotransferases. *The FASEB Journal*, 11(17), 517-525. <https://doi.org/10.1096/fasebj.11.7.9212075>

Wang, Q., Zhu, Y., Zou, X., Li, F., Zhang, J., Kang, Z., Li, X., Yin, C., & Lin, Y. (2020). Nitrogen Deficiency-Induced Decrease in Cytokinins Content Promotes Rice Seminal Root Growth by Promoting Root Meristem Cell Proliferation and Cell Elongation. *Cells*, 9(4), 916. <https://doi.org/10.3390/cells9040916>

Yamamoto, Y., Kobayashi, Y., Devi, S. R., Rikiishi, S., & Matsumoto, H. (2002). Aluminum toxicity is associated with mitochondrial dysfunction and the production of reactive oxygen species in plant cells. *Plant Physiology*, 128(1), 63–72. <https://doi.org/10.1104/pp.010417>.

Zhang, Z., Luo, J., and Guo, B., He, Y., Chun-Zhu, C., & Yang, Y., Physiological and Omics Identification of a Rice Germplasm with High Cadmium Uptake and Translocation. Available at SSRN: <https://ssrn.com/abstract=4226697> or <http://dx.doi.org/10.2139/ssrn.4226697>