



Programa de Doctorado en Bioingeniería
Universidad Miguel Hernández de Elche

**Estudio de los roles de INCURVATA11
y CUPULIFORMIS2 como proteínas
accesorias del Polycomb Repressive
Complex 2 de Arabidopsis**

Riad Nadi

Director de la tesis
José Luis Micol Molina

Elche, 2024

JOSÉ LUIS MICOL MOLINA, Catedrático de Genética de la Universidad Miguel Hernández de Elche (UMH)

HAGO CONSTAR:

Que el presente trabajo ha sido realizado bajo mi dirección y recoge fielmente la labor desarrollada por el Ingeniero Riad Nadi para optar al grado de Doctor. Las investigaciones reflejadas en esta memoria se han desarrollado íntegramente en la Unidad de Genética del Instituto de Bioingeniería de la UMH, según los términos y condiciones definidos en el Plan de Investigación del doctorando, y cumpliendo los objetivos inicialmente previstos de forma satisfactoria y lo establecido en el Código de Buenas Prácticas de la UMH.

José Luis Micol Molina

Elche, 19 de febrero de 2024

II.- RESUMEN

Como primera parte de esta tesis se hizo una revisión de las publicaciones sobre la superfamilia de las dioxigenasas dependientes de 2-oxoglutarato (también denominado 2-cetoglutarato o α -cetoglutarato) y Fe^{2+} (2OGD), que son enzimas oxidativas cuyo sitio activo incluye dos histidinas y en la mayoría de los casos un residuo de ácido aspártico o glutámico. Este motivo conservado cataliza reacciones de desmetilación, desmetilación, hidroxilación, halogenación, desaturación, ruptura o cierre de anillos, y epimerización. El genoma de *Arabidopsis* contiene más de 150 genes que codifican proteínas 2OGD, a las que se ha clasificado en los clados DOXA, DOXB, DOXC y JMJ. Las proteínas DOXA son homólogas de la dioxigenasa dependiente de α -cetoglutarato AlkB (por Alkylation B) de *Escherichia coli*, una enzima de reparación del ADN que revierte las metilaciones de los átomos N^1 de la adenina y N^3 de la citosina causadas por agentes alquilantes. Las proteínas DOXC constituyen la clase más amplia y diversa de las 2OGD, actuando en facetas del metabolismo vegetal como la biosíntesis y/o el catabolismo de la auxina y los lignanos, isoprenoides, flavonoides, glucosinolatos, alcaloides, estrigolactonas y cumarinas, y tienen importantes papeles en la homeostasis del etileno, las giberelinas, la auxina y el ácido salicílico. Las proteínas JMJ contienen un dominio Jumonji C (JmjC) y catalizan la desmetilación por hidroxilación de lisinas en las histonas. Las proteínas DOXB animales juegan un papel crucial en la síntesis del colágeno, el componente estructural mayoritario del espacio extracelular de muchos tejidos. Algunas proteínas DOXB de las plantas modifican postraduccionalmente determinadas O-glicoproteínas ricas en hidroxiprolina, como las extensinas, que son similares al colágeno. La familia CUPULIFORMIS (CP) incluye cinco proteínas DOXB: INCURVATA11 (ICU11), CP2, CP3, CP4 y CP5. A diferencia de todas las DOXB estudiadas hasta ahora, ICU11 y CP2 son nucleoplásmicas y tienen funciones epigenéticas.

Antes del comienzo de esta tesis se habían descrito dos mutantes *incurvata11* (*icu11*), que manifiestan hiponastia foliar y floración temprana. Estos rasgos morfológicos también los causan las mutaciones en genes que codifican algunos componentes de la maquinaria epigenética, como *CURLY LEAF* (*CLF*), que pertenece al grupo Polycomb (PcG). Las proteínas del PcG forman parte de dos Polycomb Repressive Complexes (PRC) heteromultiméricos con distintas actividades epigenéticas: el PRC1 es una ligasa de ubiquitina de la histona H2A, y el PRC2, una metiltransferasa de la lisina 27 de la histona H3 (H3K27). El PRC1 de *Arabidopsis* tiene cinco componentes principales, y el PRC2, ocho, uno de los cuales es EMBRYONIC FLOWER 2 (EMF2). Estos PRC también cuentan con proteínas

acesorias, que facilitan la incorporación del complejo a determinadas regiones de la cromatina. Un ejemplo de ello es EMBRYONIC FLOWER 1 (EMF1), que contribuye al depósito de la marca epigenética H3K27me3 en un subgrupo de los genes diana del PRC2 y también participa en la monoubiquitinación de la H2A por el PRC1.

El fenotipo morfológico de los mutantes *icu11* es relativamente débil, y los *cp2* son indistinguibles del tipo silvestre. Sin embargo, los dobles mutantes *icu11 cp2* muestran un fenotipo letal postembrionario muy similar al de los mutantes simples *emf1* y *emf2* más extremos: no manifiestan desarrollo vegetativo y forman órganos parecidos a flores inmediatamente después de la germinación, a los que se denomina flores embrionarias. También forma flores embrionarias el triple mutante *telomere repeat binding1-2 (trb1-2) trb2-1 trb3-2*. Las proteínas TRB1, TRB2 y TRB3 se unen a las secuencias repetitivas teloméricas para el mantenimiento de los telómeros y además son proteínas accesorias del PRC2, al que reclutan a determinados genes para el depósito de la marca H3K27me3.

Col-0 es la estirpe silvestre de referencia y la de uso experimental más común en *Arabidopsis*. La redundancia funcional entre *ICU11* y *CP2* se infirió del estudio de los fenotipos sinérgicos de las combinaciones dobles mutantes y sesquimutantes de las mutaciones *icu11* y *cp2*, cuyos fondos genéticos eran S96 (*icu11-1*), Ws-2 (*icu11-2*) y Col-0 (*cp2-1*, *cp2-2* y *cp2-3*). En consecuencia, todas sus combinaciones genéticas tenían fondos genéticos híbridos S96/Col-0 o Ws-2/Col-0. Para evitar los eventuales efectos de los modificadores presentes en los fondos S-96 y Ws-2, la segunda parte de esta Tesis ha consistido en la obtención y caracterización de cuatro nuevos alelos de *ICU11*, mutagenizando las estirpes silvestres S96 (*icu11-4* e *icu11-7*) y Col-0 (*icu11-5* e *icu11-6*) mediante la tecnología CRISPR/Cas9. Los alelos *icu11-5* e *icu11-6* son aparentemente nulos, ya que presentan pequeñas deleciones que desfasan la pauta de lectura del primer exón del gen *ICU11* y originan un codón de terminación prematura, y no parecen haber sufrido mutaciones adicionales indeseadas. También manifiestan menor hiponastia foliar que *icu11-1*, aunque similares floración temprana e interacciones genéticas con los alelos mutantes de *CP2*. Hemos usado las mutaciones *icu11-5* e *icu11-6* para confirmar que la ausencia simultánea de las proteínas ICU11 y CP2 es letal y que los fenotipos de los dobles mutantes y sesquimutantes *icu11 cp2* son independientes de su fondo genético y no manifiestan especificidad de alelo.

En nuestro estudio de los mutantes *icu11-5 cp2-1* y *emf2-3* hemos establecido que su letalidad puede ser paliada si se cultivan en medio suplementado con un 3% en sacarosa, en lugar del 1% habitual, lo que a su vez permite estudiar su fenotipo morfológico a lo largo de todo su ciclo de vida. En esta condición de cultivo, las plantas *icu11-5 cp2-1* y *emf2-3* desarrollan tallos, hojas caulinares pequeñas, flores que manifiestan transformaciones

homeóticas de sépalos y pétalos en carpelos, y silicuas que rinden semillas que solo en algunos casos resultan viables. Estas observaciones indican que la letalidad de las flores embrionarias *icu11 cp2* y *emf2-3* se debe a su escasa capacidad fotosintética, derivada de su carencia de hojas vegetativas, y que ICU11 y/o CP2 se requieren para la especificación de la identidad de los órganos florales.

En la tercera parte de esta Tesis se ha intentado dilucidar la función de ICU11 y CP2 mediante análisis interactómicos y transcriptómicos. Durante el transcurso de esta Tesis, otros autores publicaron un artículo en el que se describía un tercer alelo de *ICU11* (*icu11-3*) y se demostraba, mediante un ensayo de coimmunoprecipitación, que la proteína ICU11 interacciona con varios componentes principales del PRC2. Hemos realizado escrutinios de interactores de las proteínas ICU11 y CP2 mediante purificación por afinidad en tándem. Hemos confirmado así que ICU11 interacciona con los componentes principales del PRC2 EMF2, FERTILIZATION INDEPENDENT ENDOSPERM (FIE), SWINGER (SWN) y MULTICOPY SUPPRESSOR OF IRA1 (MSI1) y con las proteínas accesorias de este complejo EMF1, TRB1, TRB2 y TRB3, y demostrado que también lo hace con otras proteínas nucleares. CP2 no presentó interacciones con componentes principales o proteínas accesorias del PRC2, aunque sí lo hizo con TRB4 y TRB5, miembros poco caracterizados de la familia TRB, y otras proteínas nucleares.

También hemos realizado ensayos de complementación de fluorescencia bimolecular mediante transformación transitoria de hojas de *Nicotiana benthamiana*, en los que tanto ICU11 como CP2 interaccionaron con los componentes principales del PRC2 SWN y CLF, y con sus proteínas accesorias TRB1 y TRB3. No hemos encontrado interacción alguna entre ICU11 y CP2, lo que indica que no forman heteromultímeros.

Hemos llevado a cabo mediante secuenciación masiva de ARN un análisis comparativo de los perfiles transcriptómicos de plántulas de Col-0, *cp2-1* e *icu11-5*, flores embrionarias de *icu11-5 cp2-1* y *emf2-3*, e inflorescencias de Col-0. Este análisis ha revelado la gran semejanza entre los perfiles del doble mutante *icu11-5 cp2-1* y el mutante simple *emf2-3*, así como con los de otros mutantes portadores de alelos de genes que codifican componentes principales del PRC2. Muchos de los genes desregulados en las flores embrionarias *icu11-5 cp2-1* son portadores de la marca represora H3K27me3 en Col-0.

Considerados en conjunto, nuestros resultados confirman que ICU11 es una proteína accesoria del PRC2, revelan que muy probablemente CP2 también lo sea, y aportan nuevos indicios de su relación funcional y de sus funciones epigenéticas parcialmente solapantes.