

Análisis bioinformático de genes de función desconocida que modifican su expresión por cromato, en plantas de *Arabidopsis thaliana*

Santiago José Torres Ríos¹, Fátima Hernández Madrigal¹, Yazmín Carreón Abud¹ y Miguel Martínez Trujillo✉

¹Facultad de Biología, Universidad Michoacana de San Nicolás de Hidalgo. Ciudad Universitaria. C.P. 58030, Morelia, Michoacán, México.

Resumen

Los componentes de cromo son altamente tóxicos para las plantas y son perjudiciales para su crecimiento y desarrollo. En nuestro grupo de trabajo ha sido utilizada a *Arabidopsis thaliana* como modelo, y previamente se ha observado una estimulación del crecimiento y desarrollo, en concentraciones de 20 μM de cromato de potasio (Cr(VI)), así como una inhibición a partir de 100 μM . En este trabajo fue analizado el efecto de dos concentraciones de Cr(VI), 20 μM y 140 μM , en los cambios de la expresión en genes. Se hizo una selección de los genes que aparecen con función desconocida, con base en la intensidad de los cambios de expresión con respecto al control sin Cr(VI) y se determinaron los dominios conservados en cada una de las proteínas codificadas por estos genes. Los resultados obtenidos permitieron determinar que es posible encontrar información valiosa de los genes mediante herramientas bioinformáticas y que el cromato modifica selectivamente los cambios de expresión de genes en plantas dependiendo de la concentración utilizada, orientando la expresión a la adaptación para las condiciones generadas por el cromato.

Palabras clave: Cromato, genes, expresión.

Abstract

The chromium compounds are highly toxic to plants and are detrimental to their growth and development. In our working group *Arabidopsis thaliana* has been used to as a model, and it has been previously observed stimulation of growth and development in concentrations of 20 μM potassium chromate (Cr(VI)) and inhibition since 100 μM . In this paper it was analyzed the effect of two concentrations of Cr(VI), 20 μM and 140 μM , on gene expression changes. A selection of genes with unknown function was performed based on the intensity of expression change relative to the control without the metal. The conserved domains in each of the proteins encoded by the genes, were determined. Based on these domains, biological information of the potential functions of the genes were generated. The results indicate that it is possible to find valuable information of genes using bioinformatic tools and that chromate changes gene expression in plants depending on the concentration used, directing expression for adapting to the conditions generated by the chromate

Key words: chromate, genes, expression.

Introducción

El cromo (Cr) en un elemento natural y ubicuo, que se encuentra en rocas, plantas, suelos, animales, y en los humos y gases volcánicos. El incremento de las concentraciones de Cr en el ambiente debido a actividades antropogénicas ha hecho necesario entender cómo interacciona con los organismos para dar posibles soluciones a su manejo.

Los efectos biológicos del cromo dependen de su estado de oxidación y localización celular. El Cr(VI) es considerado la forma más tóxica de cromo, mientras que el Cr(III) es relativamente inocuo debido a su insolubilidad y subsecuente inhabilidad para cruzar la membrana celular (Katz y Salem, 1993). La reducción del Cr(VI) a Cr(III) ha sido reportada en muchos sistemas biológicos, involucrando la formación intermedia de Cr(V), proceso en el que se generan radicales libres que pueden causar alteraciones al ADN, entre otros efectos tóxicos (Kawanishi *et al.*, 1986).

El Cr(VI) afecta el crecimiento en plantas e induce un número de síntomas tóxicos tal como la inhibición de la germinación de la semilla, disminución o inhibición del crecimiento de la raíz primaria, clorosis y modificaciones de la arquitectura de la raíz por la inducción de raíces laterales (Ortiz-Castro *et al.*, 2007; Shanker *et al.*, 2005).

En nuestro grupo de trabajo hemos utilizado a *Arabidopsis thaliana* como modelo y se ha observado una estimulación del crecimiento y desarrollo en concentraciones de cromato 20 μM , así como una inhibición a partir de 100 μM (Hernández-Madrigal, 2011). Además, hemos observado interacciones con los nutrientes, como fosfato y sulfato, que revierten el efecto tóxico del cromato en concentraciones de 1000 μM ó mayores (Ortiz-Castro *et al.*, 2007; Ortiz-Castro *et al.*, 2009; Hernández-Madrigal, 2011).

En nuestro grupo de trabajo, mediante la técnica de microarreglos, se ha probado el efecto de concentraciones de 20 μM y 140 μM de cromato de potasio en plantas germinadas directamente en medios con el metal y se han comparado los cambios de expresión de estas concentraciones con respecto a plantas crecidas sin Cr(VI). De los genes encontrados que modificaron su expresión se encuentran aquellos para los que se tiene una

✉ Autor de correspondencia: Miguel Martínez Trujillo. Facultad de Biología, Universidad Michoacana de San Nicolás de Hidalgo. Ciudad Universitaria. C.P. 58030, Morelia, Michoacán, México. Email: codigogenetico@gmail.com.

función conocida y otros para los cuales se desconoce su función (Hernández-Madrigal, 2011).

Como parte de la estructura terciaria de una proteína, es posible determinar regiones específicas conocidas como dominios, que tienen una estructura tridimensional característica y que pueden encontrarse en diferentes proteínas. Los dominios pueden ser considerados como unidades funcionales y/o estructurales de una proteína. Los dominios son a menudo seleccionados evolutivamente, ya que normalmente poseen una función prominente en la biología de la proteína a la que pertenecen. (Murzin *et al.*, 1995; Tatusov *et al.*, 2003; Letunic *et al.*, 2004).

Una de las bases de datos de dominios conservados es la CCD (Marchler-Bauer *et al.*, 2004), que comparte información con las bases SMART (Letunic *et al.*, 2004) y COG (Tatusov *et al.*, 2003). De las tres bases de datos mencionadas, CDD es la que se encuentra más depurada, con bloques conservados, separados de las secuencias no alineadas, lo que permite obtener resultados más confiables cuando se buscan dominios a partir de la secuencia de aminoácidos de una proteína (Marchler-Bauer *et al.*, 2004).

En el presente trabajo, se analizaron los genes de función desconocida que modificaron su expresión por efecto del cromato en plantas de *A. thaliana*.

Materiales y Métodos

Se utilizaron bases de datos y herramientas bioinformáticas disponibles en las direcciones electrónicas que se describen para cada caso:

- a) Para secuencias de nucleótidos y proteínas:

<http://www.ncbi.nlm.nih.gov>

<http://www.sanger.ac.uk>

- b) Para dominios conservados de proteínas:

<http://pfam.sanger.ac.uk>

<http://pfam.janelia.org>

<http://pfam.sbc.su.se>

<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=cdd>

<http://www.expasy.org/prosite/>

Los genes de *A. thaliana* que modificaron su expresión en concentraciones de 20 μM y 140 μM de Cr(VI) se obtuvieron de Hernández-Madrigal (2011).

Resultados y Discusión

Los cambios de expresión global de genes en *A. thaliana* en respuesta al cromato fue determinada previamente en nuestro laboratorio, en dos concentraciones diferentes, 20 μM y 140 μM (Hernández-Madrigal, 2011).

Se encontró una respuesta diferencial en los genes de función desconocida que respondieron a cambios en su expresión en medios con Cr(VI). Los genes inducidos y reprimidos son diferentes, lo que indica que la planta responde de manera diferente a este metal de acuerdo a su concentración. Considerando que el número de genes que se inducen o reprimen es numeroso, se seleccionaron aquellos que modificaron su expresión en concentración de 20 μM en más de 2 veces y los genes en concentración de 140 μM que modificaron su expresión en más de 3 veces. Los loci de los genes seleccionados se listan a continuación:

a) Genes que indujeron su expresión en 20 μM de cromato: At2g07110, At5g63820, At1g44478, At2g43255, At1g20870, At3g04780, At5g53905, At2g16730, At1g01450

b) Genes que reprimieron su expresión en 20 μM de cromato: At2g12905, At1g27800, At1g56660, At5g02670, At5g51360, At3g54550, At3g21430, At5g40450, At1g23060, At3g55790, At4g04680, At2g07739, At3g07060, At4g20250, At3g29010, At5g45090, At5g53020.

c) Genes que indujeron su expresión en 140 μM de cromato: At2g42610, At4g12735, At2g14560, At5g42530, At2g25510, At5g40155, At5g61080

d) Genes que reprimieron su expresión en 140 μM de cromato:

At1g47400, At2g14247, At4g19645, At1g47330, At2g12905, At3g45800, At4g31880, At3g44140, At5g12050, At3g03010, At5g49960, At3g56910, At1g35610, At5g17960, At1g36925, At1g02540, At3g14280, At1g73120, At3g29710, At3g19800, At4g26260, At4g06740, At1g52270, At5g15120, At1g19010, At2g22122, At3g05750, At1g77890.

Para cada uno de los genes seleccionados se buscó la secuencia de la proteína hipotética reportada. Para esto se utilizó el procedimiento de búsqueda, en el cual fue necesario localizar primeramente la secuencia perteneciente a cada gen para encontrar así la secuencia de las proteínas. Con la secuencia de la proteína se procedió a determinar los dominios conservados y la familia en la cual se agrupa. Para esto se procedió a buscar las secuencias en un BLAST (Basic Local Alignment Search Tool) de alguno de los servidores ya mencionados. Teniendo las secuencias de las proteínas se buscaron características e información que tuviera relación con alguna otra especie, considerando sus dominios conservados (CDD).

Análisis de genes que aumentaron su expresión en 20 μM de Cr(VI)

De los 9 genes analizados que inducen su expresión, no se encontró alguno que estuviera relacionado con el estrés oxidativo, lo que sugiere que la planta no detecta esta concentración del metal como tóxica. De hecho, en los resultados obtenidos por Hernández-Madrigal (2011) se encontró que a estas concentraciones el crecimiento tanto de la raíz como el follaje se estimula.

El gen At1g01450 codifica para una proteína con dominio de cinasa en aminoácidos de serina, treonina y tirosina. Esta proteína está altamente conservada en otras dicotiledóneas como la vid, el sorgo y la higuera, e incluso en el arroz (83%), una planta monocotiledónea. Esta conservación sugiere que la proteína tiene una función conservada en las plantas, posiblemente relacionada en procesos de transducción de señales en los que se requiere fosforilación.

El gen At3g04780 codifica para una proteína con un dominio que se ha reportado que interacciona con el proteosoma. Este complejo multiproteico está implicado en la degradación de proteínas, en un procedimiento que requiere la unión de ubiquitina y permite eliminar proteínas que en ese momento no se requieren, ya sea para iniciar un proceso fisiológico o inhibirlo. Resalta la alta conservación de esta proteína no sólo en el reino vegetal, sino también en animales, como es el caso de *Drosophila* (98%), lo que indica que la función de esta proteína es característica de al menos los eucariontes multicelulares.

El dominio ACD, presente en la proteína codificada por el

gen At1g20870, se reporta para proteínas de choque térmico para las que se sugiere un papel como chaperonas. Estas proteínas se encargan de mantener la conformación de otras proteínas en condiciones de estrés celular, para que conserven su función. En el caso del cromato utilizado a la concentración de 20 μM , el crecimiento de las plantas se favorece, lo que se puede relacionar con el hecho de que las plantas pueden controlar su respuesta, en parte porque las proteínas pueden seguir manteniendo su conformación por la actividad de chaperonas, como la que incrementó su expresión en este trabajo. La conservación de la proteína es alta en especies relacionadas, pero se reduce en otras dicotiledóneas.

La proteína codificada por At2g16730 tiene dominios implicados en la hidrólisis de carbohidratos. Considerando que a 20 μM de Cr(VI) las plantas aumentan su crecimiento, este incremento en la expresión le permite posiblemente tener disponibles los monosacáridos necesarios para mayor obtención de energía. Al parecer esta proteína tiene una función conservada en el reino vegetal, ya que las similitudes con las especies encontradas fue del 100% ó cercanas a este valor.

El dominio de RNasa PH encontrado en la proteína codificada por At2g07110, está implicado en el procesamiento de ARN transportadores. La inducción del gen conduce posiblemente a un procesamiento más rápido de estas moléculas, que se encargan de llevar los aminoácidos a los ribosomas. Si el crecimiento de las plantas es mayor a 20 μM de Cr(VI), esto implica una mayor síntesis de proteínas, lo que se favorece por la mayor disponibilidad de los ARN transportadores. La conservación de esta proteína con similitudes altas ocurre tanto en plantas como en animales, lo que es de esperarse, ya que la síntesis de proteínas es un proceso básico de los eucariontes.

La proteína codificada por el gen At1g44478 está poco conservada en el reino vegetal, e incluso con la especie del mismo género *A. lyrata* sólo tiene una similitud de 67%. Aunque se relaciona en la morfogénesis celular, en *A. thaliana* no se presenta el dominio RRM (para unión a ARN), lo que sugiere que la función en la especie estudiada posiblemente sea diferente.

En síntesis, es posible deducir que los genes que inducen su expresión en 20 μM de Cr(VI) están implicados en respuestas adaptativas que le permiten a la planta de *A. thaliana* tener un mejor crecimiento.

Análisis de genes que redujeron su expresión en 20 μM de Cr(VI)

Para varias de las proteínas codificadas por los genes analizados, no se encontraron dominios conservados. Es posible que éstas sean proteínas características de *A. thaliana* y especies muy relacionadas. Tal es el caso de la proteína codificada por At1g23060, que sólo se conserva en *A. lyrata*. Incluso At1g56660 codifica para una proteína no relacionada con *A. lyrata*.

Tres proteínas altamente conservadas son la codificadas por los genes At3g54550, At2g07739 y At3g54550, sin embargo no se sabe de alguna posible función de sus dominios.

La proteína codificada por At3g21430 tiene un dominio presente en represores transcripcionales, particularmente implicados en el proceso del ciclo celular. Considerando lo anterior y a que las plantas aumentan su crecimiento en 20 μM de Cr(VI), la disminución de un represor en el ciclo celular permite una mayor

producción de células y por lo tanto mayor crecimiento. Esta proteína está altamente conservada en angiospermas, lo que indica una función similar en las especies de estos grupos.

El dominio de la familia TIR encontrado en la proteína codificada por el gen At5g45090, participa en interacciones proteína-proteína en procesos de señalización, sin embargo, no es posible deducir cómo se relaciona este proceso tan general con la respuesta adaptativa de la planta.

Una proteína altamente conservada es la codificada por At5g53020, que contiene dominios de proteínas del poro nuclear. Estas estructuras permiten la salida de moléculas grandes del núcleo, como es el caso de los ARN e incluso las partículas ribosomales. Su disminución no está clara en el contexto de la respuesta de la planta a un mayor crecimiento.

En síntesis, la información proporcionada por los genes de función desconocida que redujeron su expresión en 20 μM de Cr(VI) permite resaltar el hecho de que se favorece posiblemente una mayor división celular.

Análisis de genes que aumentaron su expresión en 140 μM de Cr(VI)

Varias proteínas no generaron información que pudiera relacionarse con una simple función. Tal es el caso de la codificada por At2g25510, para la que no se encontró similitud con alguna otra proteína. La proteína codificada por At5g40155 está incluso poco conservada en *A. lyrata*, una especie muy relacionada. Aunque la proteína codificada por At4g12735 parece conservada en el reino vegetal, no tiene dominios reportados. La proteína codificada por At5g42530 está medianamente conservada en el reino vegetal, sin embargo, no hay antecedentes de posible función.

El gen At2g14560 codifica para una proteína con dominio implicado en la redistribución de fosfolípidos por respuesta al daño. Esta es una clara respuesta adaptativa de la planta, ya que el Cr(VI) en 140 μM disminuyó el crecimiento.

El dominio de RNasa H presente en la proteína codificada por At5g61080 se reporta para el procesamiento de los fragmentos de Okazaki, en la replicación de las moléculas de ADN. El aumento de esta proteína implica ajustes en el proceso de replicación y aunque las plantas redujeron su crecimiento en la concentración del Cr(VI), lo que implica menor división celular, no se vió incremento en la expresión de componentes de las ADN polimerasas, de acuerdo a lo analizado por Hernández-Madrigal (2011).

En síntesis, la inducción en la expresión de genes en Cr(VI) 140 μM implica ajustes en la distribución de fosfolípidos en membranas y en la replicación del ADN. Aunque en los genes de función desconocida no hubo expresión incrementada de genes implicados en daño oxidativo ó interacción con otros nutrientes, Hernández-Madrigal (2011) reportó resultados de genes de función conocida relacionados con estos aspectos. Por lo tanto, las respuestas de las plantas están orientadas a remediar el daño causado por el metal.

Análisis de genes que disminuyeron su expresión en 140 μM de Cr(VI)

Varios de los genes analizados no generaron información relevante. Tal fue el caso de At2g12905, At1g47400 y At2g14247

At4g19645, que aunque sus proteínas se conservan en otras plantas, no se reportan dominios que indiquen alguna función probable. La proteína codificada por At3g44140 no se encontró en otras plantas, por lo que posiblemente sea característica de esta especie o especies muy relacionadas.

Otros genes están relacionados con la síntesis de proteínas. La disminución de la expresión implica una menor síntesis de estas moléculas, lo cual se puede considerar como una respuesta del menor crecimiento de las plantas en el Cr(VI) a 140 μM . La proteína codificada por At3g03010 tiene dominios relacionados con la actividad de la separación del ARN transportador de la proteína que se está sintetizando, por lo que una menor actividad retrasaría la síntesis de las proteínas; además, se encuentra altamente conservada en eucariontes. Además, una proteína relacionada con el ribosoma del cloroplasto codificada por At3g56910, también ve disminuida su expresión.

La proteína codificada por At5g49960 tiene un dominio implicado en el transporte de iones inorgánicos, por lo que su disminución indica una alteración en los mecanismos de transporte de otros iones.

La afectación de una proteína con dominio transmembranal, codificada por At1g47330, implica una posible disminución en la interacción con algún ligando, afectándose por lo tanto un proceso de señalización.

En síntesis, la afectación de genes implicados en síntesis de proteínas es un reflejo del menor crecimiento de las plantas, que se acompaña de alteraciones en el transporte de iones y posibles receptores que modifican las señales celulares.

Análisis general de los cambios en la expresión de genes

Las plantas son organismos sésiles, por lo que los cambios ambientales no pueden ser evitados y forzosamente han desarrollado estrategias adaptativas que les han permitido sobrevivir hasta nuestros días (Taiz y Zeiger, 2006). Sin embargo, las actividades antropogénicas han generado cambios ambientales a un ritmo tan acelerado, que no ha permitido a muchas especies adaptarse y llegan a desaparecer de los ecosistemas. Entre estos cambios se encuentran las acumulaciones de compuestos de Cr, en cantidades en las que de manera natural no se encuentran en los ambientes naturales (Shanker *et al.*, 2005).

Para abordar los problemas ocasionados por metales como el Cr es necesario establecer estrategias que permitan entender el efecto de sus compuestos en la fisiología de los organismos. Cervantes *et al.* (2001) han demostrado como el Cr(VI) es internalizado en bacterias mediante transportadores de sulfato y que existen determinantes genéticos que permiten expulsar a estos compuestos a través de proteínas transportadoras. Sin embargo, en plantas el trabajo ha sido más escaso, aunque se ha relacionado la interacción de algunos nutrientes con el Cr(VI), así como una reducción en el contenido de proteínas (Sherwry y Peterson, 1974; Vajpayec, 1999).

Para entender el efecto del Cr(VI) en plantas nuestro grupo de trabajo inició un análisis de la expresión global de genes mediante la técnica de microarreglos. Los genes de función conocida se han analizando con anterioridad (Hernández-Madrugal, 2011). Considerando que al menos una tercera parte de los genes

de *A. thaliana* carecen de una función clara, en este trabajo se utilizaron herramientas bioinformáticas para generar información preliminar de los genes que modificaron significativamente su expresión en los medios con Cr(VI).

Los cambios de expresión de genes se valoraron para concentraciones de Cr(VI) de 20 μM y de 140 μM , ya que en estas condiciones se estimula o se inhibe el crecimiento de las plantas, respectivamente (Hernández-Madrugal, 2011). Los genes con función desconocida que modificaron su expresión fueron centenares en los tratamientos de 140 μM , por lo que fue necesario abordar sólo los que tuvieron mayores cambios.

Los resultados de los análisis en las secciones anteriores demostraron que en la condición de estimulación del crecimiento de plantas, los genes que se indujeron o reprimieron, están implicados en respuestas adaptativas para aumentar la síntesis de proteínas, mantener la conformación de éstas y en acelerar el ciclo celular. Por lo tanto, en estas condiciones la planta no sólo se sobrepone al efecto del metal, sino que además la respuesta permite generar plantas de crecimiento más acelerado, posiblemente para una mejor adaptación. Por el contrario, algunos de los genes que se indujeron o reprimieron en las condiciones de Cr(VI), en las que las plantas disminuyen su crecimiento, están relacionados con una menor síntesis de proteínas y en redistribución de lípidos de membrana, debido a los efectos tóxicos ocasionados por el metal.

Referencias

- Cervantes C, Campos-García J, Devars S, Gutiérrez-Corona F, Tavera HL, Torres-Guzmán JC, Moreno-Sánchez R. 2001. Interactions of chromium with microorganisms and plants. *FEMS Microbiol Rev* 25: 335-347.
- Hernández-Madrugal F. 2011. Efecto del cromo (VI) en la expresión global de genes en *Arabidopsis thaliana* L. Tesis de Maestría. UMSNH. Morelia, Michoacán, México.
- Katz SA, Salem H. 1993. The toxicology of chromium with respect to its chemical speciation: a review. *J Appl Toxicol* 13: 217-224.
- Kawanishi S, Inoue S, Sano S. 1986. Mechanism of DNA cleavage induced by sodium chromate (VI) in the presence of hydrogen peroxide. *J Biol Chem* 261: 5952-5958.
- Letunic I, Copley RR, Schmidt S, Ciccarelli FD, Doerks T, Schultz J, Ponting CP, Bork P. 2004. SMART 4.0: towards genomic data integration. *Nucleic Acids Res* 32: 142-144.
- Marchler-Bauer A, Anderson JB, *et al.* 2005. CDD: a Conserved Domain Database for proteína classification. *Nucleic Acids Research* 33: 192-196.
- Murzin AG, Brenner SE, Hubbard T. 1995. SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J Mol Biol* 247: 536-40.
- Ortiz-Castro R, Martínez-Trujillo M, López-Bucio J, Cervantes C, Dubrovsky J. 2007. Effects of dichromate on growth and root system architecture of *Arabidopsis thaliana* seedlings. *Plant Science* 172: 684-691.
- Ortiz-Castro R, Martínez-Trujillo M, López-Bucio J, Cervantes C, Carreón-Abud Y. 2009. Restauración del crecimiento radical por

nutrimentos inorgánicos en *Arabidopsis thaliana* L., expuesta a cromo. *Terra Latinoamericana* 27: 97-103.

Shanker AK, Cervantes C, Loza-Tavera H, Avudainayagam S. 2005. Chromium toxicity in plants. *Environ Int* 31: 739-753.

Shewry PR, Peterson JP. 1974. The uptake of chromium by barley seedlings (*Hordeum vulgare* L.). *J Exp Bot* 25: 785-797.

Taiz L, Zeiger E. 2006. Plant physiology, 6a Ed., *Sinauer Associates Inc. Usa.* 774 pp.

Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Kiryutin B, Koonin EV, Krylov DM, Mazumder R, Mekhedov SL, Nikolskaya AN, Rao BS, Smirnov S, Sverdlov AV, Vasudevan S, Wolf YI, Yin JJ, Natale DA. 2003. The COG database: and updated version includes eukaryotes. *BMC Bioinformatics* 4: 41-65.

Vajpayec P, Sharma SC, Tripathi RD, Rai, UN, Yunus M. 1999. Bioaccumulation of chromium and toxicity to photosynthetic pigments, nitrate reductase activity and protein content of *Nelumbo nucifera* Gaertn. *Chemosphere* 39: 2159-2169.