




Universitat Autònoma de Barcelona

**ADVERTIMENT.** L'accés als continguts d'aquesta tesi queda condicionat a l'acceptació de les condicions d'ús establertes per la següent llicència Creative Commons:  [http://cat.creativecommons.org/?page\\_id=184](http://cat.creativecommons.org/?page_id=184)

**ADVERTENCIA.** El acceso a los contenidos de esta tesis queda condicionado a la aceptación de las condiciones de uso establecidas por la siguiente licencia Creative Commons:  <http://es.creativecommons.org/blog/licencias/>

**WARNING.** The access to the contents of this doctoral thesis it is limited to the acceptance of the use conditions set by the following Creative Commons license:  <https://creativecommons.org/licenses/?lang=en>



**Universitat Autònoma  
de Barcelona**

Departament de Biologia Cel·lular, de Fisiologia i d'Immunologia  
Facultat de Biociències

Memoria presentada por Adrián Tirado Herranz para optar al grado de  
Doctor en Inmunología en el Programa de Doctorado en Inmunología  
Avanzada de la  
Universidad Autónoma de Barcelona

**ANÁLISIS PROTEÓMICO Y BIOQUÍMICO  
DEL REGULADOR AUTOINMUNE:  
NUEVAS FUNCIONALIDADES  
Y VÍAS DE DEGRADACIÓN DE AIRE**

Tesis realizada en el Laboratori d'Immunologia Cel·lular  
del Institut de Biotecnologia i Biomedicina  
de la Universitat Autònoma de Barcelona  
bajo la direcció del Dr. Iñaki Álvarez

**Adrián Tirado Herranz**

Doctorando

**Iñaki Álvarez Pérez**

Director

Diciembre de 2022





# CONTENIDOS

<b>CONTENIDOS</b> .....	<b>11</b>
<b>ABREVIATURAS</b> .....	<b>17</b>
<b>RESUMEN</b> .....	<b>21</b>
<b>INTRODUCCIÓN</b> .....	<b>25</b>
<b>I. Tolerancia y Educación Tímica</b> .....	<b>25</b>
1.1 Células Epitiales Tímicas (TECs) .....	27
1.2 Selección Tímica .....	31
<b>II. Regulador Autoinmune (AIRE)</b> .....	<b>32</b>
2.1 Características bioquímicas y biofísicas de AIRE .....	34
2.2 Mutaciones en el gen AIRE .....	37
2.3 Mecanismo molecular de acción .....	39
2.4 AIRE en el ciclo celular .....	41
2.5 AIRE en periferia .....	43
<b>III. El Complejo Principal de Histocompatibilidad (MHC)</b> .....	<b>44</b>
3.1 Estructura de las moléculas de MHC .....	46
3.2 Presentación antigénica (Interacción MHC-péptido) .....	48
3.3 Vía de procesamiento antigénico .....	49
3.3.1 Vía de procesamiento antigénico de clase I .....	49
3.3.2 Vía de procesamiento antigénico de clase II .....	50
<b>IV. El Sistema Ubiquitín Proteasoma (UPS)</b> .....	<b>52</b>
<b>V. La familia E3 Ubiquitín Ligasas Seven in Absentia Homolog (SIAH)</b> .....	<b>56</b>
<b>OBJETIVOS</b> .....	<b>61</b>
<b>MATERIALES Y MÉTODOS</b> .....	<b>65</b>
<b>I. Cultivos celulares y vectores</b> .....	<b>65</b>
1.1 Líneas celulares de origen no linfoide .....	65
1.2 Condiciones de cultivo .....	65
1.3 Construcciones genéticas .....	66
1.4 Transfecciones .....	67
1.4.1 Transfección transitoria .....	67
1.4.2 Transfección estable y clonación .....	68
<b>II. Proteómica y peptidómica</b> .....	<b>69</b>
2.1 Proteómica cuantitativa (iTRAQ).....	69
2.1.1 Digestión proteica y marcaje con el reactivo iTRAQ-8-plex® .....	69

2.1.2 Cromatografía líquida y análisis por espectrometría de masas.....	70
2.1.3 Análisis de datos y estadística .....	70
2.2 Inmunoprecipitación y purificación de péptidos asociados a moléculas de HLA .....	71
2.2.1 Análisis por Espectrometría de masas .....	72
2.2.2 MS/MS Ion Search and Peptide identification .....	73
2.3 Análisis de proteínas y péptidos.....	73
<b>III. Anticuerpos .....</b>	<b>74</b>
3.1 Anticuerpos primarios.....	74
3.2 Anticuerpos secundarios .....	74
<b>IV. Electroforesis y Western blot .....</b>	<b>75</b>
<b>V. Citometría de flujo .....</b>	<b>75</b>
5.1 Tinción intracelular .....	76
5.2 Tinción celular de superficie .....	77
5.3 Ensayo de apoptosis .....	77
5.4 Análisis del ciclo celular.....	78
<b>VI. Microscopía .....</b>	<b>78</b>
6.1 Inmunofluorescencias.....	79
6.2 Ensayo de migración (Woung Healing Assay) .....	79
<b>VII. Inmunoprecipitación.....</b>	<b>80</b>
7.1 Inmunoprecipitación con beads de sefarosa .....	80
7.2 Inmunoprecipitación con beads magnéticas .....	81
7.2.1 Inmunoprecipitación con beads magnéticas .....	81
7.2.2 Inmunoprecipitación con paraformaldehído y beads magnéticas.....	82
<b>RESULTADOS .....</b>	<b>87</b>
<b>CAPÍTULO I: Análisis proteómico diferencial en transfectantes estables de AIRE en HEK293.....</b>	<b>87</b>
1.1 Generación de un modelo celular de expresión estable de AIRE para estudiar los proteomas y los peptidomas de HLA-I .....	87
1.2 Impacto de la expresión de AIRE en el proteoma celular (proteómica cuantitativa) .....	87
1.3 Identificación de los dominios de AIRE involucrados en apoptosis.....	95
1.4 La expresión de AIRE induce una mayor movilidad en células HEK293 .....	98
1.5 Caracterización de los repertorios peptídicos asociados a las moléculas de HLA-I en células que expresan AIRE de forma diferencial.....	102
<b>CAPÍTULO II: AIRE y ciclo celular. Caracterización de una nueva localización de AIRE en células post-mitóticas.....</b>	<b>109</b>
2.1 Localización celular de AIRE y de la proteína truncada conteniendo diversos dominios.....	109
2.2 AIRE se localiza en células postmitóticas cerca del midbody .....	112
2.3 Efecto de la proteína AIRE sobre el ciclo celular y mitosis en transfectantes estables	113

2.4 Análisis de la distribución de AIRE, DNA-PK y Miosina 10 en el midbody .....	115
2.5 Distribución de AIRE a lo largo del ciclo celular.....	118
<b>CAPÍTULO III: Caracterización de la interacción de AIRE y proteínas pertenecientes a la familia de E3 Ubiquitín ligasas Seven in Absentia Homologue (SIAH).....</b>	<b>123</b>
3.1 CACYBP/SIP se ve incrementada en clones de transfectantes estables de AIRE-Myc/His .....	123
3.2 AIRE interacciona con las proteínas de la familia Seven in Absentia Homologue (SIAH): SIAH1 y SIAH2 .....	127
3.3 AIRE requiere la región que contiene la primera secuencia consenso (PKALVPP) para la interacción con SIAH1 y SIAH2.....	129
3.4 AIRE co-localiza con las proteínas SIAH1 y SIAH2. ....	130
3.5 SIAH1 incrementa los niveles de AIRE poli-ubiquitinado.....	133
3.6 SIAH1 reduce los niveles intracelulares de AIRE .....	135
<b>DISCUSIÓN .....</b>	<b>141</b>
<b>CONCLUSIONES .....</b>	<b>163</b>
<b>BIBLIOGRAFÍA .....</b>	<b>167</b>









# ABREVIATURAS

**Aa:** aminoácido.

**AIRE:** Regulador Autoinmune (*AutoImmune REgulator*)

**APC:** Célula Presentadora de Antígeno (*Antigen Presenting Cell*)

**APECED:** Distrofia Ectodérmica-Candidiasis-Poliendocrinopatía Autoinmune.

**APS1:** Síndrome Poliendocrinopático Autoinmune 1.

**CACYBP/SIP:** Proteína de Unión a Calciclina/Proteína de Interacción con SIAH.

**CARD:** Dominio de Reclutamiento de Caspasas

**CBP:** Proteína de Unión a CREB

**cTEC:** Célula Epiteliales Tímica Cortical (*Cortical Thymic Epithelial Cell*)

**DC:** Células Dendrítica (*Dendritic Cell*)

**DMEM:** Dulbeco's Modified Eagle's Medium

**DMSO:** Dimetilsulfóxido

**DNA-PK:** Proteína Kinasa dependiente de DNA

**DNA-PKcs:** Subunidad catalítica de la Proteína Kinasa dependiente de DNA

**DP:** Doble Positivo (en referencia a maduración de linfocitos T en el timo)

**E1:** Enzima activadora de ubiquitina

**E2:** Enzima conjugadora de ubiquitina

**E3:** Enzima ligasa de ubiquitina

**ELL2:** Factor de Elongación de la RNA Polimerasa II

**ER:** Retículo Endoplásmico (*Endoplasmic Reticulum*)

**eTAC:** Célula que expresa Aire extratímica

**FACT:** Proteína Facilitador de la Transcripción de la Cromatina

**FBS:** Suero Fetal Bovino

**$\gamma$ H2AX:** Miembro X de la familia de H2A fosforilado

**H3K27ac:** Acetilación de la Lisina 27 de la Histona 3

**H3K4me0:** Lisina 4 no metilada de la Histona 3

**H3K4me1:** Hipometilación de la Lisina 4 de la Histona 3

**HEK293:** *Human Embryonic Kidney 293* (en referencia a línea celular)

**HIPK2:** Proteína Kinasa 2 de Interacción con Homeodominio

**HLA:** Antígeno Leucocitario Humano (*Human Leucocyte Antigen*)

**HT93:** *Human Thyroid 93* (en referencia a línea celular)

**HSR:** Región de Tinción Homogénea (*Homogeneously Staining Regions*)

**MHC:** Complejo Mayor de Histocompatibilidad (*Major Histocompatibility Complex*)

**mTEC:** Célula Epitelial Tímica Medular (*Medullary Thymic Epithelial Cell*)

**NLS:** Señal de Localización Nuclear (*Nuclear Localization Sequence*)

**PARP1:** Poli-ADP-ribosa Polimerasa

**pGE:** Expresión Promiscua de Genes (*Promiscuous Gene Expression*)

**PHD:** Homeodominio Vegetal (*Plant Homeodomain*)

**PML-NB:** Cuerpos Nucleares de la Leucemia Promielocítica

**PRR:** Región Rica en Prolinas

**pTEF $\beta$ :** Factor Transcripcional  $\beta$  de Elongación Positivo

**PVDF:** Fluoruro de Polivinilideno

**RING:** *Really Interesting New Gene*

**RNApol-II:** RNA Polimerasa II

**SAND:** Dominio Sp100, AIRE, NuP41/75, Deaf1

**SIAH:** *Seven in Absentia Homologue*

**SIRT1:** Sirtuina 1 Deacetilasa

**SCF:** Factor de Células Madre / Complejo Multiproteico Degradativo Skp1 – Cullin – F-box

**SP:** Simple Positivo (en referencia a maduración de linfocitos T en el timo)

**TCR:** Receptor de Célula T

**TEC:** Células Epiteliales Tímicas (*Thymic Epithelial Cells*)

**TOP:** DNA Topoisomerasa

**TRAs:** Antígenos Restringidos de Tejidos (*Tissue Restricted Antigen*)

**TSS:** Sitio de Inicio de Transcripción





# RESUMEN

El timo es el órgano donde se genera el repertorio de linfocitos T que circulará por la periferia. Para que este repertorio sea inmunocompetente, es decir, sea suficientemente diverso para el reconocimiento de la multitud de patógenos existente y sea mínimamente autorreactivo se requiere una selección exhaustiva en el timo. La tolerancia central evita en gran medida la maduración de linfocitos T con potencial autorreactivo. Consiste en la eliminación de timocitos autorreactivos durante la selección tímica a través de dos procesos: una selección positiva en la corteza, donde sobrevivirán timocitos que logren reconocer moléculas del Complejo Mayor de Histocompatibilidad (MHC) propias asociados a péptidos propios; y de una selección negativa en la médula, que eliminará aquellos que reconozcan complejos MHC-péptido propios con alta afinidad. Para que la selección negativa sea eficiente será necesaria la expresión de todas las proteínas presentes en el organismo. El Regulador Autoinmune (AIRE) es la proteína encargada de dirigir gran parte de la expresión ectópica de antígenos restringidos a tejidos y órganos (TRAs) por parte de las células epiteliales tímicas medulares (mTECs), donde se expresa mayoritariamente. AIRE es una proteína multidominio relacionada con la regulación de la expresión y transcripción génica, apoptosis, diferenciación celular e incluso ciclo celular. Mutaciones en el gen de AIRE provocan la denominada Poliendocrinopatía Autoinmune de tipo 1 (APS-1), enfermedad autoinmune monogénica y poliorgánica, cursando habitualmente con autoanticuerpos en sangre e infiltrados linfocitarios debido a irregularidades en el proceso de tolerancia central.

En un trabajo previo de nuestro grupo se determinaron diferencias a nivel proteómico comparando células que expresaban AIRE con las que no. Aquellas células que expresaban AIRE presentaban un aumento en la expresión de proteínas pro-apoptóticas con una consecuente inducción en la apoptosis respecto a las células control, así como un incremento en proteínas de tipo chaperona y una disminución de proteínas relacionadas con el citoesqueleto. Entre las proteínas incrementadas se

encontró la Calcyclin-Binding Protein o SIAH1-Interacting Protein (CACYPB/SIP). Es una proteína adaptadora que contiene un motivo de unión a una E3 ubiquitín ligasa, *Seven In Absentia Homologue 1* (SIAH1). SIAH1 pertenece a una familia de E3 ubiquitín ligasas de tipo RING. Posee tres miembros, dos de ellos catalíticamente funcionales: SIAH1 y SIAH2 que presentan una elevada similitud de secuencia pero con actividades antagónicas. SIAH1 es pro-apoptótica y SIAH2 pro-oncogénica. Su regulación está determinada en respuesta a señales de daño genotóxico e hipoxia, respectivamente. Ambas proteínas comparten sustratos, presentando algunos de ellos relación con la proteína AIRE. Un análisis de la secuencia proteica de AIRE permitió observar que la proteína contenía dos posibles motivos de unión a SIAH1, uno completo y otro con una mutación conservadora.

En la presente tesis se han caracterizado, mediante técnicas bioquímicas y de proteómica cuantitativa, varios aspectos relacionados con la proteína AIRE y el efecto de su expresión en líneas celulares, que han permitido encontrar nuevas posibles funciones de la proteína: En primer lugar, se han comparado mediante técnicas de proteómica cuantitativa (iTRAQ) los proteomas de células que expresaban AIRE de forma diferencial y se han comparado los peptidomas de HLA-I de las mismas células, estudio que no se había realizado hasta el momento. Adicionalmente se han mapeado los dominios de AIRE para inducir apoptosis. Finalmente, esta aproximación ha permitido mostrar que AIRE induce un incremento en la movilidad celular. En segundo lugar, se ha caracterizado una nueva localización de la proteína AIRE en células post-mitóticas, cerca del *midbody*. Asimismo, se ha logrado mostrar la localización de AIRE a lo largo del ciclo celular y, más concretamente, en las diferentes fases de la mitosis; En último lugar, se ha caracterizado bioquímica y funcionalmente la interacción entre AIRE y las E3 ubiquitín ligasas de la familia de SIAH, mostrando que AIRE interacciona con SIAH1 Y SIAH2 a través del primer motivo de interacción y que la interacción con SIAH1 permite su poliubiquitinación y degradación. El efecto de la interacción con SIAH2 no está tan claro, aunque parece que hay una menor ubiquitinación que con SIAH1.







# INTRODUCCIÓN

## I. Tolerancia y Educación Tímica

Hace décadas, los experimentos de timectomía, trasplante de timo y trasplante de injertos en ratón de Miller lograron definir la función clave de este órgano: la inducción de tolerancia en los linfocitos T<sup>1</sup>. La **tolerancia inmune** se define como el conjunto de mecanismos que asegura que el Sistema Inmune no produzca respuesta frente a antígenos propios<sup>2</sup>. El timo es un órgano crítico para el Sistema Inmune, cuya respuesta está comprometida por una red compleja de componentes celulares y moleculares divididos globalmente en Respuesta inmune innata y Respuesta inmune adaptativa<sup>3</sup>. La respuesta inmune innata se produce a través de receptores de reconocimiento de patrones codificados en línea germinal, expresados por células como macrófagos, células dendríticas y granulocitos. Una vez activadas tras el reconocimiento de antígenos actúan rápidamente (minutos u horas). Los componentes de la inmunidad innata para favorecer la respuesta inmune adaptativa, más tardía y específica. La respuesta inmune adaptativa se basa, fundamentalmente, en la respuesta y acción de linfocitos B, que reconocen el antígeno directamente y linfocitos T cuyo reconocimiento de péptidos se realiza a través de la presentación del antígeno por moléculas del complejo principal de histocompatibilidad (MHC) de clase I (MHC-I) y clase II (MHC-II)<sup>4 5</sup>.

El reconocimiento del antígeno se realiza a través de los receptores de superficie expresados en los linfocitos. En humanos, estos receptores son la inmunoglobulina de membrana que junto con el dímero CD79 $\alpha$  y CD79 $\beta$  forman el Receptor de Células B (BCR) y el Receptor de Células T (TCR) que junto con CD3 forma el complejo TCR. Estos receptores pertenecen a la superfamilia de las Inmunoglobulinas, son clonalmente diversos, sufren recombinación de sus segmentos génicos para poder reconocer virtualmente cualquier antígeno y se expresan en la membrana plasmática o, en el caso de los anticuerpos, pueden secretarse en forma soluble<sup>6</sup>. El reordenamiento de los

segmentos génicos (V(D)J) para la expresión de estos receptores es una característica exclusiva del sistema inmune adaptativo. Durante el desarrollo de las células B y T, cada una debe generar un receptor funcional mediante una recombinación génica somática de los diferentes segmentos (V, D y J)<sup>7</sup>. Si el proceso produce un receptor antigénico funcional, las células inmaduras podrán recibir señales de supervivencia y desarrollo a través del contacto con las células estromales de la médula ósea o del timo. En el caso de las células T, si el proceso de recombinación genera receptores que no reconocen el MHC propio o que reconocen antígenos propios con alta afinidad, los precursores celulares se eliminan evitando el desarrollo de linfocitos T no funcionales o potencialmente autorreactivos<sup>7 8</sup>.

Los precursores linfoides de células T que se establecen en el timo como linfocitos T inmaduros (timocitos) sufren un proceso de activación, desarrollo y maduración para transformarse en células T maduras naïve denominado **educación tímica**. Los timocitos en desarrollo recorren el tejido tímico siguiendo un gradiente de quimiocinas desde la zona cortico-medular pasando por la zona subcapsular, córtex tímico y finalizando su recorrido en la médula, desde donde se dirigen a la periferia a través de las vénulas post-capilares<sup>9</sup>. La arquitectura tímica favorece el contacto de los timocitos con múltiples tipos celulares: macrófagos, fibroblastos, células mesenquimales, células epiteliales tímicas (TECs)<sup>10</sup>, células dendríticas (DCs) y células B residentes en el timo. Las interacciones que establezcan con DCs, células B<sup>11</sup> y sobre todo, las TECs, serán esenciales para la selección tímica y la adquisición de tolerancia central, permitiendo la generación de un repertorio de células T inmunocompetentes y la eliminación de aquellas potencialmente autorreactivas<sup>12</sup>. El estroma tímico está compuesto por todas aquellas células no hemapoyéticas, independientemente de su origen o linaje, y generan la estructura del timo, una matriz celular donde se desarrollarán los timocitos. La clasificación del estroma del timo puede realizarse por dos marcadores, la ausencia o presencia del marcador pan-hematopoyético CD45 y la expresión de keratina. Las células keratina<sup>-</sup> corresponderán a un conjunto de células mesenquimales (fibroblastos, células mesenquimales no fibroblásticas, células de tejido conjuntivo formadoras de la cápsula y septum en el tejido tímico y células endoteliales de la vasculatura del timo.

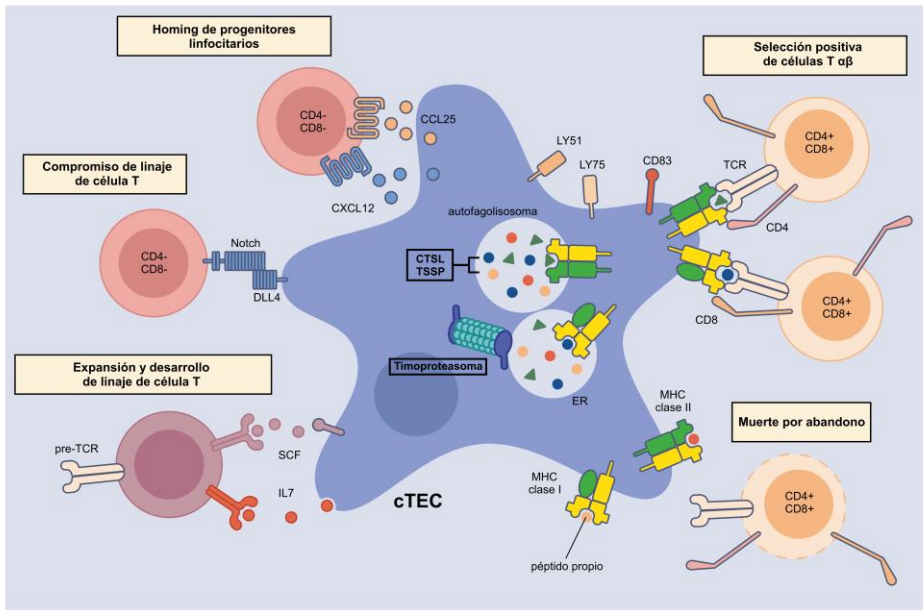
Las células keratina<sup>+</sup> constituirán el epitelio tímico, compuesto por las TECs. Las células CD45<sup>+</sup> corresponderán a DCs, macrófagos y células B, asimismo importantes componentes del estroma tímico<sup>13</sup>.

## 1.1 Células Epiteliales Tímicas (TECs)

Las TECs resultan cruciales para la generación del repertorio de linfocitos T maduros, produciendo factores de proliferación y supervivencia para los timocitos en desarrollo y presentando péptidos propios en el contexto de las moléculas de MHC durante la educación tímica<sup>12</sup>. Las TECs provienen de un único progenitor tímico epitelial común<sup>14</sup> y la diferenciación y especialización de las diferentes poblaciones y subpoblaciones se encuentran dirigidas por la inducción del factor de transcripción *forkhead-box N1* (FoxN1)<sup>16</sup>. Diferentes estímulos posteriores determinarán la diferenciación a uno u otro subtipo. Las TECs se dividen en dos tipos según su localización dentro del tejido tímico y expresión diferencial de marcadores específicos<sup>17</sup>. La expresión de estos factores, no obstante, varía en cada subpoblación siendo las TECs un tipo celular muy heterogéneo.

Las células epiteliales tímica corticales (**cTECs**) se localizan en el córtex tímico y pueden identificarse por la expresión de citokeratina 8,  $\beta$ 5t, Ly51, CD205, CXCL12 y CCL25<sup>18 19</sup>. Algunas de las moléculas que definen esta población resultan claves en la educación tímica. La diferenciación de estas células es inducida gracias a la interacción con los propios timocitos en desarrollo. La expresión de las quimiocinas CXCL12 y CCL25 induce la atracción y localización de los precursores de linfocitos T a la corteza tímica<sup>10</sup>, además las cTECs expresan el ligando NOTCH similar a Delta 4 (DLL4), el cual determina la inducción de compromiso de línea T de los precursores linfoides que llegan al timo. La producción de IL-7 y el factor de células madre (SCF) asegura la proliferación de los timocitos en desarrollo<sup>20</sup>. Las cTECs también expresan componentes esenciales en el procesamiento de antígenos que influirán en el proceso de selección tímica, como la subunidad  $\beta$ 5t que permite la generación del timoproteasoma<sup>21</sup> o algunas proteasas

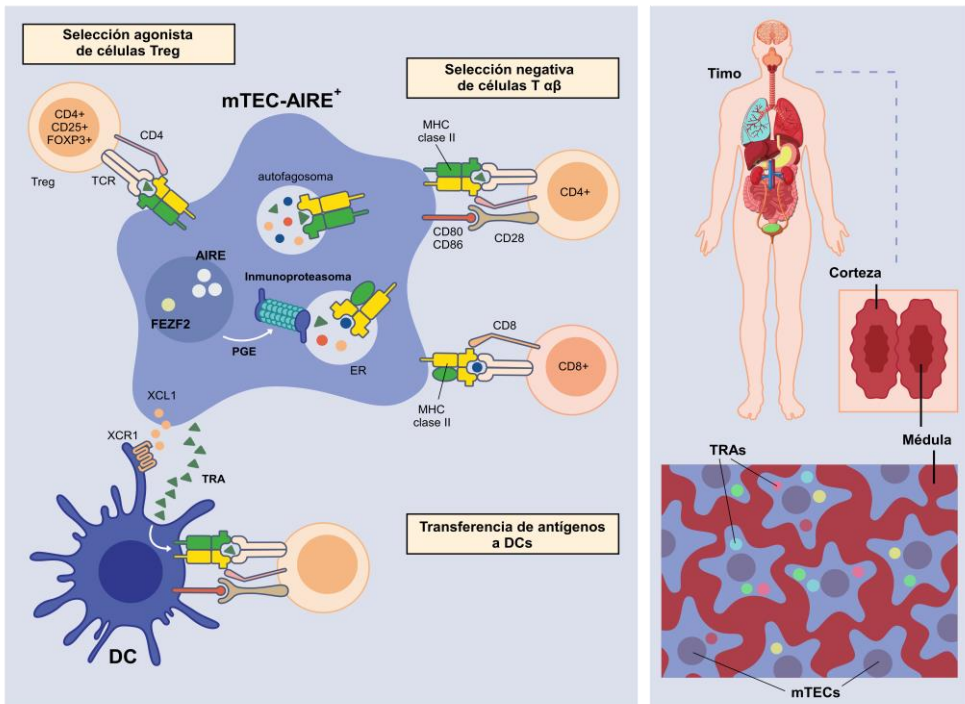
específicas como la proteasa endosomal-lisosomal Catepsina L o la serin-proteasa específica de timo<sup>22</sup> (Imagen 1).



**IMAGEN 1. Funciones y marcadores celulares específicos de las cTECs durante la educación tímica.** Catepsina L (CTSL), Serín-Proteasa-Específica de Timo (TSSP). Adaptación de *Abramsom et al. Nature Reviews Immunology (2019)*.

Las células epiteliales tímicas medulares (**mTECs**) se localizan en la médula tímica y se caracterizan por la expresión de citokeratina 5, CD80/CD86, Claudinas 3 y 4 y CCL19/21<sup>18 19</sup>. Su diferenciación está dirigida por la inducción de vías de señalización que habilitan la expresión del factor de transcripción NF- $\kappa$ B (las vías de señalización RANK y CD40)<sup>23</sup>. Existen dos subpoblaciones principales acorde al grado de expresión en superficie de MHC-II y CD80/CD86 categorizándose en células mTEC<sup>low</sup>, con baja expresión de MHC-II y CD80/CD86 y células mTEC<sup>hi</sup> con alta expresión de los mismos<sup>24</sup>. Aunque la subpoblación mTEC<sup>low</sup> son precursoras de las mTEC<sup>hi</sup>, pueden permanecer como mTEC<sup>low</sup> terminalmente diferenciadas al expresar involucrin, formando parte de los corpúsculos de Hassall o expresar CCL21 interviniendo en la selección tímica al atraer timocitos desde la corteza hacia la médula<sup>19 25</sup>. Algunas células de subpoblaciones mTEC<sup>low</sup> expresarán RANK y CD40 (receptores de la familia TNF) cuyos ligandos, RANKL y CD40L, respectivamente, serán aportados por la población de

timocitos en desarrollo. La activación del receptor RANK induce la expresión del Regulador Autoinmune (AIRE en humanos y Aire en ratón), marcador esencial de la población mTEC<sup>hi</sup> maduras y requerido para una eficiente selección tímica puesto que controla la expresión de determinados antígenos para su presentación en superficie<sup>23</sup>. Tanto la subpoblación mTEC<sup>low</sup>, como las mTEC<sup>hi</sup>-AIRE<sup>+</sup>, pueden expresar Fezf2, otro factor de transcripción con función similar a AIRE: la modulación de la antigenicidad del epitelio medular tímico<sup>26</sup> (**Imagen 2**).



**IMAGEN 2. Funciones y marcadores celulares específicos de las mTECs durante la educación tímica.**

Expresión Promiscua de Genes (PGE). Antígenos Restringidos de Tejido (TRA). Adaptación de Abramson et al. *Nature Reviews Immunology* (2019).

mTECs post-AIRE, provenientes de linaje mTEC<sup>hi</sup>-AIRE<sup>+</sup><sup>27 25</sup>. Estas células están caracterizadas por una pérdida en la expresión de AIRE y un incremento en marcadores como involucrina, desmogleina o keratina 10 y se han relacionado con la biología de corpúsculos de Hassall, que junto con las mTEC<sup>low</sup> terminalmente diferenciadas, intervendrían en la activación de DCs y generación de células T reguladoras (Tregs) al inducir linfopoyetina estromal tímica (TSLP)<sup>28 29</sup>. Por último, células mTECs post-AIRE

pueden generar otra subpoblación, las denominadas células *tuft* tímicas, muy similares a las encontradas en la mucosa intestinal. Estas células están reguladas por la expresión del factor de transcripción POU2F3 y caracterizadas por la expresión de MHC-II, kinasa DCLK1 y producción de IL25, participando en la diferenciación de células NKT2 o células linfoides residentes de intestino<sup>25</sup>.

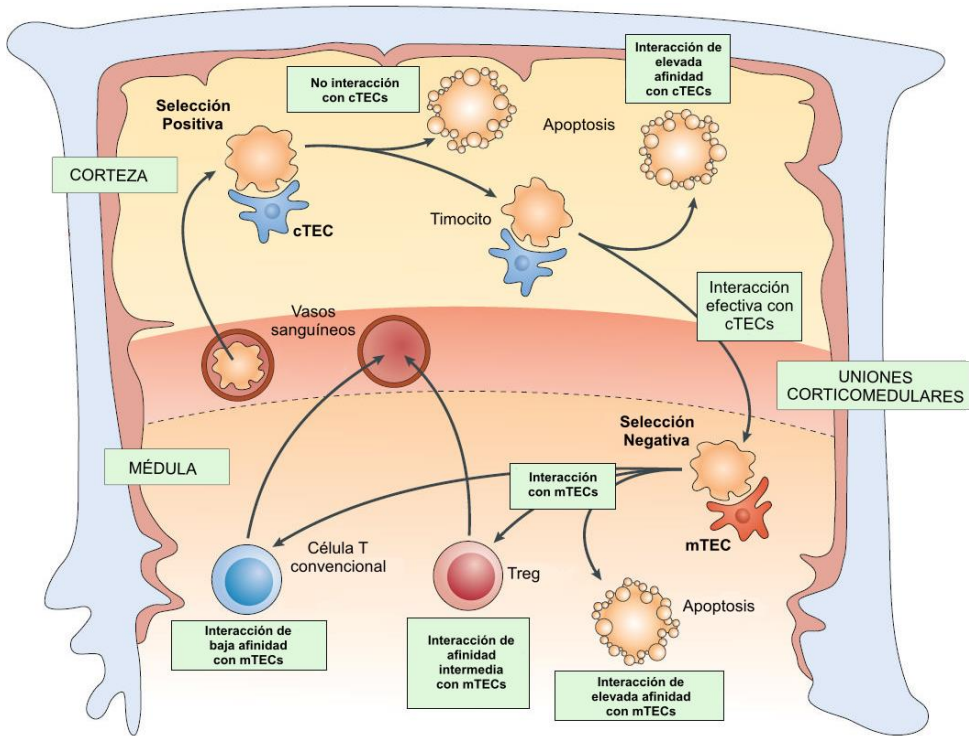
La educación tímica estará determinada, por tanto, en base a la diversidad de autoantígenos que sean accesibles para el conjunto de timocitos en desarrollo. La variedad de esos autoantígenos estará definida por la expresión de antígenos ubicuos y antígenos específicos de tejido a través de las células presentadoras de antígeno profesionales (APCs) residentes en el timo, además de aquellos epítomos propios que puedan llegar al timo a través de la circulación o asociados a células migratorias que ingresan en el mismo<sup>17</sup>. Esa diversidad de expresión antigénica está basada fundamentalmente en un proceso denominado **Expresión Promiscua de Genes** (pGE), característico y común en las TECs<sup>30 31</sup>. Las diferentes poblaciones de TECs son capaces de expresar hasta más de 19000 genes; concretamente en las mTECs. Esta diversidad de expresión génica es máxima debido a la expresión concomitante de Aire. Aire permite la regulación específica de un amplio set de genes restringidos de tejido<sup>30 31</sup>. Dichos genes generarán una amplia variedad de **antígenos restringidos de tejido (TRAs)**, definidos como aquellos antígenos que son expresados en menos de cinco tipos de tejidos como son la insulina, tiroglobulina, algunos antígenos oculares o la proteína básica de mielina pero también antígenos restringidos a estados de diferenciación, desarrollo y de expresión temporal<sup>30</sup>. Aunque AIRE constituya un determinante molecular importante implicado en la pGE, no regula el total de TRAs en las mTECs. Existen genes independientes de la actividad de este factor que son igualmente expresados en el ambiente intratímico<sup>32 26 31</sup>



## 1.2 Selección Tímica

El timo es el órgano donde los linfocitos T maduran. De manera general, el ambiente tímico atraerá progenitores de células T derivados de médula ósea y los dirigirá hacia la zona subcortical y la corteza, así, la expresión de receptores de quimiocinas CCR9 y CXCR4 en su superficie permitirá su migración para establecer contacto con las cTECs (CCL25 y CXCL12)<sup>33</sup>. Durante el recorrido por la corteza tímica, los timocitos dobles positivos (DP) CD8<sup>+</sup> y CD4<sup>+</sup> sufren una **selección positiva** en la que se establece el compromiso de linaje a través de un contacto productivo por interacción de sus receptores de célula T (TCR) con complejos péptido propio-MHC propio presentados por las cTECs. Este proceso inducirá señales de supervivencia y diferenciación a timocito simple positivo (SP) como SP CD8<sup>+</sup> o SP CD4<sup>+</sup>. Los timocitos que superan dicha etapa y no sufren una “muerte por abandono” o por señales de apoptosis derivadas de una interacción con alta afinidad, permanecen como timocitos SP. En este primer proceso se obtiene un repertorio de linfocitos T capaces de responder en periferia a células presentadoras de antígeno (APCs) propias infectadas o tumorales, pero con potencial autorreactivo, ya que pueden reconocer MHC propio cargado con péptidos propios con alta afinidad<sup>12 34</sup>. El proceso de selección positiva inducirá en los timocitos SP la expresión de CCR7 favoreciendo la migración de los mismo hacia la médula tímica atraídos por el gradiente de quimiocinas CCL19/21. En la médula, las mTECs y DCs presentarán un amplio abanico de antígenos propios, incluidos aquellos denominados como antígenos restringidos de tejido (TRA). El reconocimiento de estos antígenos permitirá una **selección negativa** en la que solamente sobrevivirán aquellos timocitos SP que reconozcan complejos MHC propio/ péptidos propios con baja afinidad<sup>33 12 34</sup>. Algunas poblaciones de timocitos SP CD4<sup>+</sup> que describen una afinidad intermedia por complejos MHC/péptido propios acabarán diferenciándose hacia células T reguladoras (Tregs) con otro fenotipo y características. La implicación de mTECs y DCs en la selección y diferenciación de poblaciones Treg es controvertida. Así, en ratones, la ablación de Aire no generaba cambios sustanciales a nivel poblacional en Tregs, sin embargo, estudios con modelos transgénicos de TCR:cognado se observó que Aire era requerido para la generación de Tregs con TCR transgénicos específicos. En humanos, la

generación de células Treg dependientes de AIRE es similar a la generada por DCs<sup>35 36 37</sup>  
<sup>38</sup>. Finalmente, los timocitos SP maduros expresarán el marcador S1P1, receptor de esfingosina 1-fosfato o S1P, que permitirá la egrésión desde el timo a periferia como linfocitos T maduros<sup>39</sup>. Por tanto, el proceso de educación tímica o doble selección asegura la eliminación de aquellos clones de linfocitos T autorreactivos frente a autoantígenos presentados por moléculas de MHC propias, evitando su liberación a la periferia y el consecuente desarrollo de respuestas autoinmunes (**Imagen 3**).



**IMAGEN 3. Selección Tímica.** Representación esquemática de la selección tímica. Adaptación de *Handel et al. Nature Reviews Neurology (2018)*.

II. F

En 1997 se identificó que el gen *AIRE*, que codificaba para la proteína con el mismo nombre, estaba relacionado con la Poliendocrinopatía Autoinmune tipo 1 (APS1), también llamada Síndrome Poliendocrinopatía Autoinmune-Candidiasis-Distrofia Ectodérmica (APECED), una enfermedad rara autoinmune monogénica de carácter autosómico recesivo<sup>40 41</sup>. Este síndrome provoca respuestas autoinmunes órgano-específicas con el consecuente deterioro y destrucción de multitud de tejidos, principalmente tejidos de tipo endocrino. Las manifestaciones principales de esta enfermedad se dividen en tres componentes: candidiasis crónica, hipoparatiroidismo autoinmune y enfermedad de Addison, la identificación de al menos dos de ellos son necesarios para el diagnóstico<sup>40 42 43</sup>. Los factores etiológicos que provocan la enfermedad se producen por la pérdida de la tolerancia central debido a fallos en el proceso de selección tímica. La pérdida en la correcta expresión de *AIRE* ya sea por la producción de una proteína aberrante o una proteína con dominios no funcionales, convergen en el impedimento de la regulación en la expresión de TRAs y, por tanto, de la adquisición de una correcta antigenicidad en la médula del timo, provocando en última instancia la liberación a periferia de timocitos potencialmente autorreactivos y afectaciones en las poblaciones de linfocitos Treg para el control de la tolerancia periférica<sup>43 44</sup>. Todo ello provoca la generación de una plétora de autoanticuerpos dirigidos contra autoantígenos comunes entre pacientes con este síndrome, así como infiltrados linfocitarios en diversos órganos y tejidos. Es una enfermedad rara pero especialmente prevalente en algunas poblaciones, como población finlandesa (1/25000), población sarda (1/14000) y judíos iraníes (1/9000)<sup>40</sup>. El estudio del papel de *AIRE* en esta enfermedad y con la función ejercida en el timo fueron posibles gracias a estudios de ratones deficientes en dicho gen, puesto que mostraban un fenotipo similar a APECED en humano<sup>45</sup>.

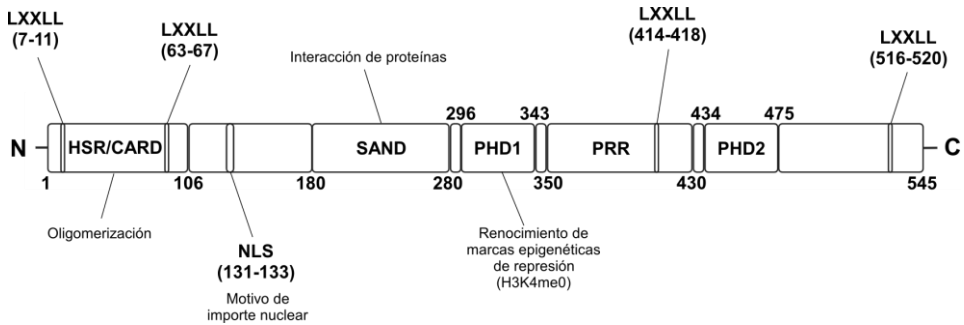
El gen *AIRE* se encuentra localizado en la región q22.3 del cromosoma 21. Además de la implicación en la regulación de la pGE, su expresión en las mTECs ha sido relacionada con la diferenciación y maduración de dichas células, y por tanto, con la alteración de la arquitectura tisular de la médula del timo. En última instancia, la deficiencia de *AIRE*

provoca alteraciones en la migración y selección de timocitos en desarrollo, afectando directamente al número y maduración de mTECs<sup>23 46 47</sup>. AIRE también es expresado por otras poblaciones celulares, como las DCs y células B residentes en el timo<sup>11</sup>, en las denominadas células extratímicas residentes en órganos linfoides secundarios<sup>48</sup>, en células de la línea germinal masculina<sup>49 50</sup> y en células embrionarias<sup>51 52</sup>.

## 2.1 Características bioquímicas y biofísicas de AIRE

El gen *AIRE* codifica para una proteína estructuralmente homóloga a proteínas nucleares que intervienen en procesos de regulación transcripcional y capaces de unirse a cromatina<sup>41</sup>, es una proteína multidominio de 545 aminoácidos. A lo largo de su secuencia, presenta diferentes dominios y motivos funcionales que permiten su actividad como regulador transcripcional. Los motivos de AIRE son (**Imagen 4**):

—**Región de Tinción Homogénea/Dominio de reclutamiento de Caspasas (HSR/CARD)**: se extiende a través de los primeros aminoácidos (1-106). Inicialmente se describió como HSR por la elevada similitud y homología de secuencia al dominio HSR-SAND presente en las proteínas de la familia Sp100, análisis de secuencias basados en estructuras tridimensionales posteriores identificaron que el dominio presentaba una elevada homología con el Dominio de Reclutamiento de Caspasas (CARD) de Apaf-1. El dominio CARD adquiere una estructura de seis  $\alpha$ -hélices que disponen de una serie de aminoácidos hidrofóbicos altamente conservados hacia el interior de la estructura<sup>53</sup>. Este dominio permite la oligomerización de la proteína<sup>54</sup>, característica necesaria para el desarrollo de su función principal como factor activador de la transcripción génica<sup>55</sup>. La presencia de dominios CARD en gran cantidad de proteínas relacionadas con muerte celular programada sugirió que este dominio estaba relacionado con la inducción de apoptosis sobre las mTECs, favoreciendo un rápido *turnover* de las mismas<sup>56</sup>. Este hecho favorecería el proceso de presentación antigénica de TRAs mediante la *cross*-presentación en DCs de los péptidos liberados por dichas mTECs apoptóticas<sup>56</sup>.



**IMAGEN 4.** Representación esquemática de la secuencia de la proteína AIRE. Se identifican los distintos dominios y motivos que componen el Regulador Autoinmune y algunas funciones y actividades que desempeñan. Adaptación de *Peterson et al. Nature Reviews Immunology (2008)*.

—**Señal de Localización Nuclear (NLS):** permite el transporte hacia el núcleo celular gracias a la acción de adaptadores y receptores pertenecientes a la familia de las carioferinas como la importina  $\alpha$ . En AIRE, este motivo se identificó como bi-parcial ( $_{110}\text{RKGRK}_{114}$  y  $_{131}\text{KRK}_{133}$ , separados por un *linker* de 16aa) aunque solamente el segundo motivo es el funcional<sup>57 58</sup>.

—**Dominio SAND:** se encuentra localizado entre las posiciones 181-280, su secuencia contiene una serie de aminoácidos básicos sugiriendo interacción con ligandos de carga negativa como los grupos fosfato del DNA<sup>59</sup>. En AIRE, este dominio no contiene específicamente las secuencias canónicas KDWK o NKWK, por lo que probablemente ejerza su función como zona de anclaje e interacción con diversas proteínas. Independientemente de su secuencia, es un dominio esencial para permitir la transcripción génica y es fundamental para su correcta localización celular<sup>60 61</sup>.

Todos estos dominios reúnen la mayoría de residuos de lisina dentro de la estructura proteica, residuos susceptibles de sufrir acetilación. Esta modificación resulta determinante para la actividad de esta proteína puesto que regula la capacidad de participación de AIRE en complejos multiproteicos y, por tanto, de su correcta distribución dentro de la célula. Proteínas acetiltransferasas como CBP/p300 inhiben su

capacidad de interacción con proteínas heterólogas, mientras que la deacetilasa Sirtuina 1 (SIRT1) facilita dicha actividad (es expresada altamente en células mTECs)<sup>62 63</sup>.

—**Homeodominio Vegetal 1 y 2 (PHD1 y PHD2)**: se extienden desde las posiciones 296-343 y 434-475, respectivamente, y están separados por una región rica en prolinas PRR. Los dominios PHD pertenecen a la familia de dedos de zinc y están caracterizados por presentar motivos ricos en cisteínas organizados en una secuencia de 4 cisteínas / 1 histidina / 3 cisteínas que permiten la coordinación de dos iones de zinc. Este tipo de dominios están relacionados con el reconocimiento de marcas epigenéticas sobre cromatina, concretamente, el grado de metilación de la Histona 3 (H3). Las características bioquímicas del dominio PHD1 permite el reconocimiento específico selectivo de la lisina 4 no metilada de la H3 (H3K4me0)<sup>64</sup>. El dominio PHD2, sin embargo, aunque comparte elevada similitud con el dominio PHD1 presenta una superficie con carácter electrostático positivo, imposibilitando la interacción con histonas. No obstante, la integridad del dominio PHD2 es crucial para la regulación de la transcripción génica ejercida por AIRE, la modificación por mutagénesis dirigida o la delección completa de este dominio aboga la inducción en la expresión de aquellos genes dependientes de AIRE<sup>65</sup>, frente a la pérdida de eficiencia de aquellas mutaciones inducidas exclusivamente en el dominio PHD1<sup>55 66</sup>. AIRE ha sido relacionado con eventos de modificación post-traduccionales como la ubiquitinación o con el sistema Ubiquitín-Proteasoma debido a la presencia de dos dominios PHD en su secuencia<sup>67</sup>. De manera general, los dominios PHD se encuentran conservados entre proteínas cuya actividad está asociada a la regulación de la transcripción mediada por interacción con cromatina<sup>68</sup>. Los dominios PHD de AIRE presentan la estructura canónica de dominio PHD<sup>64</sup> y ésta, resulta claramente distinta al plegamiento de un dedo RING típicamente presente en enzimas E3 Ubiquitín Ligasa<sup>69</sup>.

—**Motivos LXXLL**: AIRE contiene cuatro de estos motivos distribuidos sobre las posiciones 7-11, 63-67, 414-418 y 516-520. Participan en interacciones proteína-proteína y aparecen comúnmente en proteínas coactivadoras y correpresoras que se unen a receptores nucleares e influyen en la transcripción génica<sup>70</sup>. El motivo LXXLL del

extremo carboxilo terminal es esencial en la secuencia de AIRE para la correcta transactivación de genes<sup>71</sup>.

## 2.2 Mutaciones en el gen AIRE

Hasta el momento se han descrito más de 100 mutaciones en el gen *AIRE*<sup>72</sup>. La mayoría de las mutaciones identificadas **se concentran en los dominios funcionales de la proteína**. Las mutaciones predominantes aparecen sobre la región codificante para el dominio HSR/CARD y los dominios PHD<sup>44 73</sup>. Algunas de ellas consisten en mutaciones *nonsense* o de cambio de marco de lectura, generando polipéptidos truncados. Otras son mutaciones que afectan a un solo aminoácido, generando la proteína completa pero no funcional<sup>73</sup>. De manera general, es determinante que AIRE logre localizarse en el núcleo y generar oligómeros para ejercer su función, la integridad del resto de dominios, aunque fundamental para lograr una correcta regulación de la transcripción génica, puede resultar prescindible en casos de heterocigosis mientras se co-exprese una parte de la proteína AIRE sana<sup>73</sup>. Hasta la fecha se han identificado muy pocas mutaciones que operan de manera dominante negativa, entre ellas: algunas que afectan al dominio SAND (p.G228W o p.R247C) y otras que recaen sobre los dominios PHD1 (p.C311Y) y PHD2 (p.C446G)<sup>73 74</sup>. En estudios *in vitro* se comprobó el efecto dominante negativo sobre genes regulados por AIRE tras la co-transfección de AIRE WT y AIRE con diversas mutaciones sobre el dominio PHD1, incluida p.C311Y sobre líneas celulares. Otras mutaciones relativas al dominio HSR/CARD y la mutación p.C311\* mostraron un efecto recesivo. La mutación p.R471C referente al dominio PHD2 no ofrecía variaciones a nivel de regulación sobre genes dependientes comparado con AIRE WT<sup>73 74 75</sup>. En estudios *in vivo* en modelos animales de ratón se confirmaron estos efectos y características: ratones con mutaciones monoalélicas (p.C311Y) y (p.C446G) presentaban un impedimento para la expresión de TRAs dependientes de AIRE en mTECs y una reducción en el número de células Treg en el timo, así como respuestas autoinmunes de carácter leve<sup>76</sup>. Las mutaciones dominantes negativas muestran una penetrancia incompleta y los pacientes portadores presentan cuadros clínicos alejados

de los comunes establecidos para APS-1 de carácter recesivo<sup>77</sup>. Esto es debido a que, de manera general, las mutaciones dominantes-negativas generan proteínas disfuncionales, actuando como inhibidores competitivos de la función de la proteína WT. La formación de hetero-oligómeros (MT/WT) u homo-oligómeros (MT/MT)<sup>54</sup> provocaría, por una parte, la incapacidad de AIRE para formar oligómeros funcionales y una localización celular óptima para ejercer su actividad como factor de transcripción y, por otra, la inhibición de la interacción con otras proteínas que colaboran con AIRE regulando su función<sup>73</sup>. Algunas mutaciones sin sentido p.R257\* o p.C311\* generarían proteínas truncadas que establecerían hetero-oligómeros con proteína AIRE WT, manteniendo el *core* funcional (dímeros de AIRE WT activo), ello explicaría el carácter recesivo de éstas mutaciones<sup>74 73</sup>. Afectaciones a los dominios PHD1 o PHD2 en las mutaciones p.C311Y y p.C446G generarían homo-oligómeros o hetero-oligómeros incapaces de ejercer su actividad sobre la cromatina pese a la correcta localización subcelular<sup>76</sup>.

La distribución y localización subcelular de AIRE en mTECs queda relegada al núcleo en forma de *dots* nucleares morfológicamente muy similares a los denominados cuerpos nucleares de la Leucemia Promielocítica o PML-NBs<sup>78 79</sup>. La dificultad para el estudio del mismo a nivel celular *in situ* en el tejido tímico o para el aislamiento de mTECs-AIRE<sup>+</sup> debido a la baja proporción de células en comparación con otras poblaciones celulares del tejido, llevó al estudio de la proteína en líneas celulares, transfectando AIRE WT, AIRE truncado en diversos dominios o portando mutaciones comunes en pacientes de APECED. La distribución de AIRE en líneas celulares recordaba a la distribución natural observada en las mTECs, con la diferencia de la presencia también en el citoplasma formando estructuras fibrilares que recordaban a las generadas por microtúbulos y filamentos intermedios, además co-localiza con proteínas estructurales del citoesqueleto como son la  $\alpha$ -tubulina y la vimentina<sup>79 80 81</sup>.



## 2.3 Mecanismo molecular de acción

La expresión génica inducida por AIRE parte de la interacción con la cromatina y con diversas proteínas relacionadas con la maquinaria de reparación génica e inicio de la transcripción. La interacción se produce gracias al reconocimiento de marcas epigenéticas de remodelación de cromatina. El dominio PHD1 interacciona con la histona H3 atraído por la no-metilación de H3K4me0 y/o repelido por hipermetilación H3K4me3<sup>82 64</sup>. Estudios de secuenciación de RNA de células únicas sobre mTECs determinaron que AIRE también es capaz de reconocer otras marcas epigenéticas de represión, como la H3K27me3 regulado por diferentes metiltransferasas<sup>31</sup>.

A nivel molecular, AIRE se localizaría en regiones amplias de cromatina donde se aloja un elevado número de factores de transcripción generales y específicos de cada tipo celular y caracterizados por presentar marcas epigenéticas de represión, los denominados *super-enhancers*<sup>83</sup>. Estas estructuras fomentan la transcripción de regiones intra-cromosómicas o inter-cromosómicas de manera simultánea al organizar proteínas y factores de transcripción sobre los sitios de inicio de transcripción (TSSs) de dichas regiones<sup>84</sup>. En los *super-enhancers*, AIRE actuaría formando parte de complejos multiproteicos en el que intervienen numerosas proteínas, principalmente la maquinaria de respuesta a daño genotóxico y de inicio y elongación de la transcripción.

En estas regiones, AIRE interacciona con la DNA Topoisomerasa 1 (TOP1) favoreciendo, por un lado, el inicio y estabilización de las roturas de doble cadena en el DNA; por otro, el reclutamiento de componentes de la maquinaria de respuesta a daño genotóxico ( $\gamma$ H2AX, DNA-PK, Ku80 y PARP-1). Además, componentes esenciales para la transcripción génica como la RNA Polimerasa II (RNA-PolII), la proteína bromodominio 4 (BRD4) y el factor de transcripción de elongación positivo  $\beta$  ( $\beta$ TAF $\beta$ ) también son reclutados generando, en suma, grandes complejos multiproteicos orquestados por AIRE que sirven de unión entre las actividades desarrolladas por las maquinarias de transcripción y *splicing*<sup>63 83 85</sup>. Estos *super-enhancers* pueden actuar como puertos de

factores de transcripción para las TSSs, particularmente aquellos que se encuentran asociados a la RNA-PolIII. Esta proteína se encuentra inicialmente interaccionando con el factor de transcripción de elongación negativo (NELF) y el factor inducido sensible a DRB (DSIF) inhibiendo su actividad y en estado pausado de inicio de la transcripción<sup>86 87</sup>. La movilización de pTEF $\beta$  a zonas próximas a RNA-PolIII inactiva permitiría la fosforilación de DSIF y la estabilización de su actividad para la transcripción génica. En estado transcripcional abortivo, la RNA-PolIII puede promover roturas de doble cadena de ADN a través de TOP2a favoreciendo, paralelamente, la generación de complejos proteicos orquestados por AIRE<sup>85</sup>.

Este escenario describiría una regulación de transcripción génica **estocásticamente determinista** sobre multitud de genes, muchos de ellos, codificantes de TRAs. AIRE reconoce marcas epigenéticas de represión. Por tanto, es esperable que esa inducción en la expresión génica recaiga sobre genes que, de manera general, se encuentran reprimidos en células diferenciadas<sup>88</sup>. Sumado a ello, se ha de tener en cuenta el elevado grado de pGE en todas las TECs, independientemente de su localización anatómica o grado de maduración, es decir, las TECs *per se* presentan unas características biológicas de remodelación epigenética que facilita el proceso de transcripción de prácticamente todo el genoma independientemente de la expresión o no de AIRE, aunque los mecanismos que las confieren aún son desconocidos. AIRE es capaz de regular un amplio conjunto de genes determinados. Aunque la expresión de los mismos es inducida de manera aleatoria, inducidos en un momento dado y en una mTEC determinada, la transcripción del gen llega a ser hasta 16 veces superior comparada con el nivel de transcripción medio del mismo gen en la población de mTECs<sup>31</sup>. Este hecho facilita la eficiencia en la selección negativa de timocitos: la presentación antigénica de un determinado TRA en una determinada mTECs será máxima y más efectiva que si todas las mTECs presentasen todos los TRAs dependientes de AIRE de manera uniforme. Por contrapartida, las distintas poblaciones de mTEC<sup>hi</sup>-Aire<sup>+</sup> generarán un linaje determinado con un perfil epigenético exclusivo, es por ello que en mTECs de distinto linaje se vea expresada una parte determinada del

total de los genes dependientes de AIRE<sup>88</sup>. Esta situación genera una gran diversidad antigénica en timos de diferentes individuos; la generación del repertorio de linfocitos T siempre será diferente aun presentando “el mismo genoma”. A nivel evolutivo tiene sentido, puesto que se evitaría la generación del mismo agujero de reconocimiento de antígenos en todos los individuos de una misma especie, el precio a pagar es la susceptibilidad a generar respuestas autoinmunes.

## 2.4 AIRE en el ciclo celular

El ciclo celular supone un proceso complejo y altamente controlado en el cual los componentes celulares son cuidadosamente segregados en dos células hijas<sup>89</sup>. Se puede dividir en la interfase, donde la célula podrá permanecer de manera quiescente realizando las funciones biológicas asociadas a su estado de diferenciación (fase G0) o permanecer en un estado de proliferación activo en el cual se prepara para dar lugar a dos células hijas (fase G1, fase S y fase G2)<sup>89 90</sup>. El ciclo celular culmina con el proceso de división celular, la fase M o mitosis, a su vez dividida en cinco estadios (profase, metafase, anafase, telofase y citocinesis)<sup>91</sup>.

El ciclo celular ha de estar minuciosamente controlado, la célula en división ha de superar una serie de puntos de control o *checkpoints* que aseguren:

- El crecimiento suficiente de la célula para poder llevar a cabo la división celular, incluyendo la síntesis de proteínas necesarias para la compleción de la misma y la integridad del material genético<sup>90 92</sup>.
- La correcta duplicación del material genético<sup>90 92</sup>
- La correcta formación del huso mitótico y posterior anclaje de sus microtúbulos al centrómero de los cromosomas duplicados a través del cinetocoro<sup>90 91</sup>.

En el ciclo celular participan numerosas proteínas reguladoras y está principalmente controlado por la actividad fluctuante de complejos proteicos compuestos por las

quinasas dependientes de ciclinas (CDKs) y las ciclinas<sup>93</sup>. La asociación de ciclinas con las CDKs permitirá su activación fosforilando sustratos clave para la síntesis de DNA y la progresión de la mitosis<sup>93</sup>. Las CDKs fosforilan la proteína de retinoblastoma o pRB permitiendo la liberación y activación de factores de transcripción de la familia E2F, necesarios para la entrada de la célula de fase G1 a fase S<sup>94 92</sup>. Tras la duplicación del DNA durante la fase S pueden generarse roturas de doble hebra (DSBs) que provocan la activación de la maquinaria de reparación de DNA (DNA-PK) y las kinasas ATM y ATR<sup>95</sup>, causando el arresto de la célula en fase G2 para asegurar la integridad del material genético que será repartido durante la mitosis<sup>90</sup>.

La fase M o mitosis se divide en profase, donde se compacta el DNA generando los cromosomas, formados por dos cromátidas hermanas unidas por el centrómero. Durante esta fase se genera el huso acromático que permitirá el reparto de cromátidas entre las células hijas. La envoltura nuclear es disgregada y el huso acromático se extiende hacia los polos celulares. Durante la metafase las cromátidas hermanas se alinean en el plano ecuatorial, los microtúbulos del huso acromático se anclan al centrómero que une las cromátidas a través de los cinetocoros. En la anafase, las cromátidas son separadas simultáneamente hacia los centrómeros opuestos del huso acromático situados en el polo celular. Finalmente, en la telofase, se comienza a sintetizar la envoltura nuclear alrededor de los cromosomas separados que comienzan a descondensarse para generar el núcleo celular en cada célula hija<sup>91</sup>.

El proceso de división celular finaliza con la citocinesis y escisión de la membrana celular, que permiten la separación física entre ambas células hijas. El huso mitótico sufre una reorganización que genera una estructura denominada huso central, localizada en la placa ecuatorial de la célula en división, que dictamina la posición del surco de corte y es clave para la progresión de la citocinesis<sup>96</sup>. La actividad del huso central es iniciada gracias a la actividad del complejo Centralspindlina (MKLP1 y MgcRacGAP) que recluta el factor de intercambio de guanina ECT2 activando a las GTPasas RhoA y Rac. La activación de estas GTPasas provoca una cascada de señalización que favorece la ingresión del surco de corte para la escisión. El complejo

Centralspindlina junto con el Complejo Pasajero del Cromosoma (CPC) conformado por Aurora Kinasa B, INCENP, Borealina y Survivina inician y controlan la progresión de citocinesis activando a proteínas remodeladoras del citoesqueleto y favoreciendo la actividad de proteínas motoras<sup>97</sup>. Una vez que el surco de corte ha ingresado se genera el *midbody*. Esta es una estructura centralizada en el puente citoplasmático que sigue uniendo las células hijas, está compuesto por microtúbulos muy empaquetados junto con material amorfo electrodensito. La Centralspindlina queda altamente concentrada en las regiones próximas al *midbody*<sup>98</sup>, mientras que las zonas más distales quedan relegadas a los microtúbulos y proteínas del complejo ESCRT que permitirán la escisión<sup>99</sup>. Algunas proteínas como MYH10 o DNA-PK participan activamente en el proceso de citocinesis<sup>100 101 102</sup>.

La localización subcelular de AIRE dentro de la célula es dependiente del ciclo celular y la formación de los características *dots* nucleares sería posterior a la división celular, desapareciendo poco antes de la entrada en mitosis y durante el proceso de división<sup>103</sup>. Este hecho sugeriría la imposibilidad de una implicación de AIRE en división celular. No obstante, estudios en células embrionarias de ratón identificaron AIRE como regulador de las proteínas del huso mitótico permitiendo la correcta maduración, duplicación y organización del huso mitótico<sup>52</sup>.

## 2.5 AIRE en periferia

Fuera del timo, se ha detectado la expresión de AIRE en órganos linfoides secundarios a nivel de los nódulos linfáticos. En ese microambiente, células que expresan AIRE extratímicas (eTACs) consiguen establecer interacciones con aquellos linfocitos T naïve potencialmente autorreactivos favoreciendo la inducción de tolerancia<sup>104 48 105</sup>. Estas eTACs constituyen una población de células APC derivadas de médula ósea con precursor común a la población de DCs, presentan marcadores como CD45<sup>low</sup>, MHC-II<sup>hi</sup>, CD80<sup>low</sup>, CD86<sup>low</sup>, EpCAM<sup>hi</sup> con capacidad para inducir tolerancia mediante anergia clonal. AIRE también participa en el proceso de renovación y pluripotencialidad de

células madre embrionarias, favoreciendo un estado de hiperactividad transcripcional y el mantenimiento de un correcto ciclo celular de las mismas<sup>106 52</sup>.

La actividad de AIRE puede trasladarse a diferentes tipos celulares, su expresión genera un impacto sobre amplios *sets* de genes pero los *loci* afectados son diferente en cada tipo celular<sup>107</sup>. Cada tipo de célula y su estado de maduración presenta un conjunto de diferentes promotores actuando sobre los *super-enhancers*, asimismo, los *super-enhancers* suelen distribuirse regulando genes que definen el tipo y el estado de diferenciación celular, además, la distribución de la RNA-PolIII no productiva a lo largo de la cromatina también es variable<sup>108</sup>. En líneas celulares monocíticas, AIRE logra inducir la expresión de genes propios de procesos de maduración de células dendríticas<sup>109</sup>. En líneas celulares epiteliales transfectadas con AIRE, se incrementa la expresión de genes relacionados con la expresión de proteínas pro-apoptóticas y chaperonas mientras que algunas proteínas asociadas al citoesqueleto se ven disminuidas<sup>110</sup>.

### III. El Complejo Principal de Histocompatibilidad (MHC)

La capacidad del sistema inmune adaptativo para responder frente a lo propio y lo extraño está principalmente controlada por el número extremadamente alto de receptores únicos. En el caso de las células T, el TCR reconoce los antígenos peptídicos asociados a los productos proteicos de un conjunto de genes codificados en el denominado Complejo Principal de Histocompatibilidad (*Major Histocompatibility Complex*, MHC) o Antígeno Leucocitario Humano (*Human Leucocyte Antigen*, HLA) en humanos<sup>111</sup>.

En humanos la región del MHC se encuentra localizado en el brazo corto del cromosoma 6 y codifican para diferentes proteínas que intervienen tanto en la respuesta inmune innata y adaptativa como en el procesamiento y presentación antigénica<sup>112</sup>. Se encuentra dividido en tres regiones:

**Clase I:** situados en la región más telomérica, codifican para la cadena  $\alpha$  de las moléculas de MHC clásicas de clase I (HLA-A, B y C) altamente polimórficas; moléculas similares a clase I (denominadas no clásicas, *class I-like* o de tipo Ib) entre las que se encuentran HLA-E, F y G, y ligandos del receptor NKG2 (MIC-A y MIC-B) fundamentales para la respuesta inmune citotóxica<sup>112</sup>.

**Clase II:** localizados en la región centromérica, codifican para las moléculas MHC clásicas de clase II (HLA-DR, DQ y DP) y para moléculas no clásicas relacionadas con el ensamblaje del MHC (HLA-DM y DO), así como el transportador asociado con procesamiento antigénico (TAP-1 y TAP-2) y las subunidades del inmunoproteasoma LMP2( $\beta$ 1i) y LMP7 ( $\beta$ 5i)<sup>112</sup>.

**Clase III:** situada entre las regiones Clase I y Clase II. Aúna genes que codifican proteínas con función diversa aunque relacionadas con respuestas de tipo inmune e inflamatoria como citoquinas del factor de necrosis tumoral, proteínas chaperonas o componentes del sistema de complemento<sup>113</sup>.

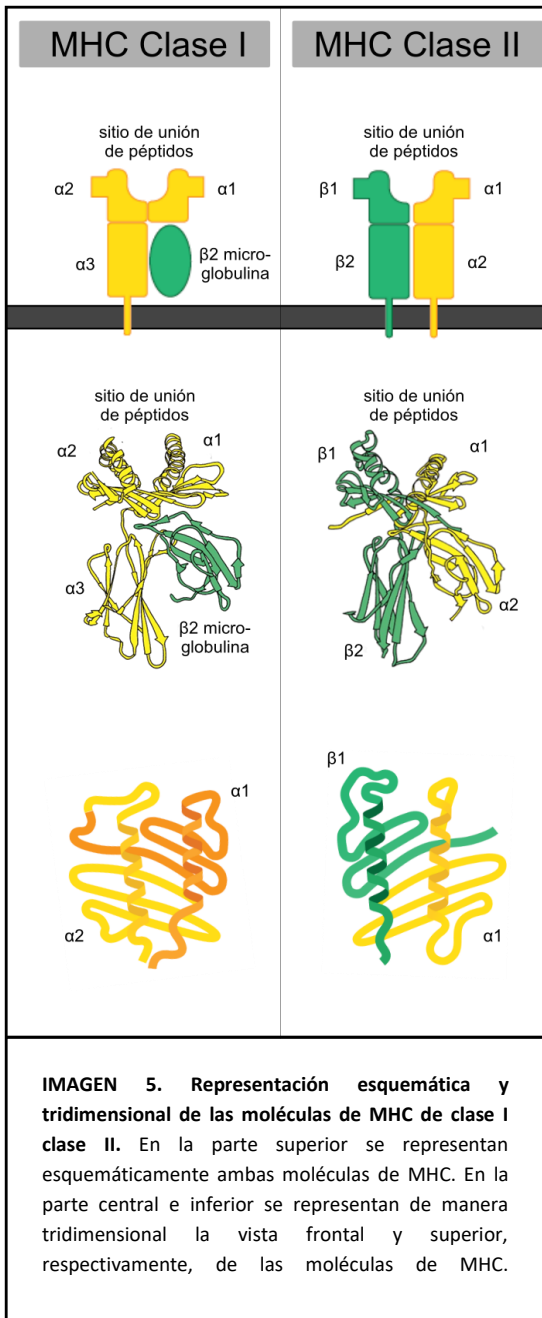
El descubrimiento de los *loci* que alojaban los genes relativos al MHC tras la observación de procesos de rechazo hiperagudo en trasplantes, permitió definir su función, la cual permite generar una respuesta específica del sistema inmune frente a lo extraño, siendo el elemento restrictivo a través del cual los antígenos pueden ser reconocidos por el receptor de células T (TCR)<sup>114</sup>. La presión evolutiva para generar una respuesta inmune frente a virus y otros patógenos ha podido conferir a los genes que codifican las moléculas clásicas de MHC un elevado polimorfismo, permitiendo asegurar una respuesta frente al mayor número de elementos externo. Esta variedad confiere una resistencia o susceptibilidad a desarrollar diversas patologías de tipo infecciosa o autoinmune<sup>115 116</sup>.

### 3.1 Estructura de las moléculas de MHC

Los genes de las moléculas de MHC clásicas codifican para una serie de glicoproteínas encargadas de la presentación en superficie de péptidos antigénicos para ser reconocidos por linfocitos T<sup>117</sup>. Las moléculas clásicas de MHC de clase I se componen de un heterodímero unido no covalentemente de cadenas de diferente tamaño: la **cadena  $\alpha$**  de mayor tamaño (sobre 43kDa), codificada en humanos por tres genes de la región de clase I (HLA-A, -B y -C), la cual presenta tres dominios extracelulares denominados  $\alpha 1$ ,  $\alpha 2$  y  $\alpha 3$  (de la superfamilia de las inmunoglobulinas), una región transmembrana y una cola intracitoplasmática corta; y una subunidad de menor tamaño denominada  **$\beta 2$  microglobulina** ( $\beta 2m$ , 12kDa). La  $\beta 2m$  es codificada en el cromosoma 15, fuera de la región de HLA<sup>118</sup>. La cadena  $\alpha$  no es estable en membrana a no ser que establezca, a través del dominio  $\alpha 3$  una interacción con la  $\beta 2m$ . Las regiones  $\alpha 1$  y  $\alpha 2$  conforman el sitio de unión a péptido y la región  $\alpha 3$  constituye la zona de reconocimiento por el co-receptor CD8. El sitio de unión a péptido en estas moléculas se genera a partir del plegamiento de la cadena  $\alpha$ , generando un hueco rígido y estanco que permite el anclaje de péptidos de 8 a 11 aminoácidos, con preferencia por nonámeros. Las regiones  $\alpha 1$  y  $\alpha 2$  presentan residuos muy polimórficos mientras que el dominio  $\alpha 3$  y la  $\beta 2m$  poseen residuos relativamente conservados<sup>119 120</sup> (**Imagen 5**).

Las moléculas clásicas de MHC de clase II se componen de un heterodímero unido no covalentemente de cadenas de tamaño similar: la **cadena  $\alpha$**  (35kDa) y la **cadena  $\beta$**  (30kDa), codificadas por los tres pares de genes de la región de clase II (HLA-DR, -DQ y -DP en humanos), las cuales presentan dos dominios extracelulares,  $\alpha 1$  y  $\alpha 2$  y  $\beta 1$  y  $\beta 2$  (del superdominio de inmunoglobulinas), una región transmembrana y una región intracitoplasmática. Las regiones  $\alpha 1$  y  $\beta 1$  generan el sitio de unión a péptido y la región  $\beta 2$  constituye la zona de reconocimiento por el correceptor CD4. El sitio de unión a péptido en éstas moléculas se genera por la interacción de ambas cadenas  $\alpha$  y  $\beta$  generando un hueco más flexible y abierto en sus extremos que permite el anclaje de péptidos de mayor tamaño, principalmente de entre 12-18aa<sup>121 122</sup> (**Imagen 5**).





A pesar de las diferencias, en las moléculas de MHC clase I y clase II el sitio de unión a péptido es estructuralmente similar y está constituido por una hendidura generada a partir de dos  $\alpha$ -hélices (paredes del surco) y una base formada por una lámina  $\beta$  compuesta por ocho hebras  $\beta$  antiparalelas. Como se ha comentado, la estructura se conforma por los dominios  $\alpha 1$  y  $\alpha 2$  en clase I y  $\alpha 1$  y  $\beta 2$  en clase II. En el surco, algunos residuos generan bolsillos o *pockets* donde podrán acomodarse las cadenas laterales de los aminoácidos del ligando peptídico. Las propiedades fisicoquímicas de los residuos que constituyen los *pockets* determinarán qué ligandos peptídicos podrán unirse<sup>123 124</sup>. El polimorfismo característico de los genes que codifican para las moléculas clásicas se sitúa principalmente sobre los residuos que interactúan con el péptido a

presentar<sup>125</sup>. Esto permite que diferentes alotipos puedan presentar diferentes mezclas de péptidos (inmuno-peptidomas) y que, por tanto, cada alotipo presente un repertorio peptídico específico<sup>124</sup>.

### 3.2 Presentación antigénica (Interacción MHC-péptido)

En las moléculas de MHC de clase I los péptidos interaccionan con algunos de los residuos que componen el sitio de unión. Este sitio es cerrado en los extremos y unirá péptidos de entre 8 y 12 aminoácidos, principalmente nonámeros en conformación extendida<sup>126 127 128</sup>. Aquellos péptidos de mayor tamaño sobresaldrán por la zona central del surco hacia el exterior, mostrándose al TCR<sup>129 130 131 132 133</sup>. Dentro del sitio de unión se distribuyen seis *pockets* de unión denominados como A-F. Los ligandos de clase I se fijarán al surco por los extremos amino y carboxilo terminal además de interaccionar con algún otro residuo de los diferentes *pockets*. En la mayoría de alotipos los bolsillos de unión B y F alojan los **residuos de anclaje** de la secuencia del péptido: el residuo en posición 2, P2 (*pocket* B) y el residuo carboxilo terminal PΩ (*pocket* F). Ambos residuos P2 y PΩ juegan un papel importante en la estabilización de la interacción péptido-HLA<sup>128 133 125 134</sup>. El resto de bolsillos acomodarán de forma más laxa cadenas laterales del resto de posiciones de la secuencia del péptido. La posición carboxilo terminal del ligando peptídico suele contener residuos alifáticos, aromáticos o básicos<sup>123 135</sup>. Las características bioquímicas de los residuos que conforman los *pockets* de unión y la secuencia del péptido determinarán si la unión péptido-HLA podrá establecerse de manera efectiva y con mayor o menor afinidad. El conjunto de los diferentes residuos que puede presentar una molécula de HLA clase I determinada se denomina **motivo de anclaje**.

En moléculas de HLA de clase II, los péptidos presentados se distribuirán a lo largo del surco de unión, abierto y flexible, los extremos de la secuencia del péptido podrán sobresalir fuera del mismo. Los ligandos de clase II presentan habitualmente entre 10 a 20 residuos de longitud (o más largos), con preferencia por aquellos de entre 12 y 18 aminoácidos<sup>136</sup>. Al igual que las moléculas de clase I, existe un *core* de unión de unos nueve residuos compartido y restrictivo que determinará el conjunto de péptidos presentado. En la estabilización del *core* intervienen las posiciones P1, P4, P6/P7 y P9 en la mayoría de los alelos<sup>121 137</sup>. Las moléculas de clase II podrán presentar familias de péptidos o *nested sets* formadas por péptidos derivados de la misma proteína que

conservan el *core* de unión en su secuencia aunque varía la longitud de sus extremos amino o carboxilo terminal<sup>136 138 139</sup>.

### 3.3 Vía de procesamiento antigénico

La finalidad de las moléculas de MHC de clase I y clase II será presentar diferentes péptidos para generar una respuesta inmune determinada. El tipo de moléculas de MHC-I y MHC-II que exprese una célula y el origen de los péptidos que constituyen el complejo péptido-MHC determinará el tipo de célula inmune que reconozca el complejo y el tipo de respuesta inmune que activará<sup>140</sup>. Se pueden diferenciar dos vías principales de procesamiento antigénico: la vía de procesamiento antigénico de clase I y la vía de procesamiento antigénico de clase II:

#### 3.3.1 Vía de procesamiento antigénico de clase I

Todas las células nucleadas del organismo expresan moléculas de MHC-I, que serán las encargadas de presentar péptidos a linfocitos T CD8<sup>+</sup><sup>135</sup>. Los péptidos presentados por clase I derivan del procesamiento de proteínas citosólicas o nucleares, bien de proteínas propias (mutadas o no) u originadas por patógenos intracelulares, también pueden derivar de los denominados productos ribosomales defectivos (DRiPs)<sup>141</sup>. Todas ellas serán procesadas en el citosol por el proteasoma. Adicionalmente, se ha descrito que otras proteasas pueden jugar un papel en la generación de determinados antígenos. Este sería el caso de la Tripeptidil Peptidasa II (TPPII)<sup>142 143</sup>, la enzima convertidora de angiotensina (ACE)<sup>144</sup> o la oligopeptidasa Thimet (TOP)<sup>145</sup>. Los péptidos de entre 8 y 16aa son trasladados al lumen del Retículo Endoplásmico (ER) mediante los Transportadores Asociados al Procesamiento Antigénico (TAP-1 y TAP-2), los cuales seleccionarán, principalmente en cuanto a tamaño y residuo hidrofóbico carboxilo terminal, los sustratos que potencialmente se unirán a las moléculas de clase I<sup>146 147</sup> presentes en el lumen del ER. Una vez en el lumen, estos péptidos podrán ser

procesados por las proteasas específicas del ER ERAP1/2<sup>148 149 150 151</sup> antes de ser presentados por las moléculas de MHC de clase I. La acción de estas aminopeptidasas logra definir, por tanto, el repertorio peptídico de clase I<sup>149</sup> (**Imagen 6**).

Las moléculas de clase I funcionales se generan en el ER. Inicialmente la cadena  $\alpha$  se sintetizará en el lumen del ER rugoso donde interaccionará con calnexina<sup>152</sup>, una chaperona que la mantendrá parcialmente plegada. La interacción de la cadena  $\beta$ 2m con la cadena  $\alpha$  de clase I provocará la disociación de la molécula de clase I de calnexina generándose el heterodímero que será estabilizado por otra chaperona, calreticulina<sup>153 154</sup>. El trímero se asociará junto con otras chaperonas tapasina<sup>155</sup> y Erp57<sup>156 157</sup> además de TAP, formando el denominado Complejo de Carga de Péptido (PLC). Los péptidos que ingresan y/o son procesados en el lumen del ER son cargados en las moléculas de clase I de manera eficiente gracias a la acción de la tapasina, los péptidos que se unan con menor afinidad al surco de unión de clase I que otros, serán reemplazados gracias a ella<sup>158</sup>. Tapasina hace de proteína adaptadora entre la molécula de clase I en ensamblaje y TAP. Algunos alelos son más dependientes de tapasina que otros<sup>159 160</sup>. Otra chaperona estructuralmente muy similar y que se une de manera análoga al PLC, la TAPBPR, podría participar en la carga de péptidos de manera independiente a tapasina o participar en el proceso de *cross*-presentación<sup>161</sup>. La unión de un péptido con suficiente afinidad logrará estabilizar la molécula de MHC clase I para la presentación en superficie<sup>162</sup>, donde será presentado a los linfocitos T CD8<sup>+</sup>.

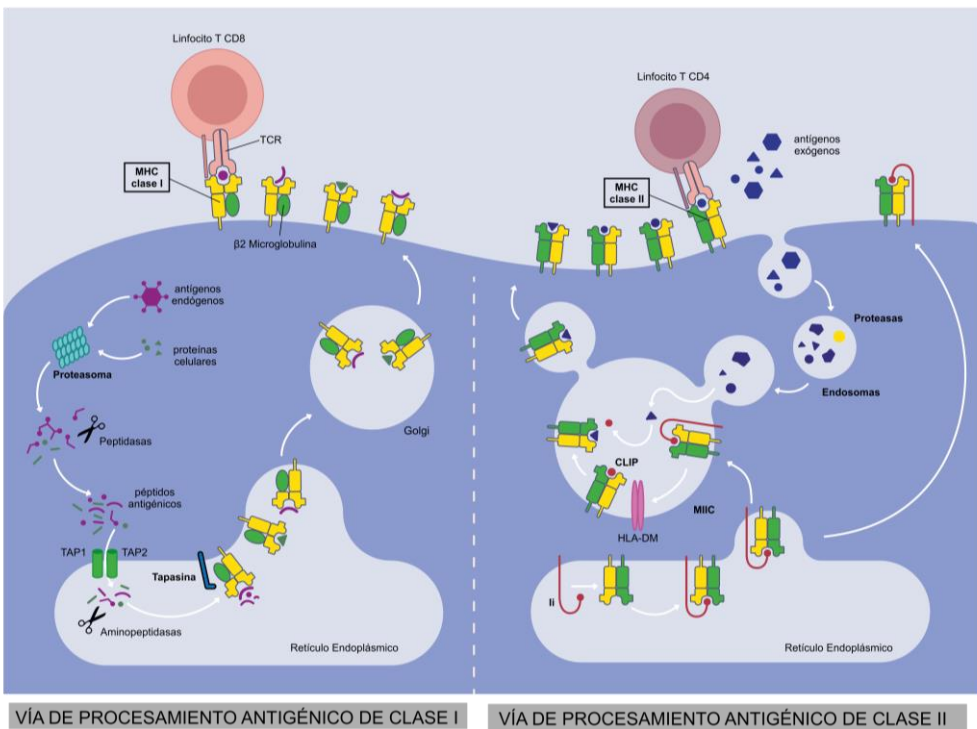
### 3.3.2 Vía de procesamiento antigénico de clase II

Las células presentadoras de antígeno profesionales (APCs) como células dendríticas, macrófagos y células B, así como las cTECs y mTECs expresan moléculas de clase II en superficie, éstas serán las encargadas de presentar antígenos a linfocitos T CD4<sup>+112 163</sup>. La expresión de estas moléculas es inducida en respuesta a IFN $\gamma$  y otros estímulos generados por células no inmunes<sup>164</sup>. Los péptidos presentados por ellas derivan de proteínas presentes en la membrana plasmática, localizadas en compartimentos

vesiculares u orgánulos de la vía endocítica o secretada y provenientes del exterior celular (internalizadas en la vía vesicular a través de fagocitosis, endocitosis, pinocitosis,...). Las cadenas  $\alpha$  y  $\beta$  del heterodímero de clase II son sintetizadas y ensambladas en el ER, donde se asocian a una glicoproteína transmembrana que actúa como pseudo-péptido, la denominada cadena invariante (Ii o CD74)<sup>165 166</sup>. La cadena invariante permite: la estabilización de las moléculas de clase II vacías, evita la unión de péptidos no provenientes de endocitosis desde el lumen del RE hasta que ingresa en la vía secretora y dirige las moléculas de clase II hacia la vía de secreción debido a la presencia de un motivo de dileucinas presente en la cola citoplasmática. Ii forma trímeros donde cada subunidad se une de forma no covalente a un heterodímero de clase II dando lugar a nonámeros  $Ii_3(\alpha\beta)_3$ <sup>167</sup>. Durante el ensamblaje de este complejo, los componentes que lo forman se asocian con calnexina<sup>168</sup>. Una vez generados, los nonámeros se disocian de calnexina y se dirigen a la vía secretora a través de compartimentos endosómicos con un pH bajo, donde ocurrirá la unión con el péptido. Estos compartimentos reciben el nombre de MIIC<sup>169</sup>. En los compartimentos MIIC las diferentes proteasas residentes hidrolizan las cadenas Ii de los nonámeros dejando una fracción peptídica denominada *Class II Associated Ii Peptide* (CLIP) que recorre el surco de unión de las moléculas de clase II. CLIP mantendrá la estabilidad de las moléculas de clase II en estos compartimentos hasta que sea sustituido por un péptido para la presentación en membrana. Las proteínas exógenas que ingresen en MIIC serán procesadas por las proteasas residentes generando los péptidos a presentar<sup>170</sup>. La molécula de HLA clase II no clásica no polimórfica HLA-DM permitirá la regulación del intercambio de CLIP por péptidos derivados de proteínas procesadas en compartimentos endocíticos, así como la protección de las moléculas de MHC-II frente a la inactivación<sup>171 172</sup>. HLA-DO es expresada en células B y médula tímica, modulando la actividad de HLA-DM<sup>173 174</sup>. Una vez conformado el complejo péptido-MHC-II, se transporta a la membrana plasmática mediante la vía secretora donde se presentará en superficie pudiendo ser reconocida por los linfocitos T CD4<sup>+</sup> (**Imagen 6**).

En condiciones normales, la mayoría de los péptidos presentados por ambas clases de MHC, derivan de proteínas autólogas<sup>123 175</sup>. Esta presentación resulta esencial para la

homeostasis del repertorio de linfocitos T circulante. Los procesos de *cross-presentation* permiten la presentación de péptidos generados a partir de proteínas exógenas a través de moléculas de clase I, asegurando la activación de linfocitos T CD8<sup>+</sup> frente a antígenos exógenos<sup>176</sup>. Esta dislocación de la vía de presentación podría deberse a una traslocación de la maquinaria de clase I en la vía endosomal, procesos de transporte retrógrado o a la intervención de TAP transfiriendo péptidos de la vía endosomal al lumen del ER<sup>177</sup>.



**IMAGEN 6. Representación de las vías de procesamiento antigénico de clase I y clase II.** Adaptación de Groettrup et al. *Nature Reviews Immunology* (2010) y Neefjes et al. *Nature Reviews Immunology* (2011).

## IV. El Sistema Ubiquitín Proteasoma (UPS)

Uno de los procesos necesarios para la generación de péptidos potencialmente presentables por el MHC es la degradación y procesamiento de proteínas<sup>178 179</sup>. La concentración intracelular de proteínas está extremadamente regulada. La vida media de cada una varía dependiendo de la naturaleza de la proteína, algunas de ellas como

los DRiPs, los factores de transcripción o enzimas metabólicas presentan una vida media muy corta. El catabolismo de proteínas generan aminoácidos libres para la homeostasis proteica y péptidos citosólicos, algunos de los cuales se unirán a MHC-I<sup>180</sup>. El Sistema Ubiquitín Proteasoma (UPS) es el principal sistema de degradación de proteínas del interior celular. Es un proceso que requiere de una modificación postraduccional multifacética y dinámica que afecta a casi todos los aspectos de la biología eucariótica, la denominada **ubiquitinación** o unión covalente de ubiquitina. La proteína contiene 76 aa, con siete residuos de lisina (K) y un extremo amino terminal que serán susceptibles de modificaciones posteriores (sucesivas ubiquitinaciones, acetilaciones, fosforilaciones, SUMOilaciones,...). Todo ello genera una plétora de señales con diversa funcionalidad y finalidad conocido como el código de ubiquitinación<sup>181</sup>.

La ubiquitinación comienza con la unión del extremo carboxilo terminal de una única molécula de ubiquitina sobre un residuo de Lys de un sustrato. La ubiquitina es unida a sustratos a través de una cascada enzimática en tres pasos muy sofisticada. En ella participan las denominadas enzimas E1 ubiquitín-activadoras, E2 ubiquitín-conjugadoras y E3 ubiquitín ligasas. Las distintas modificaciones posteriores marcarán la finalidad de la ubiquitinación. E1 activa la ubiquitina de manera ATP-dependiente y la transfiere a E2, formando un enlace tioéster entre el sitio activo de cisteína presente en la E2 y el extremo carboxilo terminal de ubiquitina. La actividad simultánea de E2-ubiquitina con la E3 y el sustrato específico culmina en la formación de un enlace isopeptídico entre el grupo amino de la cadena lateral de lisina del sustrato con el carboxilo terminal de la ubiquitina<sup>181 182</sup>. Las enzimas E3 resultan componentes críticos en la cascada puesto que controlan estrictamente la eficiencia y especificidad del proceso de ubiquitinación. Existen dos tipos principales de E3 ubiquitín ligasas en eucariotas, definidos por el tipo de dominio catalítico que portan, el dominio RING (*Really Interesting New Gene*) o dominio HECT (*Homologous to the E6AP Carboxyl Terminus*). Se diferencian en estructura y tipo de procesamiento catalítico<sup>183</sup>.

Las **enzimas E3 RING** permiten la transferencia directa de la ubiquitina activada desde la E2 al sustrato. Resulta la familia más grande de E3 ubiquitín ligasas (más de 600 miembros), su dominio catalítico y sitio de unión a sustrato puede presentarse en una misma subunidad o en diferentes subunidades de un complejo E3<sup>184</sup>. La estructura de los dominios tipo RING se compone de una secuencia de ocho residuos de cisteínas o histidinas que acomplejan iones  $Zn^{2+}$ , aunque la secuencia de residuos puede variar. La presencia de iones  $Zn^{2+}$  tampoco resulta necesaria para el caso de las enzimas E3 U-box, pertenecientes a esta familia ya que la estructura terciaria adoptada a través de interacciones no covalentes entre los residuos del *core* es suficiente para unir la enzima E2<sup>185 186</sup>. Las **enzimas E3 HECT**, en cambio, generan un enlace tioéster intermediario con la ubiquitina de la E2 antes de ser agregada al sustrato<sup>187</sup>.

La función de todas las enzimas E3 no reside exclusivamente en su actividad catalítica, también por su capacidad de reconocimiento específico de sustrato. Debido a la presencia de modificaciones por ubiquitinación en prácticamente todas las vías celulares es necesario poder distinguir sustratos específicos del resto de proteínas<sup>181</sup>. La mayoría de sustratos son reconocidos por las E3 ubiquitín ligasas a través de una pequeña secuencia denominada *degron*. En otros casos, el reconocimiento específico se realiza a través de dominios proteicos<sup>183</sup>.

En el UPS, las proteínas son marcadas para la degradación por la unión covalente de varias unidades de ubiquitina a través del residuo K48 o K11 de la ubiquitina<sup>181</sup>. Este proceso generalmente conocido como poliubiquitinación es el más común de todo el código de ubiquitinación y marca los sustratos para su traslocación al proteasoma, al ser reconocidas por proteínas receptoras de ubiquitina con dominios de unión a ubiquitina, y su posterior degradación. Las proteínas ubiquitinadas serán hidrolizadas a oligopéptidos por el proteasoma, un complejo multicatalítico presente en el citosol de manera libre o asociado al ER y también en el núcleo de células eucariotas<sup>179 188</sup>.



El proteasoma es un complejo multiproteico encargado de la degradación y procesamiento de proteínas en el citoplasma generando aminoácidos y péptidos<sup>189</sup>. Su actividad se produce tras la asociación de un *core* catalítico, el proteasoma 20S, junto con otros complejos reguladores. La forma más común en células de mamíferos es el denominado **proteasoma 26S**, resultado de la asociación del proteasoma 20S con una o dos subunidades del complejo regulador 19S. Los complejos reguladores facilitan el acceso del sustrato y coordinan el reconocimiento y la traslocación del mismo al *core* catalítico 20S<sup>190 191</sup>. El proteasoma 20S presenta una estructura de barril, formada por cuatro anillos heptaméricos externos estructurales y dos anillos centrales catalíticos. Los anillos externos se componen de siete subunidades  $\alpha$  ( $\alpha$ 1- $\alpha$ 7). Los anillos internos, de siete subunidades  $\beta$  ( $\beta$ 1- $\beta$ 7)<sup>192</sup>. De ellas, las subunidades  $\beta$ 1,  $\beta$ 2 y  $\beta$ 5 son catalíticas y aunque actúan mediante un mecanismo de hidrólisis común, presentan especificidades de sustrato diferentes. La subunidad  $\beta$ 1 presenta función caspasa (hidrólisis detrás de aminoácidos ácidos), la subunidad  $\beta$ 2 con función trípica (hidrólisis después de aminoácidos básicos) y la subunidad  $\beta$ 5 con función quimiotrípica (hidrólisis después de aminoácidos hidrofóbicos)<sup>193 194</sup>. En algunos tejidos linfoides y células expuestas a IFN $\gamma$  y TNF $\alpha$ , se produce el reemplazo de las subunidades catalíticas del proteasoma 26S por otras tres subunidades:  $\beta$ 1i,  $\beta$ 2i y  $\beta$ 5i, que reciben el nombre de inmunosubunidades. La sustitución de las subunidades del proteasoma constitutivo por las inmunosubunidades genera el conocido como **inmunoproteasoma**, con un perfil de actividad diferente y más eficiente en la generación de algunos ligandos de clase I<sup>195 196</sup>  
<sup>197 198</sup>. Ve disminuida su actividad caspasa en comparación con sus actividades trípica y quimiotrípica<sup>193 196</sup>. En las cTECs presentes en la corteza del timo existe el **timoproteasoma**. En él, la subunidad  $\beta$ 5i se sustituye por la  $\beta$ 5t constituyendo su *core* catalítico por  $\beta$ 1i,  $\beta$ 2i y  $\beta$ 5t. La adición de  $\beta$ 5t reduce la actividad quimiotrípica mientras que la actividad caspasa y trípica se mantienen. El timoproteasoma es importante en el establecimiento de un repertorio inmunocompetente de linfocitos T CD8<sup>+</sup><sup>199 21</sup>.

El UPS es, por tanto, esencial para el mantenimiento de la homeostasis proteica, asegurando los niveles proteicos y la eliminación de proteínas aberrantes. La proteólisis ejercida por el proteasoma también regula un gran número de procesos biológicos como apoptosis, señalización o diferenciación celular, entre otros, permitiendo la progresión del ciclo celular y la supervivencia de la célula, al generar péptidos antigénicos presentados por el MHC de clase I<sup>179 200 180</sup>.

## V. La familia E3 Ubiquitin Ligasas *Seven in Absentia Homolog* (SIAH)

La familia *Seven in Absentia Homolog* (SIAH) es una familia de E3 ubiquitín ligasas ortóloga a la familia *Seven in Absentia* (SINA) en invertebrados. En vertebrados, esta familia se compone de tres miembros SIAH1, SIAH2 y SIAH3, las cuales presentan sustratos específicos y procesos de regulación y actividades diferentes<sup>201</sup>. SIAH1 es considerada supresora tumoral debido a que en varios tipos de cáncer su expresión se encuentra reducida. SIAH2, en cambio, tiene una actividad proto-oncogénica y se encuentra incrementada en melanoma, cáncer de próstata, mama, pulmón, páncreas e hígado<sup>202</sup>. SIAH3, en cambio, presenta la mayor divergencia de secuencia y debido a la pérdida del dominio catalítico RING se postula que podría actuar como posible inhibidor de la actividad del resto de miembros de SIAH<sup>203</sup>.

Las proteínas de la familia SIAH presentan regiones altamente conservadas en su secuencia, pudiéndose identificar un dominio catalítico tipo RING para la unión con enzimas E2, una región rica en cisteínas que conforman dos dedos de zinc tipo SIAH indicando una posible zona mediadora de interacciones proteína-proteína o con DNA, una región de dimerización denominada DIMER y una región de unión a sustrato (SBD) aportando la especificidad<sup>204</sup>. SIAH dimeriza adquiriendo una estructura cuaternaria en forma de S, donde la zona central estaría constituida por el enfrentamiento de ocho láminas  $\beta$  antiparalelas de cada monómero de SIAH generando un sándwich y los dedos de zinc, distales<sup>205</sup>. Esta estructura genera huecos o ranuras de interacción para el

reconocimiento de motivos específicos de sustratos y adaptadores de SIAH<sup>206 207</sup>. Las E3 ubiquitin ligasas de SIAH presentan un reconocimiento específico por el denominado **motivo de unión a SIAH**, un *degron* compuesto por una secuencia general de siete aminoácidos PxAxVxP (siendo X cualquier aminoácido). Parece que la valina y la prolina finales (VxP) son los residuos *core* del motivo de unión<sup>208 209 210</sup>. Las proteínas SIAH son capaces de interactuar aisladas o asociadas a diferentes adaptadores. SIAH1 interactúa con el complejo multiproteico degradativo SCF (Sk1-Cullin1-Fbox complex) a través de la proteína adaptadora CACYBP/SIP, este complejo reconoce  $\beta$ -catenina y la ubiquitina para su degradación<sup>211 212 209</sup>. La Proteína de Interacción con SIAH1 o de unión a calciclina CACYBP/SIP (Siah1-Interacting Protein/Calcyclin-Binding Protein) es una proteína adaptadora de pequeño tamaño, se encuentra implicada en un amplio abanico de procesos celulares (defosforilación, ubiquitinación, dinámica del citoesqueleto, regulación de la expresión génica, diferenciación e incluso tumorigénesis)<sup>213</sup>.

SIAH1 y SIAH2 son capaces de auto-ubiquitinarse regulando sus niveles, siendo estabilizadas mediante fosforilación de kinasas como p53, p38, Akt o HIPK2<sup>202</sup>. La familia SIAH también ha sido relacionada con el sistema inmune innato. Las proteínas de la familia SIAH resultan estructuralmente similares a los factores asociados a receptores TNF (TRAF)<sup>205</sup>. El homólogo a SIAH1 en ratón (Siah1b) es capaz de activar NF- $\kappa$ B; por otro lado, SIAH2 es capaz de interactuar con TRAF2 y degradarlo, inhibiendo la capacidad de inducir factores de transcripción como JNK o NF- $\kappa$ B en condiciones de estrés<sup>214</sup>. Algunos estudios también han logrado demostrar la interacción de SIAH1 con MyD88 (proteína adaptadora esencial para la vía de señalización de TLRs) y la enzima E2 conjugadora Ube2d2 que interviene en la poliubiquitinación de TRAF6<sup>215 184</sup>. Los miembros pertenecientes a SIAH, por tanto, están involucrados en la regulación de varias vías de señalización relacionadas con estrés oxidativo, respuesta a hipoxia, apoptosis, supresión tumoral, regulación del ciclo celular, regulación de la transcripción, respuesta a daño genotóxico, espermatogénesis y señalización vía TNF $\alpha$ <sup>216 217 218 202</sup>, presentando un rol principal en homeostasis y respuesta a estrés celular. Como E3 ubiquitin ligasas, pueden regular respuestas

celulares mediando en la estabilidad proteica, su recambio, su localización subcelular o su marcaje para degradación vía proteasoma.





# OBJETIVOS

En un trabajo previo de nuestro laboratorio se realizó un estudio proteómico comparativo de las células epiteliales de origen tiroideo HT93 transfectadas con AIRE y las mismas células sin transfectar. En este estudio se identificó un patrón de proteínas incrementadas y disminuidas, lo que nos permitió confirmar que AIRE induce apoptosis en células epiteliales<sup>110</sup>. Sin embargo, no se realizó este estudio proteómico en otras células que confirmaran que el efecto de AIRE en apoptosis es general y que nos permitiera observar cambios en la abundancia proteica que pudieran sugerir nuevas funciones de AIRE. Adicionalmente, hasta el momento no se ha analizado la influencia de AIRE en los peptidomas presentados por las moléculas de HLA-I, proceso fundamental para el reconocimiento por parte de los timocitos y para mostrar el estado funcional de la célula.

Asimismo, se ha descrito que en células embrionarias de ratón, AIRE puede jugar un papel durante la mitosis<sup>52</sup>. En nuestro conocimiento no se ha realizado ningún estudio en este sentido en células humanas.

En el estudio proteómico previo de nuestro laboratorio se observó que células epiteliales que expresaban AIRE mostraron un incremento en los niveles de CACYBP/SIP<sup>110</sup>. Este hecho se ha confirmado en el estudio realizado en esta tesis con las células HEK293. Este hecho, junto con el papel adaptador de CACYBP/SIP respecto a SIAH1, papel esencial que facilita la ubiquitinación de sustratos, nos permitió plantearnos el papel de AIRE en la ruta que involucra a las proteínas SIAH.

La presente tesis comprende el estudio de tres aspectos novedosos de AIRE: en primer lugar, el análisis de los proteomas y de los inmunopeptidomas de HLA-I en células HEK293 y dichas células transfectadas con AIRE; en segundo lugar el papel de AIRE en el ciclo celular, durante la mitosis y la citocinesis; y en tercer lugar, el papel de AIRE en el procesamiento antigénico a través de su interacción con las proteínas SIAH.

Los objetivos de la tesis son:

- **Objetivo 1:** Estudiar el papel de AIRE en la conformación de los proteomas celulares y los inmunopeptidomas asociados a HLA-I.
- **Objetivo 2:** Estudiar el papel de AIRE en el ciclo celular y su localización durante la mitosis y la citocinesis.
- **Objetivo 3:** Analizar el papel de AIRE en el procesamiento antigénico mediante la caracterización de su interacción con proteínas de la vía de degradación de las E3 ubiquitín ligasas pertenecientes a la familia de proteínas SIAH.







# MATERIALES Y MÉTODOS

## I. Cultivos celulares y vectores

### 1.1 Líneas celulares de origen no linfoide

- a) HEK293: *Human Embryonic Kidney 293*. Línea celular de riñón embrionario humano. Son células adherentes de morfología estrellada y fácilmente transfectables<sup>219</sup>.
- b) HT93: *Human Thyroid 93*. Línea celular de tejido tiroideo humano. Células adherentes de morfología estrellada<sup>220</sup>
- c) HeLa: *Henrietta Lack*. Línea celular derivada de un carcinoma de cuello de útero. Son células adherentes con una morfología estrellada<sup>221</sup>.
- d) SK-Hep-1: línea celular derivada del líquido ascítico de un paciente con adenocarcinoma hepático. Células adherentes de origen endotelial<sup>222</sup>.

### 1.2 Condiciones de cultivo

Todas las células se cultivaron en medio DMEM suplementado con 10% suero fetal bovino (FBS) en un incubador en condiciones de esterilidad a 37°C y 5% de CO<sub>2</sub> controlando periódicamente la confluencia, la morfología, el estado del medio y la viabilidad de las mismas.

Los continuos pases celulares (ciclo de 3/4 días) se realizaron mediante un proceso de tripsinización: una vez las células llegaron a monocapa (máxima confluencia celular) se aspiró el medio de cultivo y se lavó la superficie con 10mL de tampón fosfato salino (PBS 1X) estéril y se añadieron 3mL de Tripsina-EDTA 1X (0.25% Tripsina, 0.02% EDTA, Sigma<sup>TM</sup>). Las células se incubaron 5 min a 37°C y se resuspendieron, conservando un

volumen variable dependiendo de las necesidades. Los volúmenes de PBS 1X y Tripsina fueron escalados para cultivos de mayor superficie. Finalmente se añadió medio fresco. El recuento celular se realizó en cámara de Neubauer.

Cuando se requirió, las células HEK293 transfectadas con construcciones génicas de interés de expresión estable se mantuvieron bajo las mismas condiciones de cultivo anteriormente descritas, en DMEM, 10% FBS y 1 mg/mL Geneticina (G418) (Gibco™).

Se congelaron *pellets* celulares periódicamente de cada tipo celular en cultivo. Para ello, las células se recuperaron en esterilidad mediante tripsinización y se centrifugaron a 500 RCF. A continuación se lavaron 3 veces con PBS 1X frío a 500 RCF, el último lavado de 10 min de duración. Una vez retirado el sobrenadante se sumergieron en unos 10mL de N<sub>2</sub> líquido y se almacenaron a -80°C hasta su utilización.

### 1.3 Construcciones genéticas

Las células HEK293 se transfectaron con varias construcciones generadas en el vector de expresión pcDNA3.1 (con gen de resistencia a geneticina):

Se clonaron los genes *AIRE*, completo o las secuencias que codificaban diferentes dominios de la proteína, y *CACYBP* usando primers específicos en el vector de expresión eucariota pcDNA3.1. Los genes se amplificaron por PCR y se ligaron usando la T4 DNA ligasa. Asimismo, se clonaron dichos genes añadiendo un tag de Myc/His o FLAG en el extremo 3'. Las construcciones que se usaron fueron las siguientes:

—**AIRE**; **AIRE-Myc/His**: vector pcDNA3.1 que presenta el gen AIRE insertado de manera aislada o asociado a una cola Myc/His.

—**AIRE-(1-106/1-180/1-280/1-343/1-475)-Myc/His**: vector pcDNA3.1 que presentan respectivamente el gen AIRE truncado a lo largo de los diferentes dominios de la proteína hacia el C-terminal insertado asociado a una cola Myc/His.

—**AIRE-(107-545/181-545/281-545/344-545/476-545)-Myc/His**: vector pcDNA3.1 que presentan respectivamente el gen AIRE truncado a lo largo de los diferentes dominios de la proteína hacia el N-terminal insertado asociado a una cola Myc/His.

—**FLAG-SIAH1**; **FLAG-SIAH2**: vector pcDNA3.1 que presenta el gen SIAH1 y SIAH2, respectivamente insertado asociado con una cola FLAG. Ambas construcciones fueron cedidas por el Dr. Hagen T. (*National University of Singapore* (NUS) – Singapur)

—**FLAG-Ub**: vector pcDNA3.1 que presenta el gen de Ubiquitina insertado asociado a una cola FLAG en el extremo N-terminal.

## 1.4 Transfecciones

### 1.4.1 Transfección transitoria

Aproximadamente  $2 \times 10^5$  células se plaquearon en cada pocillo de placas p6 (Corning™) con 2 mL de medio de cultivo DMEM, 10% FBS. Tras 24h a 37°C y 5% CO<sub>2</sub>, y 30 min previo a la transfección, las células fueron deprivadas de suero retirando el medio anterior y añadiendo 2mL de medio de cultivo DMEM no suplementado. Las soluciones de transfección se generaron diluyendo, por un lado, 1µg de las distintas construcciones con el DNA de interés en medio de cultivo DMEM sin suplementar, por otro, el volumen de Polietilenimina (PEI, Polysciences™) requerido en medio de cultivo DMEM sin suplementar con una relación PEI:DNA de 3:1. Se aplicó el volumen de la solución de transfección con PEI sobre el volumen de la solución de transfección con las construcciones de DNA. Las soluciones de transfección se mantuvieron a temperatura ambiente durante 30 min y se aplicaron sobre cada pocillo. Las células se incubaron con la solución de transfección durante 4h a 37°C, 5% CO<sub>2</sub>. Finalmente el medio de cultivo

se cambió por medio fresco DMEM, 10% FBS. Tras 48h de incubación a 37°C 5% CO<sub>2</sub> las células se recuperaron mecánicamente o mediante tripsinización para el análisis correspondiente.

#### **1.4.2 Transfección estable y clonación**

El proceso de transfección estable se realizó de manera similar al de la transfección transitoria anteriormente descrita. Tras 48h de incubación a 37°C, 5% CO<sub>2</sub>, las células se mantuvieron en medio de selección DMEM, 10% FBS, 1 mg/mL G418. Se comprobó periódicamente el nivel de expresión del gen de interés mediante citometría de flujo.

Para obtener poblaciones puras de los transfectantes se procedió a la clonación de los transfectantes con varias poblaciones por dilución límite. Las células se tripsinizaron y se contaron en una cámara de Neubauer. Se realizaron diluciones seriadas en medio de selección DMEM 10% FBS, 1 mg/mL G418. Posteriormente se plaquearon 100uL de las diferentes diluciones celulares con una concentración celular aproximada de 30 cél/mL y 3 cél/mL en placas p96 de fondo plano (Corning™) que presentan previamente 100uL de medio de selección DMEM 10% FBS, 1mg/mL G418. Las placas p96 se mantuvieron en condiciones estériles en el incubador a 37°C, 5% CO<sub>2</sub> revisando periódicamente el crecimiento de pocillos por la acidificación del medio. Aquellos pocillos en los que crecieron células fueron posteriormente transferidos a placas de cultivo p6. Finalmente la expresión del gen de interés y presencia de una única población celular se comprobó por citometría de flujo.

## II. Proteómica y peptidómica

### 2.1 Proteómica cuantitativa (iTRAQ)

#### 2.1.1 Digestión proteica y marcaje con el reactivo iTRAQ-8-plex®

La concentración de proteína total se determinó usando el Pierce 660 nm *protein assay kit* (Pierce). Para la digestión 40µg de cada réplica y condición se precipitó con el método del metanol/cloroformo. Los *pellets* proteicos se resuspendieron y desnaturalizaron en 20µL de 7M Urea / 2M Tiourea, 0.1M TEAB, pH 7.5 (SERVA Electroforesis), reducido con 1µL de 50mM Tris (2-carboxietil) fosfina (TCEP, AB SCIEX), pH 8.0, a 37°C durante 60 min. Tras este paso, 2µL de 200mM reactivo de bloqueo de cisteínas (metil metanotiosulfonato (MMTS, Pierce) fue añadido durante 10 min a RT. Las muestras se diluyeron hasta 120µL para reducir la concentración de Urea/Tiourea con 50µL TEAB. Se añadieron 2µg de tripsina modificada (Pierce) a cada muestra a una relación 1/20 (p/p) y se incubó O/N a 37°C en un agitador. Las muestras digeridas se evaporaron a sequedad.

Cada muestra tripsinizada se reconstituyó con 80µL de 70% etanol / 50mM TEAB, se marcaron a RT durante 2h con media unidad de *iTRAQ Reagent 8-plex kit* (AB SCIEX, Foster City, CA, USA). El marcaje con iTRAQ se llevó a cabo de acuerdo al siguiente esquema:

<i>Reporter Ion</i>	<i>Sample</i>	
	Experimento 1	Experimento 2
113	HEK293 WT R1	HEK293 WT R3
114	HEK293 WT R2	HEK293 WT R4
115	HEK AIRE R1	HEK AIRE R3
116	HEK AIRE R2	HEK AIRE R4

Después del marcaje, las muestras se combinaron y la reacción se paró por evaporación en un Speed Vac. Las muestras se desalaron usando un SepPak C18 cartridge (Waters).

### **2.1.2 Cromatografía líquida y análisis por espectrometría de masas**

Una alícuota de cada muestra se analizó por 2D-nano-LC ESI-MSMS usando un sistema de cromatografía líquida nano (Ultimate 3000, Thermo) acoplado a un espectrómetro de masas Thermo Orbitrap Exploris OE240 con una fuente EasySpray. El volumen de inyección fue 5 $\mu$ L. La columna analítica usada fue una columna de fase reversa C18 basada en sílica de 75 $\mu$ m x 15cm (Thermo), 2 $\mu$ m de tamaño de partícula. La *trap column* es de 5 $\mu$ m de diámetro de partícula, 100 Å de tamaño de poro, conectado online con la columna analítica. La carga de muestra fue con 0.1% ácido fórmico en agua a 2 $\mu$ L/min. La nano-bomba proveyó una tasa de flujo de 250 nL/min y operó bajo un gradiente, usando 0.1% ácido fórmico en agua como fase móvil A y 0.1% ácido fórmico en 80% acetonitrilo como fase móvil B. La longitud de gradiente fue de 120 min.

La adquisición de los datos se llevó a cabo con un sistema Orbitrap Exploris OE240. Todos los datos se adquirieron en modo Data-dependent acquisition (DDA) con TuneExploris (Thermo). Para los parámetros DDA, un MS survey scan in el rango de masas 350-1250 Da (FWHM resolution 60000), seguido por 20 scans MS/MS de los más intensos precursores MS1. El tiempo de exclusión de los iones previos fue de 45 seg.

### **2.1.3 Análisis de datos y estadística**

Los espectros MS/MS se exportaron a formato mgf usando el software Proteome Discoverer 2.5 y la búsqueda se realizó usando Mascot Server 2.5.1, OMSSA 2.1.9, X!TANDEM 2013.02.01.1 y Myrimatch 2.2.140 contra una base de datos compuesta



target/decoy construida a partir del proteoma de referencia de *Homo sapiens* en Uniprot knowledgebase, junto con contaminantes que aparecen con asiduidad. Después de la recalibración de las medidas del *patent ion mass* usando *high-scoring X!TANDEM hits*, los motores de búsqueda se configuraron a los potenciales candidatos peptídicos, con una tolerancia de error en la masa de 10 ppm y tolerancia del fragmento iónico de 0.02 Da. Se permitieron la pérdida de hasta 2 sitios de ruptura triptica y error del isotopo ( $^{13}\text{C}$ ) de 1. Se consideraron modificaciones MMTS fijas de cisteína y las modificaciones variables de oxidación de metionina, ácido piroglutámico a partir de glutamina o ácido glutámico en el extremo amino, acetilación de los extremos amino de la proteína y modificación de lisina, tirosina y el extremo amino peptídico con los reactivos de iTRAQ 8-plex. Se usó una False Discovery Rate (FDR)  $\leq$  0.01 (a nivel peptídico)<sup>223</sup>. Aproximadamente un 5% de las señales con la menor calidad se eliminaron antes del análisis. La regulación diferencial se midió usando modelos lineales<sup>224</sup> y la significancia estadística se midió usando q-values (FDR). Todos los análisis se realizaron con el software de Proteobotics (Madrid).

## 2.2 Inmunoprecipitación y purificación de péptidos asociados a moléculas de HLA

Los péptidos asociados a las moléculas de HLA se obtuvieron mediante cromatografía de afinidad, elución ácida y ultrafiltración. Todo el proceso se llevó a cabo a 4°C. Para ello, se utilizaron los *pellets* celulares de  $5 \times 10^8$  células de la línea HEK293 y HEK293-AIRE y se lisaron en un *buffer* de lisis (*Lysis buffer (LB)*): 20mM Tris HCl pH 7.6, 150 mM NaCl, 1% Igepal CA-630, Inhibidor de proteasas (C0mplete)) y se mantuvieron en rotación durante 1.5h a 4°C. Posteriormente se centrifugó el lisado (1300 RCF, 10 min, 4°C). Se recuperó el sobrenadante y se ultracentrifugó (100000 RCF, 1h, 4°C). El sobrenadante de la ultracentrifugación se pasó por una pre-columna de *beads* de sefarosa previamente equilibrada con *LB*. El *flow through* se pasó por una columna de *beads* de sefarosa acopladas al anticuerpo monoclonal específico W6/32 (anti-HLA-I).

Se realizaron tres lavados: un primer lavado con 50mL de *Wash buffer* (20mM Tris HCl pH 7.6, 150 mM NaCl, 0.5% NP40, Inhibidor de proteasas (CComplete)); un segundo lavado con 200mL de *Wash buffer* (20mM Tris HCl pH 7.6, 150 mM NaCl, 1% NP40, Inhibidor de proteasas (CComplete)); y un tercer lavado con 500mL de *Wash buffer* (20mM Tris-HCl pH 7.6, 5mM NaCl Inhibidor de proteasas (CComplete)). Los complejos péptidos-HLA se eluyeron con 30 mL de 0.1% ácido trifluoroacético (TFA) (Sigma-Aldrich<sup>TM</sup>), en fracciones de 1 mL. Se midió la presencia de proteínas mediante tinción de Bradford. Las fracciones que contenían proteína se evaporaron en un SpeedVac, se juntaron y se purificaron por ultrafiltración. La fracción peptídica se concentró en un SpeedVac. Las muestras se desalaron en un ZipTip (C18), se secaron en un SpeedVac y se congelaron en un SpeedVac hasta su uso.

### 2.2.1 Análisis por Espectrometría de masas

Las muestras se resuspendieron en 0.1% ácido fórmico (Sigma-Aldrich<sup>TM</sup>) y se analizaron por LC-MS/MS en un Ultimate 3000 nano-LC  $\mu$ HPLC (Thermo) acoplado online con un espectrómetro de masas Orbitrap Exploris 240 (Thermo) a través de una fuente iónica EASY-Spray (Thermo). El  $\mu$ HPLC se equipó con una *trapping column* Pepmap 300 C18 (5 $\mu$ m, 0.3 x 5 mm, Thermo) y una columna Pepmap RSLC C18 (2 $\mu$ m; 75 $\mu$ m x 50 cm, Thermo). Los solventes A y B fueron 0.1% ácido fórmico en agua y 0.1% ácido fórmico en acetonitrilo, respectivamente. Los péptidos se separaron a un flujo de 250 nL/min a 50°C bajo las siguientes condiciones de elución: 4% B durante 3 minutos, un incremento lineal hasta 35% B en 120 min, un incremento lineal hasta 90% B en 1 min, 90% B durante 5 min, una disminución lineal hasta 4% B en 1 min y 4% B durante 20 min. Cada ciclo de adquisición consistió en un *survey scan* a una resolución de 60000 (FWHM), 350 and 1200 m/z, y un máximo de 25 espectros MS2 a una resolución de 15000 (FWHM) generados con una colisión de energía del 30%. Solo se seleccionaron para ser fragmentadas aquellas especies con cargas 1+ a 5+.

### 2.2.2 MS/MS Ion Search and Peptide identification

Los datos de MS/MS crudos se convirtieron a archivos mgf con el programa Proteome Discoverer 2.5 (Thermo) y se buscaron con Peaks Studio 7.5 (Bioinformatics Solutions) contra el proteoma humano de referencia descargado de Uniprot. Se seleccionaron los siguientes parámetros: no enzima, tolerancia MS de 15 ppm, tolerancia MS/MS de 0.02 Da. Como modificaciones variables se seleccionaron la oxidación de la metionina y la formación de ácido piroglutámico a partir de glutamina. Las identificaciones se filtraron a una FDR < 1% a nivel peptídico.

## 2.3 Análisis de proteínas y péptidos

El análisis posterior de los datos de proteómica y peptidómica fue realizado comparando las localizaciones y características de las proteínas inferidas en el análisis mediante iTRAQ y las proteínas asociadas a cada péptido mediante las bases de datos UniProt (<https://www.uniprot.org/>), Human Protein Reference Database (<https://www.hprd.org/>) y The Human Protein Atlas (<https://www.proteinatlas.org/>). Se seleccionaron aquellos péptidos con longitudes entre 8y 12 aminoácidos. Se eliminaron los péptidos duplicados y las proteínas duplicadas (distintos péptidos asociados a una misma proteína) para el análisis. La localización de cada proteína se estableció según predominase en núcleo, citoplasma, membrana, retículo endoplásmico (ER), mitocondria, vesículas, Golgi, extracelular o no determinado (ND).

### III. Anticuerpos

#### 3.1 Anticuerpos primarios

Anticuerpos Primarios
anti-c-Myc mouse monoclonal IgG1 (kappa light chain) antibody (Santa Cruz Biotechnology™)
anti-CacyBP (D-8): sc-166195 mouse monoclonal IgG2a antibody (Santa Cruz Biotechnology™)
anti-Flag M2 mouse monoclonal IgG1 antibody (Sigma-Aldrich™)
anti-DYKDDDDK Tag (8H8L17) rabbit monoclonal antibody (Invitrogen™)
anti-6x His Tag (MA1-21315) mouse monoclonal antibody (Invitrogen™)
anti-AIRE-1 (PA5-78747) rabbit polyclonal IgG antibody (Invitrogen™)
anti-Vinculina (MA5-11690) mouse monoclonal IgG antibody (Invitrogen™)
anti-Caldesmon mouse monoclonal IgG1 antibody (Chemicon™)
anti-HSP70 mouse monoclonal IgG1 antibody (BD Biosciences™)
anti-Fascina mouse monoclonal antibody (Chemicon™)
anti-Citokeratina 18 (CKT18) mouse monoclonal antibody (Chemicon™)
anti-β-catenina (CAT-5H10) mouse monoclonal antibody (Invitrogen™)
anti-Miosina 9 (MYH9) rabbit polyclonal antibody (Invitrogen™)
anti-Miosina 10 (MYH10) (GT246) mouse monoclonal antibody (Invitrogen™)
anti-DNA-PK (SAB4502385) rabbit polyclonal antibody (Sigma-Aldrich™)
anti-Ubiquitina (U5379-1VL) rabbit polyclonal antiserum (Sigma-Aldrich™)
anti-HLA-I (W6/32) mouse monoclonal antibody (Hibridoma)
anti-MHC Clase I (HLA-I) APC mouse monoclonal IgG1 antibody (MACS™)
anti-CD44 (HDCAM) FITC mouse monoclonal IgG1 antibody (BD Biosciences™)
anti-CD49d (Integrina α4) PE mouse monoclonal IgG1 antibody (BD Biosciences™)
anti-CD54 (ICAM-1) PE mouse monoclonal IgG1 antibody (BD Biosciences™)
anti-CD95 (FAS) PE mouse monoclonal IgG1 antibody (BD Biosciences™)
anti-CD209 (ICAM-3) PE mouse monoclonal IgG1 antibody (eBiosciences™)

#### 3.2 Anticuerpos secundarios

Anticuerpos Secundarios
ECL anti-Mouse IgG, Horseradish Peroxidase-Linked whole antibody from sheep (GE Healthcare™)
Precision Protein™ StrepTactin-HRP Conjugate 5,000x (BioRad™)
ECL anti-Rabbit IgG, Horseradish Peroxidase-Linked whole antibody from donkey (GE Healthcare™)
Alexa Fluor 488 anti-Mouse IgG, goat polyclonal antibody IgG (H+L) (Invitrogen™)
Alexa Fluor 488 anti-Rabbit IgG, goat polyclonal antibody IgG (H+L) (Invitrogen™)
Alexa Fluor 568 anti-Mouse IgG, goat polyclonal antibody IgG (H+L) (Invitrogen™)
Alexa Fluor 568 anti-Rabbit IgG, goat polyclonal antibody IgG (H+L) (Invitrogen™)

## IV. Electroforesis y *Western blot*

Las muestras proteicas se separaron mediante electroforesis en condiciones desnaturalizantes (SDS-PAGE) en geles de poliacrilamida al 12% o al 10%. La concentración proteica de los lisados celulares se calculó mediante el kit de cuantificación BCA (Pierce™ BCA Protein Assay Kit (Thermo Scientific™)) y un lector de placas por colorimetría (Victor3™ Plate Reader (PerkinElmer™)). Las proteínas se transfirieron a membranas de polivinilideno (Immun-Blot® PVDF Membrane (BioRad™)) activadas previamente 5 min con metanol durante 45min a un voltaje de 100V. Posteriormente las membranas se incubaron en solución de bloqueo (T-PBS (PBS, 0.1% Tween20 (Panreac™)), 5% leche en polvo desnatada) durante 1h en agitación suave. Tras el proceso de bloqueo, las membranas se lavaron un mínimo de 3 veces con T-PBS durante 5 min en agitación. Seguidamente se incubó con el anticuerpo primario a una dilución 1/1000 en T-PBS. Las membranas se incubaron durante toda la noche en agitación a 4°C. Tras lavar 3 veces con T-PBS se añadieron los anticuerpos secundarios a una concentración de 1/10000 diluidos en T-PBS y se incubaron las membranas durante 1h en agitación a temperatura ambiente.

Finalmente las membranas se lavaron 4 veces con T-PBS durante 5 min en agitación. Tras el último lavado se aplicó la solución de revelado (1:1 kit Clarity Western ECL Blotting Substrate (BioRad™)) y las proteínas se detectaron mediante quimioluminiscencia a través del Molecular Imager® VersaDoc™ (BioRad™) y el *software* QuantityOne. Cada membrana se expuso durante 20 min tomando un total de 10 fotografías. Para los análisis de Densitometría se utilizó el *software* Fiji (ImageJ).

## V. Citometría de flujo

Para ensayos de citometría se empleó el citómetro FACScalibur™ (BD-Biosciences) o FACSCanto™ (BD-Biosciences) para la adquisición de datos y el software FlowJo v10

para el análisis de los mismos. Se partió de una suspensión celular de  $2 \times 10^6$  células en PBS. Tras un proceso de tinción intracelular o tinción de superficie, se seleccionó la región R1 correspondiente a las células vivas y singletes del análisis de tamaño celular y complejidad (FCS/SSC). Se determinó el grado de expresión de las proteínas de interés en transfectantes transitorios y estables mediante el análisis de fluorescencia en el canal FL1, FL2 o FL3 en función del anticuerpo secundario empleado o del agente intercalante fluorescente.

## 5.1 Tinción intracelular

Las células se tripsinizaron y se plaquearon 200 $\mu$ L (a una concentración de  $2.5 \cdot 10^6$  cél/mL) en los pocillos requeridos de placas p96 de fondo cónico (Corning). Las placas se centrifugaron (840 RCF, 2 min, RT) y se eliminó el sobrenadante. La placa se agitó con el vórtex brevemente para después resuspender las células en 200 $\mu$ L de PBS, las placas se centrifugaron de nuevo (840 RCF, 2 min, RT), se eliminó el sobrenadante y las células se fijaron durante 20 min con 200 $\mu$ L de PBS paraformaldehído (PFA) 3.7% (Sigma-Aldrich). Tras la incubación, se centrifugaron (840 RCF, 2 min, RT) y se realizaron dos lavados con 200 $\mu$ L de PBS y un lavado con 200 $\mu$ L de Solución de Tinción Intracelular (PBS, 0.5% Triton X100, 2%FBS, 0.05% Azida Sódica (Panreac). Se incubaron las células con 50 $\mu$ L de Solución de Tinción + anticuerpo primario pertinente (a la dilución adecuada) o solamente con 50 $\mu$ L Solución de Tinción como control interno negativo durante 45 min durante 1h a RT. Tras la incubación se realizaron dos lavados con 200 $\mu$ L de Solución de Tinción y se incubaron en oscuridad con 50 $\mu$ L de Solución de Tinción + anticuerpo secundario pertinente (a una dilución 1/200) durante 45 min, 1h a RT. Finalmente se realizaron tres lavados con 200 $\mu$ L de Solución de Tinción un lavado con 200 $\mu$ L de PBS. Las células se resuspendieron en 400 $\mu$ L de PBS en tubos de citometría.

## 5.2 Tinción celular de superficie

Las células se tripsinizaron y se plaquearon en 200 $\mu$ L (a una concentración de  $2.5 \cdot 10^6$  cél/mL) en los pocillos requeridos de placas p96 de fondo cónico. Las placas se centrifugaron (840 RCF, 2 min, 4°C) y se eliminó el sobrenadante. Las células se resuspendieron en 200 $\mu$ L de PBS frío y se centrifugaron de nuevo. Se descartó el sobrenadante y se realizaron dos lavados con 200 $\mu$ L de PBS, 2% FBS frío. Se incubaron las células con 50 $\mu$ L de Solución de Tinción + anticuerpo primario pertinente (a la dilución adecuada) o solamente con 50 $\mu$ L de PBS, 2% FBS como control interno negativo durante 20 min en hielo. Tras la incubación se realizaron dos lavados con 200 $\mu$ L de PBS, 2% FBS frío y se incubaron en oscuridad con 50 $\mu$ L de PBS, 2% FBS + anticuerpo secundario pertinente (a una dilución 1/200) durante 20 min en hielo. Finalmente, se realizaron tres lavados con 200 $\mu$ L de PBS, 2% FBS frío y un lavado con 200 $\mu$ L de PBS frío. Las células se resuspendieron en 400 $\mu$ L de PBS frío en tubos de citometría.

## 5.3 Ensayo de apoptosis

Se plaquearon  $1 \times 10^5$  células/pocillo en placas p6. Tras 48h se aplicaron 50 $\mu$ M de etopósido (Sigma-Aldrich™) y el equivalente en DMSO para la condición control durante 24h a 37°C 5% CO<sub>2</sub> con la finalidad de inhibir la síntesis de ADN favoreciendo la muerte celular programada. El ensayo de apoptosis se realizó siguiendo las indicaciones del kit para la detección de apoptosis con Annexina V-Alexa Fluor 488 y yoduro de propidio (IP) (ThermoFisher Scientific™): Las células se tripsinizaron de manera habitual recogiendo el medio de cultivo en tubos de 15mL para concentrar las células apoptóticas (300 RCF, 5 min, 4°C). Se descartó el sobrenadante y se realizó un lavado con 5mL de PBS 1X frío (300RCF, 5 min, 4°C). Las células se resuspendieron en 200 $\mu$ L de *buffer* de unión a Annexina V diluyéndolas a  $2 \cdot 10^6$  cél/mL. Se distribuyeron 50 $\mu$ L de la suspensión celular en tubos de citometría y se resuspendieron con 50 $\mu$ L de *buffer* de

unión a Annexina V con una dilución 1/40 de Annexina V-Alexa Fluor 488 y se incubaron a temperatura ambiente y en oscuridad durante 15 min. Posteriormente se añadió a la suspensión 1/100 de IP (100µg/mL). Finalmente, las diferentes suspensiones celulares se diluyeron con 400µL de *buffer* de unión a Annexina V adicionales y se analizaron las muestras por citometría de flujo.

## 5.4 Análisis del ciclo celular

Se analizaron  $2 \times 10^5$  células/condición. Las células no sincronizadas fueron resuspendidas mediante tripsinización y fueron lavadas dos veces con 10mL de PBS frío (300 RCF, 5 min, 4°C). Las células fueron resuspendidas en 100µL de PBS frío y se fijaron gota a gota sobre un tubo de citometría con 900µL de 70% etanol frío en agitación mediante vórtex. Una vez resuspendidas en la solución de 70% EtOH se traspasaron a tubos Eppendorf y se mantuvieron al menos 2h a -20°C para su fijación. En el momento del análisis las células fijadas se lavaron dos veces con 1mL de PBS frío. Finalmente, se descartó el sobrenadante y las células se resuspendieron con 250µL de solución de marcaje de DNA (PBS, 0.1% Triton X100; 0.2mg/mL RNasa A; 2% de Yoduro de Propidio a 1mg/mL) durante 15 min 37°C en oscuridad. Las diferentes condiciones fueron analizadas en el canal FL3 del citómetro.

## VI. Microscopía

Se emplearon el microscopio confocal Olympus Fluoview 1000 para el análisis de inmunofluorescencias, el microscopio confocal Leica TCS SP5 para el análisis de co-localización y el microscopio de epifluorescencia NIKON Ts2R para el análisis de migración (*wound healing assay*). Las imágenes fueron procesadas mediante el software Fiji (ImageJ).



## 6.1 Inmunofluorescencias

Se plaquearon aproximadamente  $5 \times 10^4$  cél/pocillo en una placa p24 (Corning) en la que previamente se dispusieron cubreobjetos estériles en 1mL de DMEM, 10% FBS. Tras 24h, se procedió a la tinción intracelular. Cada pocillo se lavó tres veces con 700µL de PBS y se fijaron las células con 500µL de PBS PFA 3.7% durante 20 min. Se realizan tres lavados con 700µL de PBS y dos lavados con 700µL con Solución de Tinción Intracelular (PBS, 0.5% Triton X100, 2% FBS y 0.05% Azida sódica) y se incubaron con 400µL de Solución de Tinción Intracelular + anticuerpo primario/s (a una dilución adecuada) durante 1h a RT. Pasado ese tiempo se realizaron cuatro lavados con 700µL de Solución de Tinción Intracelular y se incubaron las células en oscuridad con 400µL de Solución de Tinción Intracelular + anticuerpo secundario/s (a una dilución 1/400) durante 1h a RT. Finalmente se lavó cada pocillo tres veces con 700µL de Solución de Tinción Intracelular y dos veces con 700µL de PBS. Cada cubreobjetos se fijó a través de la Solución de montaje antidecoloración ProLong® Gold con DAPI (Thermo Fisher Scientific) sobre un portaobjetos, se dejó secar O/N y se conservaron en oscuridad a 4°C hasta su análisis.

## 6.2 Ensayo de migración (*Woung Healing Assay*)

Se plaquearon  $1 \cdot 10^6$  cél. según las condiciones requeridas en placas de cultivo p6. 24h después una vez las células alcanzaron la máxima confluencia se realizaron heridas con un raspador celular o *cell scrapper* (ThermoFisher). Se lavaron las células con 2mL de PBS y se añadió DMEM sin suero o DMEM 10% FBS, según la condición establecida. Se tomaron imágenes representativas (n=30) a lo largo de las heridas, nada más realizar la herida (t=0h) y a las 24, 48 y 72h. Se calcularon las distancias medias de cada imagen tomando tres valores de distancia (superior, media e inferior) de cada imagen con el software de procesamiento de imágenes ImageJ (Fiji).

## VII. Inmunoprecipitación

### 7.1 Inmunoprecipitación con *beads* de sefarosa

Las células transfectadas se traspasaron a tubos Falcon de 15mL y se centrifugaron a 300 RCF durante 5 min a 4°C. Las células se incubaron previamente a 20µM del inhibidor de proteasoma MG-132 (Sigma-Aldrich™) durante 6h a 37°C y 5% CO<sub>2</sub> con la finalidad de inhibir la ruta de degradación vía proteasoma. Los *pellets* generados se lavaron con PBS frío. Los precipitados celulares se resuspendieron en 1mL de PBS frío y se traspasaron a tubos Eppendorf de 1.5mL y se centrifugaron de nuevo. El *pellet* se resuspendió en 200µL de *buffer* de lisis celular de Inmunoprecipitación (*IP Lysis buffer*: 10mM HEPES pH 7.4 (Sigma-Aldrich™), 1.5 mM MgCl<sub>2</sub> (Sigma-Aldrich™), 10mM KCl (Sigma), 5mM EDTA (Sigma-Aldrich™), 1% NP40 (Fluka Biochemika™), Inhibidor de proteasas (CComplete Tablets Mini, Roche™)) y se incubó durante 15 min en hielo. Posteriormente, los lisados celulares se centrifugaron (1000 RCF, 10 min, 4°C) y el *pellet* se resuspendió en 200µL de *buffer* de extracción nuclear (*IP Nuclear Extraction buffer*: 20mM HEPES pH 7.4, 300mM NaCl, 20mM KCl, Inhibidor de proteasas (CComplete)) y se incubó durante 1 hora en hielo. A continuación, las muestras se centrifugaron (13000 RCF, 10 min, 4°C) y el sobrenadante se traspasó a nuevos tubos Eppendorf. Se guardaron 10µL de cada muestra como lisado total (INPUT) del experimento. El resto de los volúmenes de las muestra se diluyeron en 800µL de *buffer* de lisis celular sin NP40 (*IP Wash buffer*: 10mM HEPES pH 7.4, 1.5 mM MgCl<sub>2</sub>, 10mM KCl, 5mM EDTA). Las muestras se incubaron a 4°C con el anticuerpo correspondiente durante 4h en rotación.

Durante ese período de tiempo se acondicionaron las *beads* de proteína G-sefarosa según instrucciones de la casa comercial (Protein G-sepharose, GE Healthcare™). Pasado el tiempo de incubación con el anticuerpo, se añadió la proteína G-sefarosa y se incubó en rotación durante 2h o toda la noche (*overnight*, O/N). Finalmente, las muestras se centrifugaron (4000 RCF, 2 min, 4°C) y los precipitados se lavaron 3 veces con *IP Wash buffer* y 1 vez con PBS 1X frío usando el volumen necesario hasta completar 1mL del tubo Eppendorf y en las mismas condiciones de centrifugación

inicial. Por último, tras la última centrifugación, se resuspendieron los precipitados en 200  $\mu$ L de *buffer* de carga de electroforesis (RBS: 24mM Tris-HCl pH 6.8, 10% Glicerol (Sigma-Aldrich<sup>TM</sup>), 5.76mM 2-Mercaptoetanol (Sigma-Aldrich<sup>TM</sup>), 0.8% SDS (Sigma-Aldrich<sup>TM</sup>), 0.04% Azul de Bromofenol (Fluka Biochemika<sup>TM</sup>), se centrifugaron brevemente y se conservaron a -20°C hasta el momento de realizar el Western Blot. Este último paso se realiza igualmente para las muestras INPUT generados con cada experimento.

## 7.2 Inmunoprecipitación con *beads* magnéticas

### 7.2.1 Inmunoprecipitación con *beads* magnéticas

Las células transfectadas se traspasaron a tubos Falcon de 15mL y se centrifugaron a 300 RCF durante 5 min a 4°C. Las células se incubaron previamente a 20  $\mu$ M del inhibidor de proteasoma MG-132 (Sigma-Aldrich<sup>TM</sup>) durante 6h a 37°C y 5% CO<sub>2</sub>. Los *pellets* generados se lavaron con PBS frío. Los precipitados celulares se resuspendieron en 1mL de PBS frío y se traspasaron a tubos Eppendorf de 1.5mL y se centrifugaron de nuevo. El *pellet* se resuspendió en 200  $\mu$ L de *buffer* de lisis celular de Inmunoprecipitación ((*IP Lysis buffer*: 50Mm Tris-HCl pH7.5, 150Mm NaCl, 0.5% NP40, 1% Triton X100 (Sigma-Aldrich<sup>TM</sup>), Inhibidor de proteasas) frío y se incubaron durante 1h en hielo. Durante este tiempo se acondicionaron las *beads* magnéticas con anticuerpos anti-c-Myc (Pierce<sup>TM</sup> Anti-c-Myc Magnetic Beads (Thermo Scientific<sup>TM</sup>)) o anti-FLAG (Anti-FLAG<sup>®</sup> M2 Magnetic Beads (MERCK<sup>TM</sup>)) según instrucciones de la casa comercial. Tras el proceso de lisis y acondicionamiento de las *beads* magnéticas, las muestras se centrifugaron (13000 RCF, 10 min, 4°C) y el sobrenadante se traspasó a nuevos tubos Eppendorf: se guardaron 20  $\mu$ L de cada muestra como lisado total (INPUT) del experimento, se resuspendieron en 20  $\mu$ L de RBS, se centrifugaron brevemente y se conservaron a -20°C hasta su análisis; el resto del volumen (180  $\mu$ L) de cada muestra se resuspendió en los tubos Eppendorf con las *beads* magnéticas acondicionadas. Se

incubaron 30 min en rotación a temperatura ambiente (RT). Tras el proceso de captura, las muestras se lavaron tres veces con 300µL de *IP Wash buffer* frío y una vez con 300µL de agua MilliQ<sup>®</sup> fría usando el separador magnético. Finalmente, las *beads* magnéticas de cada muestra se resuspendieron en 40µL de RBS y se hirvieron durante 15 min a 98°C. El volumen de RBS con el contenido de la captura de las *beads* magnéticas se recuperó usando el separador magnético. Las muestras se conservaron a -20°C hasta el momento de realizar el *Western Blot*.

### 7.2.2 Inmunoprecipitación con paraformaldehído y *beads* magnéticas

Las células se traspasaron a tubos Falcon de 15mL y se centrifugaron (300 RCF, 5 min, 4°C). Cuando se requirió, las células se incubaron previamente a 20µM de MG-132 durante 6h a 37°C 5% CO<sub>2</sub>. Los *pellets* generados se resuspendieron en 10mL de PBS frío y se centrifugaron de nuevo. Los precipitados celulares se resuspendieron en 2mL de PBS, 0.5% PFA y se incubaron 5 min a RT. El proceso de fijación se bloqueó añadiendo 10mL de PBS, 1.25M Glicina (Sigma-Aldrich<sup>TM</sup>) frío y se centrifugaron (300 RCF, 5 min, 4°C). Los precipitados celulares se resuspendieron en 1mL de PBS frío y se traspasaron a tubos Eppendorf de 1.5mL. Las muestras se centrifugaron de nuevo, se descartó el sobrenadante y se añadió a cada tubo 200µL de *buffer* de lisis celular de inmunoprecipitación (*IP Lysis buffer*: 50mM Tris-HCl pH7.5, 150mM NaCl, 1mM EDTA, 1% NP40, 0.1% SDS (Fisher Scientific<sup>TM</sup>), Inhibidor de proteasas) frío y se incubaron durante 1h en hielo. Las muestras se sonicaron a 60Hz durante 10 seg, 3 ciclos en total.

Durante este tiempo se acondicionaron las *beads* magnéticas con anticuerpos anti-c-Myc o anti-FLAG. Las muestras se centrifugaron (13000 RCF, 10 min, 4°C) y el sobrenadante se traspasó a nuevos tubos. 20µL de cada muestra se guardaron como lisado total (INPUT) del experimento, se resuspendieron en 20µL de RBS, se centrifugaron brevemente y se conservaron a -20°C hasta su análisis. El resto del volumen (180µL) de cada muestra se resuspendió en los tubos Eppendorf con las *beads*

magnéticas acondicionadas. Se incubaron O/N en rotación a RT. Las muestras se lavaron una vez con 300µL de *IP Lysis buffer* y una vez con 300µL de *IP Wash buffer* frío usando el separador magnético. Finalmente, las *beads* magnéticas de cada muestra se resuspendieron en 40µL de RBS y se hirvieron durante 15 min a 98°C. El volumen de RBS con el contenido de la captura de las *beads* magnéticas se recuperó usando el separador magnético. Las muestras se conservaron a -20°C hasta el momento de realizar el *Western Blot*.







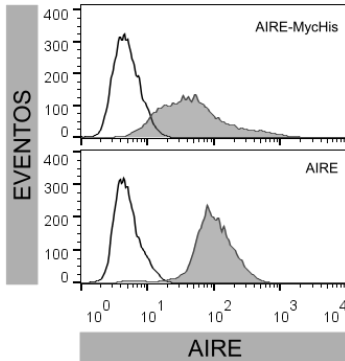


# RESULTADOS

## CAPÍTULO I: Análisis proteómico diferencial en transfectantes estables de AIRE en HEK293

### 1.1 Generación de un modelo celular de expresión estable de AIRE para estudiar los proteomas y los peptidomas de HLA-I

Se clonó el gen de AIRE completo asociado o no a una *tag* de Myc/His en su extremo C-terminal en el vector de expresión en células eucariotas pcDNA3.1. Se generaron diferentes transfectantes estables en células HEK293 con ambas construcciones. Los transfectantes se clonaron por dilución límite, con lo que se generó una serie de clones celulares de expresión estable de AIRE con una única población positiva (**Figura 1**).



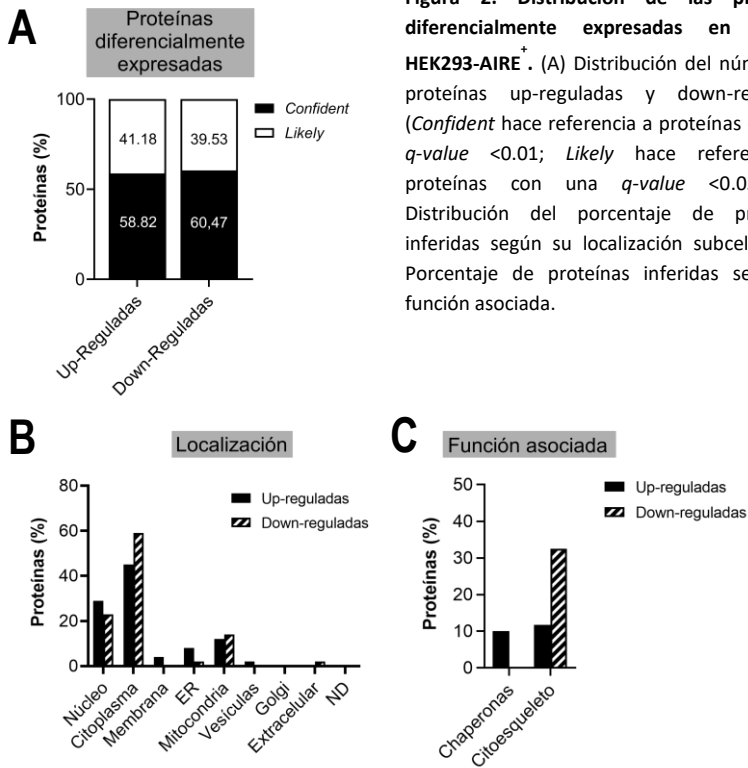
**Figura 1.** AIRE WT y AIRE-MycHis se expresan de manera estable en transfectantes de HEK293. Los niveles de expresión de AIRE para cada transfectante estable se midieron por citometría de flujo. El histograma sin rellenar muestra la fluorescencia asociada a la expresión de AIRE en células HEK293 sin transfectar como control. (AIRE=FITC)

### 1.2 Impacto de la expresión de AIRE en el proteoma celular (proteómica cuantitativa)

Para analizar si los efectos producidos por la expresión de AIRE en el proteoma celular descrito previamente<sup>110</sup> era debido exclusivamente a AIRE se analizó por *Isobaric tags*

*for relative and absolute quantitation* (iTRAQ) los proteomas de la línea HEK293 y el transfectante que expresaba el gen AIRE humano asociado al *tag Myc/His*. El análisis permitió la identificación de 3928 proteínas. El análisis estadístico definió grupos de proteínas con expresión incrementada o disminuida en las células AIRE<sup>+</sup> respecto a las células control con diferentes grados de probabilidad estadística. Así, se encontraron 30 proteínas incrementadas y 26 disminuidas con alta probabilidad (*confident up or confident down* ( $q\text{-value} < 0.01$ , *positive or negative log fold change*)), 21 probablemente incrementadas y 17 probablemente disminuidas (*likely up or likely down* ( $q\text{-value} < 0.05$ , *positive or negative log fold change*)) y 414 y 296 supuestamente incrementadas o disminuidas, respectivamente (*putative up or putative down* ( $q\text{-value} > 0.05$  but  $p\text{-value} < 1-\pi_0$ , *positive or negative log fold change*)) en los transfectantes de AIRE. Las 3119 proteínas restantes se detectaron sin verse estadísticamente incrementadas o disminuidas. Aun siendo conscientes de la posible pérdida de información, nos hemos centrado en las proteínas con un valor de  $q\text{-value}$  inferior a 0.05 (Tablas 1 y 2). Las proteínas con la calificación *putative up or putative down* solo se han evaluado por el interés de la misma en otras observaciones. Así, la proteína CACYBP (proteína de interés analizada posteriormente en el Capítulo III) se detectó como *putative up*. Teniendo en cuenta únicamente aquellas proteínas diferencialmente expresadas con una  $q\text{-value} < 0.05$ , el 58.8% de las incrementadas o up-reguladas y el 60.5% de las disminuidas o down-reguladas se detectaron como “*confident*” y el resto como “*likely*” (Figura 2A). Se determinó la localización de las proteínas diferencialmente expresadas. La localización de las proteínas se definió utilizando las bases de datos: UniProt, Human Protein Reference Database y The Human Protein Atlas. Cuando una proteína se encontraba localizada en diferentes compartimentos celulares, se eligió el más común asociado a su función biológica. Las proteínas diferencialmente expresadas se distribuyeron principalmente en núcleo, citoplasma y mitocondria (Figura 2B). La mayoría de las proteínas con expresión modificada eran principalmente citoplásmicas. De ellas, el porcentaje de proteínas disminuidas en las células AIRE<sup>+</sup> era mayor en proteínas pertenecientes a citoplasma y mitocondria. En cambio, entre las proteínas incrementadas en el transfectante de AIRE, había más que se localizaban en el núcleo, ER y membrana plasmática (Figura 2B).

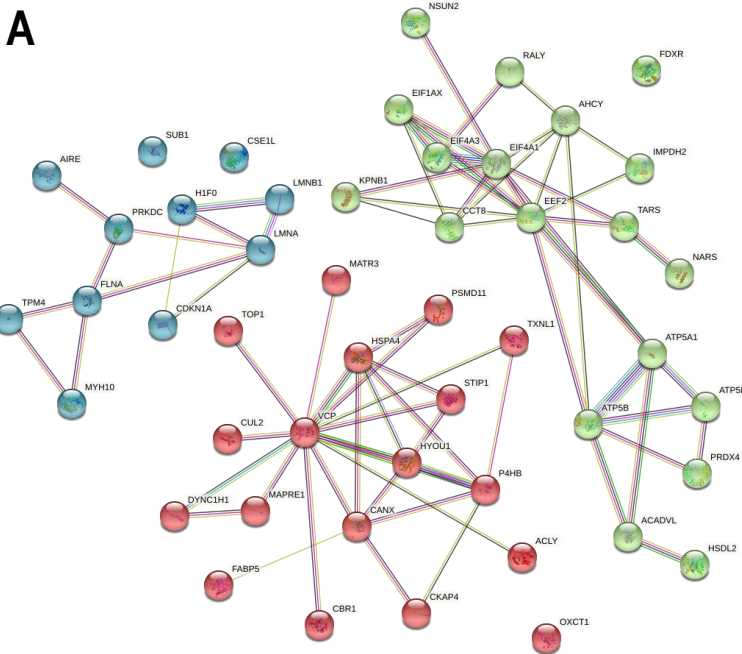
Entre las proteínas incrementadas se encontraron principalmente chaperonas y entre las disminuidas predominaban proteínas relacionadas con el citoesqueleto (**Figura 2C**, **Tabla 1**). Estos datos están en consonancia con el análisis previo con otras técnicas de proteómica cuantitativa realizado en nuestro laboratorio usando la línea epitelial de tiroides HT93<sup>110</sup>.

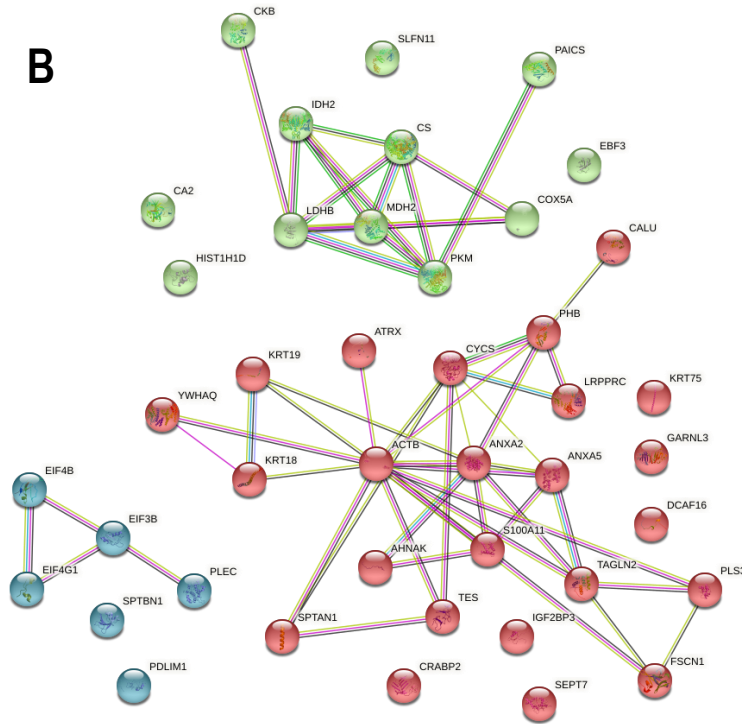


**Figura 2. Distribución de las proteínas diferencialmente expresadas en células HEK293-AIRE<sup>+</sup>.** (A) Distribución del número de proteínas up-reguladas y down-reguladas (*Confident* hace referencia a proteínas con una *q-value* <0.01; *Likely* hace referencia a proteínas con una *q-value* <0.05). (B) Distribución del porcentaje de proteínas inferidas según su localización subcelular (C) Porcentaje de proteínas inferidas según su función asociada.

Se analizaron las proteínas según sus interacciones usando el *software* STRING. Dicho *software* predice interacciones proteína-proteína, incluyendo asociaciones directas (físicas) e indirectas (funcionales) obtenidos de varias fuentes (genómica, datos de experimentos *high-throughput*, co-expresión y literatura). En general, las proteínas que ven incrementada su expresión están relacionadas con procesos biológicos de traducción, remodelación del DNA, plegamiento de proteínas y apoptosis (**Figura 3A** y **Tabla 3**, parte superior). Las proteínas que ven disminuida su expresión están

relacionadas con la generación de uniones celulares y citoesqueleto y el metabolismo aeróbico (Figura 3B y Tabla 3, parte inferior). Entre las proteínas incrementadas, se encontraron las topoisomerasas 1 y 2A y la kinasa DNA-PK. También se ve aumentada la expresión de factores como el factor de elongación 2 (EF2), el factor de iniciación de traducción IF1AX (EIF1AX) o el factor de traducción eucariótico 4A1 (EIF4A1). Estos datos sugieren una potenciación de la actividad de remodelación de la estructura del DNA propia de la actividad de transcripción, puesto que dichos elementos participan activamente en el proceso. Asimismo, el incremento en estas proteínas es congruente con la actividad biológica de AIRE, puesto que tanto TOP1 y TOP2A y DNA-PK establecen una interacción con AIRE<sup>83 225</sup>. El incremento de proteínas como la Subunidad β1 de la Importina (KPN1) están relacionadas con procesos de apoptosis<sup>226</sup>. La disminución en la expresión de proteínas como transgelina, fascina o keratina 18, 19 y 75 sugieren procesos relacionados con la remodelación del citoesqueleto<sup>110</sup>.





**Figura 3. Modelización de la red de interacción de proteínas diferencialmente expresadas en células que expresan AIRE.** Las proteínas diferencialmente expresadas en las muestras de HEK293-AIRE<sup>+</sup> fueron analizadas mediante el software de predicción de interacciones STRING. (A) Red de interacción modelizada de proteínas up-reguladas. Se diferencian tres nodos: Topoisomerasa 1A y 2A (TOP1A y TOP2A) y DNA-PK (PRKDC); proteína de respuesta a choque a térmico Hsp74 (HSPA4) y proteína p97 (VCP); el factor de elongación 2 (EEF2) y factor de traducción eucariótico 4A1 (EIF4A1). (B) Red de interacción modelizada de proteínas down-reguladas. Se diferencian dos nodos principales centralizados por la malato deshidrogenasa 2 (MDH2), citrato sintasa (CS) y lactato deshidrogenasa B (LDHB);  $\beta$ -actina (ACTB). Se usó un algoritmo de *clustering* K-means=4.

Categoría	Log <sub>2</sub> Ratio	p Value	e Value	q Value	Número de réplicas	Número de péptidos	Identificador	Nombre del gen	Nombre de la proteína
Confident up	0.534	9.99999E-07	0	0	4	6	P07305	H1-0	Histona H1.0
Confident up	0.442	9.99999E-07	0	0	4	15	O43918	AIRE	Regulador Autoimmune
Confident up	0.251	9.99999E-07	0	0	4	27	P25705	ATP5F1A	Subunidad α ATP-sintasa mitocondrial
Confident up	0.224	9.99999E-07	0	0	4	15	P27824	CANX	Calnexina
Confident up	0.178	9.99999E-07	0	0	4	29	P07237	PDI1A	Disulfuro isomerasa
Confident up	0.161	9.99999E-07	0	0	4	27	P06576	ATP5F1B	Subunidad β ATP-sintasa mitocondrial
Confident up	0.106	9.99999E-07	0	0	4	31	P02545	LMNA	Prelamina A
Confident up	0.176	9.99999E-07	0	0.001	4	121	P78527	PRKOC	DNA-PK (Subunidad catalítica)
Confident up	0.139	9.99999E-07	0	0.001	4	36	P20700	LMNB1	Lamin-B1
Confident up	0.164	2E-06	0.01	0	4	38	P55060	XPO2	Exportina 2
Confident up	0.128	2E-06	0.01	0	4	23	P26639	TARS1	Treonina-tRNA ligasa 1
Confident up	0.23	4E-06	0.01	0.001	4	48	P35580	MYH10	Miosina no muscular 10 / Miosina no muscular IIB
Confident up	9.04	5E-06	0.02	0.001	1	14	P55809	OXCT1	SCOT
Confident up	0.213	8.99999E-06	0.03	0.001	1	1	P38936	CDKN1A	Inhibidor 1 de la kinasa dependiente de ciclina
Confident up	0.103	1.21803E-05	0.04	0.001	4	12	P67936	TPM4	Cadena α4 de tropomiosina
Confident up	0.217	1.32802E-05	0.05	0.001	4	43	P55072	TEFRA	ATPasa transicional del retículo endoplásmico
Confident up	0.14	1.74216E-05	0.06	0.002	4	14	O43776	SYNC	Asparagina-tRNA ligasa
Confident up	0.124	1.88679E-05	0.06	0.002	4	31	P34932	HSP44	Proteína de choque térmico 4 de 70kDa (HSP74)
Confident up	0.194	2.35294E-05	0.08	0.002	4	34	Q14974	KPN1A	Subunidad β1 de importina
Confident up	0.152	2.55101E-05	0.09	0.002	4	17	Q8Y4L1	HYOU1	Proteína 1 up-regulada por hipoxia
Confident up	0.286	2.77777E-05	0.1	0.002	4	22	Q08123	NSUN2	RNA citosina C(5)-metiltransferasa NSUN2
Confident up	0.248	3.70369E-05	0.13	0.003	4	9	P49748	ACADVL	acil-CoA deshidrogenasa específica de muy larga cadena (VLCAD)
Confident up	0.059	4.40527E-05	0.15	0.004	4	123	Q14204	DYNC1H1	Peroxi-redoxina 4
Confident up	0.284	5.74709E-05	0.2	0.005	4	8	O43396	TXN1L	Cadena pesada 1 de la dineína 1 citoplasmática
Confident up	0.081	8.26439E-05	0.28	0.006	4	25	P05090	CC18	Proteína 1 similar a torredoxina
Confident up	0.203	0.000138887	0.48	0.008	4	44	P13639	EF2	Subunidad β de la proteína 1 del complejo T
Confident up	0.133	0.000169489	0.58	0.01	4	11	P16152	CBR1	Factor de elongación 2
Confident up	0.195	0.000175436	0.6	0.01	4	16	P23526	SAHH	Carbamil reductasa (NADPH) 1
Likely up	0.097	0.000192979	0.66	0.011	4	15	P13388	TOP2A	Adenosilhomocisteína hidrolasa
Likely up	0.282	0.000196425	0.68	0.011	4	34	P53396	ACYL	Topoisomerasa Iα (TOP2A)
Likely up	0.203	0.000243896	0.84	0.013	4	6	P47813	E1F1AX	ATP-citrato sintasa
Likely up	0.218	0.000314277	1.08	0.017	4	12	Q13617	CUL2	Factor de iniciación de traducción eucariótico Iβ1AX
Likely up	0.107	0.00032257	1.11	0.016	4	9	P22570	FDXR	Cullin-2
Likely up	0.233	0.000379297	1.3	0.019	4	22	P43243	MATR3	NAPDH-Adrenodoxina oxidoreductasa mitocondrial
Likely up	0.121	0.000499975	1.72	0.023	4	8	P53999	SUB1	Matrin-3
Likely up	0.137	0.000499975	1.72	0.024	4	19	P11387	TOP1	Coactivador p15 de la RNA-polIII
Likely up	0.118	0.000555525	1.91	0.025	4	15	P12268	IMPDH2	DNA Topoisomerasa I (TOP1)
Likely up	0.095	0.000666663	2.29	0.029	4	24	P49419	ALDH7A1	IMP deshidrogenasa 2
Likely up	0.17	0.000733284	2.52	0.031	4	22	Q07065	CKAP4	Isoforma 2 de AASA-dehidrogenasa
Likely up	0.05	0.000909008	3.13	0.036	4	11	Q9UKM9	RALY	Proteína 4 asociada a citoesqueleto
Likely up	0.146	0.000909008	3.13	0.037	4	81	P21333	FLNA	Proteína de unión a RNA Raly
Likely up	0.138	0.0009999	3.44	0.038	4	13	Q135691	MAPRE1	Filamina A
Likely up	0.426	0.0009999	3.44	0.039	4	13	Q133110	PABP4	Proteína EB1 de unión a APC
Likely up	0.144	0.0009999	3.44	0.04	4	2	Q01469	FABP5	Proteína 4 de unión a Poly(A)
Likely up	0.216	0.00109989	3.78	0.041	4	2	P60842	E1F4A1	Proteína 4 de unión a ácidos grasos
Likely up	0.088	0.001249844	4.3	0.044	4	14	O75947	ATP5H	Factor de iniciación eucariótico 4A-1 (eIF-4A-1)
Likely up	0.192	0.001249844	4.3	0.046	4	8	P31948	STP1	Subunidad 6 ATP-sintasa mitocondrial
Likely up	0.112	0.001374828	4.73	0.048	4	35	Q6YM16	HSDL2	Fosfoproteína 1 inducida por estrés
Likely up					4	10	O00231	P5MD11	Proteína 2 similar a hidroxisteroide deshidrogenasa
Likely up					4	17			Subunidad reguladora RPNG del proteosoma 26S

Tabla 1. Lista de proteínas diferencialmente up-reguladas en el análisis proteómico mediante ITRAQ. Las proteínas consideradas como *confident up* presentan una q-value <0.01, *positive log fold change*; las proteínas consideradas como *likely up* presentan una q-value <0.05, *positive log fold change*.

Categoría	Log <sub>2</sub> Ratio	p Value	e Value	q Value	Número de réplicas	Número de péptidos	Identificador	Número del gen	Nombre de la proteína
Likely down	-0.124	0.00021387	0.75	0.012	4	21	P08758	ANXA5	Annexina A5
Likely down	-0.678	0.00032257	1.11	0.017	4	2	P46100	ATRX	Regulador transcripcional ATRX
Likely down	-0.199	0.00032257	1.11	0.017	4	11	P00151	PDLM1	Proteína 1 de dominios PDZ y LIM
Likely down	-0.088	0.000379297	1.3	0.019	4	60	Q15149	PLEC	Plectina
Likely down	-0.324	0.000384601	1.32	0.019	4	6	P29373	CRABP2	Proteína 2 de unión a ácido retinoico celular
Likely down	-0.555	0.000454525	1.56	0.022	4	1	P60709	ACTB	Actina
Likely down	-0.169	0.000499975	1.72	0.023	4	13	A0A1C7CYX9	DYSL2	Proteína 2 relacionada con la dhidropirimidinas
Likely down	-0.121	0.00054543	1.88	0.025	4	17	P35232	PHB	Prohibitina
Likely down	-0.356	0.000611077	2.1	0.027	4	5	Q7Z7L1	SLEFN1	Miembro 11 de la familia Schlaflan
Likely down	-0.142	0.000687457	2.36	0.029	4	14	Q16181	SEPT7	Septina 7
Likely down	-0.122	0.000687457	2.36	0.03	4	22	P55884	EIF3B	Subunidad B del factor 3 de iniciación de traducción eucariótico
Likely down	-0.19	0.000785658	2.7	0.033	4	8	O43852	CALU	Callunina
Likely down	-0.174	0.0009999	3.44	0.038	4	15	P48735	IDH2	Isocitratato deshidrogenasa (NADP) mitocondrial
Likely down	-0.101	0.0009999	3.44	0.039	4	26	Q04637	EIF4G1	Factor 4γ1 de iniciación de traducción eucariótico
Likely down	-0.126	0.00109989	3.78	0.041	4	16	O75390	CS	Citrato sintasa mitocondrial
Likely down	-0.109	0.001249844	4.3	0.045	4	15	P40926	MDH2	Malato deshidrogenasa mitocondrial
Likely down	-0.142	0.001374828	4.73	0.048	4	16	P23588	EIF4B	Factor 4B de iniciación de traducción eucariótico
Confident down	-0.661	9.99999E-07	0	0	4	37	P13797	PLS3	Plastina
Confident down	-0.612	9.99999E-07	0	0	4	26	Q16658	FSCN1	Fascina
Confident down	-0.611	9.99999E-07	0	0	4	6	P09318	CA2	Anhidrasa carbónica 2
Confident down	-0.549	9.99999E-07	0	0	4	4	P16402	H13	Histona H1.3
Confident down	-0.524	9.99999E-07	0	0	4	25	P05783	KRT18	Keratina 18 (Keratina tipo I citoesquelética 18)
Confident down	-0.3	9.99999E-07	0	0	4	25	P07355	ANXA2	Annexina A2
Confident down	-0.275	9.99999E-07	0	0	4	17	P12277	CKB	Creatina Kinasa
Confident down	-0.086	9.99999E-07	0	0	4	78	Q13813	SPTAN1	Cadena α de spectrina no eritrocítica
Confident down	-0.337	9.99999E-07	0	0.001	4	22	P08727	KRT19	Keratina 19 (Keratina tipo I citoesquelética 19)
Confident down	-0.156	9.99999E-07	0	0.001	4	39	P14618	PKM	Piruvato Kinasa PKM
Confident down	-0.281	9.99999E-07	0	0.002	4	4	P07195	LDHB	Cadena B de la L-lactato deshidrogenasa
Confident down	-0.198	9.99999E-07	0	0.003	4	140	Q09666	AHNAK	Proteína AHNAK asociada a la diferenciación de neuroblasto
Confident down	-0.273	4E-06	0.01	0.001	4	12	O95678	KRT75	Keratina 75 (Keratina tipo I citoesquelética 75)
Confident down	-0.208	1.4245E-05	0.05	0.002	4	13	O00425	IGF2BP3	Proteína 3 de unión a mRNA IGF2
Confident down	-0.106	2.07468E-05	0.07	0.002	4	48	P42704	LRPPRC	Proteína con motivo rico en leucinas y motivo PPR
Confident down	-0.311	2.29885E-05	0.08	0.002	4	7	P99999	CYC	Citocromo C
Confident down	-0.337	6.45157E-05	0.22	0.005	4	5	P31949	S100A11	S100-A11
Confident down	-0.267	6.47055E-05	0.22	0.005	4	7	P20674	COX5A	Subunidad 5A de la Citocromo C oxidasa
Confident down	-1.72	7.51874E-05	0.26	0.005	2	1	Q9HAW6	COE3	Co-factor transcripcional COE3
Confident down	-0.161	9.61529E-05	0.33	0.007	4	16	P37802	TAGLN2	Transgelina 2
Confident down	-0.222	9.70864E-05	0.34	0.007	4	10	Q9UGI8	TES	Testina
Confident down	-0.081	9.90089E-05	0.33	0.007	4	54	Q01082	SPTBN1	Cadena β de Spectrina no eritrocítica
Confident down	-0.901	9.9999E-05	0.34	0.006	4	1	Q9NKF7	DCAF16	Factor 16 asociado a DDB1 y CUL4
Confident down	-1.159	0.000103092	0.35	0.007	4	1	Q3VWV2	GARNL3	Proteína 3 con dominio similar a GTPasa activadora Rap/Ran
Confident down	-0.18	0.000104166	0.36	0.007	4	12	P27348	YWHAQ	14-3-3 β
Confident down	-0.178	0.000123455	0.42	0.008	4	14	P22234	PAICS	Proteína multifuncional ADE2

Tabla 2. Lista de proteínas diferencialmente down-reguladas en el análisis proteómico mediante iTRAQ. Las proteínas consideradas como *confident down* presentan una *p-value* < 0.01 *nonrativus Inn fold change* las proteínas consideradas como *likely down* presentan una *p-value* < 0.05 *nonrativus Inn fold change*

Función biológica incrementada	Strength	False Discovery Rate	Proteínas relacionadas de la red de interacción
Escisión embrionaria	2.19	0.0214	TOP1, TOP2A
Cambios topológicos en el DNA	1.94	0.0418	TOP1, TOP2A
Respuesta celular derivada de factor neurotrófico	1.89	0.0479	EIF4A3, EEF2
Síntesis de ATP mitocondrial acoplado a cadena de electrones	1.75	0.0084	ATP5B, ATP5H, ATP5A1
Apoptosis	1.67	0.0112	KPNB1, H1FO, TOP2A
Formación de cresta mitocondrial	1.58	0.0162	ATP5B, ATP5H, ATP5A1
Biosíntesis de bases nitrogenadas	1.53	0.0031	ATP5B, ATP5H, IMPDH2, ATP5A1
Respuesta celular a IL-7	1.45	0.0296	ATP5B, P4HB, STIP1
Respuesta a plegamiento de proteínas inducida por IRE-1	1.33	0.0500	LMNA, ACADVL, HYOU1
Organización del huso mitótico	1.29	0.0130	KPNB1, DYNC1H1, VCP, FLNA
Respuesta a plegamiento de proteínas del ER	1.27	0.0034	CANX, VCP, LMNA, ACADVL, HYOU1
Importe proteico al núcleo	1.21	0.0209	CSE1L, KPNB1, LMNA, CDKN1A
Respuesta a proteínas mal plegadas	1.15	0.0021	CANX, HSPA4, VCP, LMNA, ACADVL, HYOU1
Importe de proteínas	1.13	0.0098	CSE1L, KPNB1, HSPA4, LMNA, CDKN1A
Organización del huso mitótico	1.13	0.0099	KPNB1, DYNC1H1, VCP, FLNA, MAPRE1
Síntesis de ribonucleoproteínas	1.12	0.0103	ACLY, ATP5B, ATP5H, IMPDH2, ATP5A1
Respuesta a estrés del ER	1.03	0.0020	CANX, EEF2, P4HB, VCP, LMNA, ACADVL, HYOU1
Respuesta celular a hipoxia	1.02	0.0214	PSMD11, P4HB, LMNA, CUL2, HYOU1
Plegamiento de proteínas	0.96	0.0319	CANX, CCT8, P4HB, VCP, PRDX4

Función biológica reducida	Strength	False Discovery Rate	Proteínas relacionadas de la red de interacción
Unión a cadherina en adhesión célula-célula	2.02	0.0010	EIF3B, EIF4B, EIF4G1
Unión a proteína S100	2.0	0.0031	KRT75, PLEC, SPTBN1, ACTB, KRT19, SPTANI, AHNAK, KRT18, SEPT7
Actividad de factor transcripcional de iniciación	1.45	0.0412	HIST1H1D, IGF2BP3, LRPPRC, PKM, PLEC, MDH2, CS, ANXA2, SPTBN1, TES, EIF3B, AHNAK, FSCN1, KRT18...
Citoesqueleto	1.35	0.0023	PLEC, SPTBN1, ACTB, KRT19, SPTANI
Unión a cadherina	1.29	2.59e-11	LRPPRC, PLEC, ANXA2, PLS3, SPTBN1, ACTB, PDLM1, SPTANI, FSCN1
Regulación de actividad transcripcional	1.15	0.0412	S100A11, PKM, PLEC, ANXA2, SPTBN1, TES, TAGLN2, PDLM1, SPTANI, AHNAK, FSCN1, KRT18, SEPT7, PAICS
Actividad estructural de moléculas	0.82	0.0031	S100A11, ANXA2, PDLM1, KRT18
Unión a RNA	0.68	2.03e-05	S100A11, ANXA2, AHNAK
Unión a proteínas del citoesqueleto	0.63	0.0412	LRPPRC, S100A11, CKB, PKM, PLEC, MDH2, ANXA2, PLS3, SPTBN1, ACTB, TES, EIF3B, TAGLN2, EBF3, PDLM1...
Unión a ácidos nucleicos	0.37	0.0154	HIST1H1D, IGF2BP3, LRPPRC, CKB, CYCS, PKM, PLEC, MDH2, IDH2, CS, ANXA2, SPTBN1, ACTB, TES, EIF3B, EBF3...
Unión a compuestos heterocíclicos	0.32	0.0054	HIST1H1D, IGF2BP3, LRPPRC, CKB, CYCS, PKM, PLEC, MDH2, IDH2, CS, ANXA2, SPTBN1, ACTB, TES, EIF3B, EBF3...
Unión a compuestos orgánicos cíclicos	0.31	0.0060	HIST1H1D, IGF2BP3, LRPPRC, CKB, CYCS, PKM, PLEC, MDH2, IDH2, CS, ANXA2, SPTBN1, ACTB, TES, EIF3B, EBF3...
Unión de proteínas	0.25	0.0404	HIST1H1D, IGF2BP3, LRPPRC, S100A11, CAZ, ANXA5, CKB, CYCS, COX5A, PKM, PLEC, MDH2, IDH2, CS, ANXA2...
Unión	0.16	0.0054	IGF2BP3, EIF3B, EIF4B, EIF4G1

**Tabla 3. Procesos biológicos modificados según la red modelizada de las proteínas diferencialmente expresadas.** Tabla de procesos biológicos incrementados (parte superior) y reducidos (parte inferior) en función de las proteínas diferencialmente expresadas a través del software STRING. *Strength* (log10) ratio entre número de proteínas en la red modelizada asociadas a un proceso biológico y el número de proteínas de una red modelizada aleatoria similar asociadas a ese proceso biológico. *False Discovery Rate* significancia del enriquecimiento. Los p-valores están corregidos para cada categoría usando el procedimiento Benjamini-Hochberg.

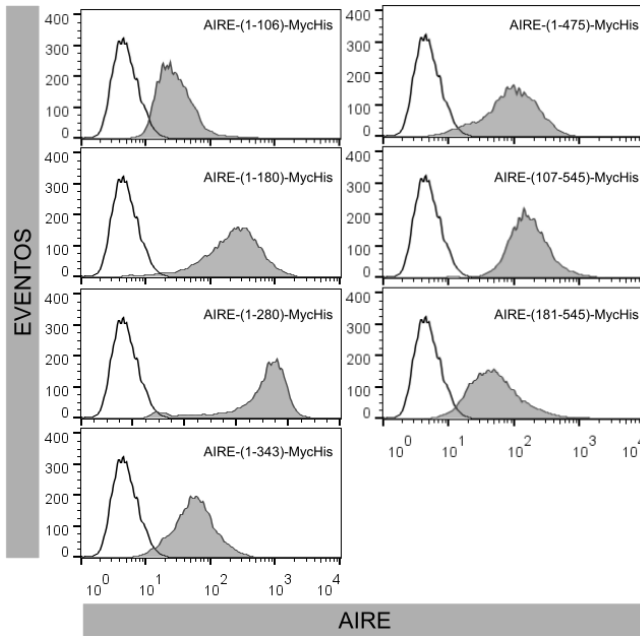


### 1.3 Identificación de los dominios de AIRE involucrados en apoptosis

Las elevadas evidencias de la relación de AIRE con la inducción de apoptosis inducida por daño genotóxico<sup>110 227 228</sup> y los resultados que se obtuvieron en la modelización de red de interacciones tras el análisis proteómico cuantitativo nos llevó a identificar la relación de AIRE con muerte celular programa en nuestros transfectantes estables. Para ello se comprobó los niveles de apoptosis espontánea tras un tratamiento con DMSO y los niveles de apoptosis inducida por daño genotóxico al aplicar un tratamiento con etopósido, conocido inhibidor de la TOP2, su aplicación favorece el bloqueo en la replicación del DNA al incrementar críticamente DSBs.

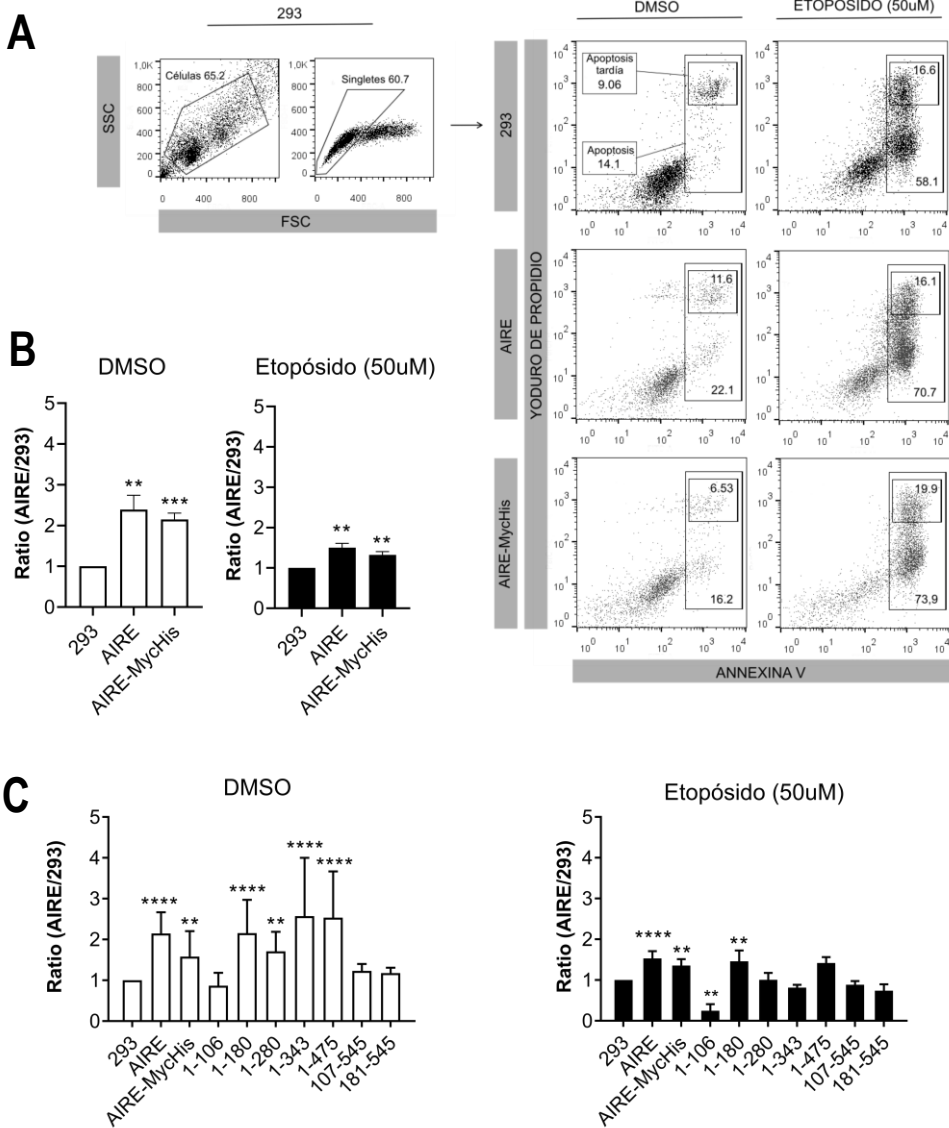
De nuevo, se clonó el gen de AIRE con los diversos dominios de la proteína truncados asociados al *tag* de Myc/His en su extremo C-terminal en el vector de expresión eucariota pcDNA3.1. Se generaron diferentes transfectantes estables en células HEK293 con las diferentes construcciones. Los transfectantes se clonaron por dilución límite. El resultado es un abanico de clones celulares de expresión estable de AIRE con una única población y de los distintos dominios de AIRE tanto en sentido N-terminal comenzando por AIRE (1-106) conteniendo el dominio CARD/HSR, AIRE (1-180) conteniendo el primer dominio y el motivo NLS, AIRE (1-280) conteniendo el primer dominio, motivo NLS y el dominio SAND, AIRE (1-343) conteniendo el dominio CARD/HSR, motivo NLS, dominio SAND y el dominio PHD1; AIRE (1-475) conteniendo el dominio CARD/HSR, motivo NLS, dominio SAND, el dominio PHD1, la región PRR y el dominio PHD2; como en sentido C-terminal: AIRE (107-545) conteniendo todos los dominios excepto el primer dominio HSR/CARD y AIRE (181-545) conteniendo todos los dominios excepto el primer dominio y el dominio SAND. No se han podido generar transfectantes estables de la proteína truncada de AIRE para las tres últimas construcciones: AIRE (281-545), AIRE (344-545) y AIRE (476-545) debido, probablemente, a la expresión de un producto proteico no estable que rápidamente puede ser degradado por el sistema UPS. Todos

ellos expresaron de manera estable y como única población las diferentes proteínas (Figura 4).



**Figura 4. AIRE WT y las diferentes proteínas truncadas se expresan de manera estable en transfectantes de HEK293.** Niveles de expresión de AIRE para cada clon transfectante estable. El histograma sin relleno indica la fluorescencia asociada a células HEK293 sin transfectar como control. (AIRE=FITC).

En primer lugar se realizaron experimentos de apoptosis con las construcciones de AIRE y AIRE-Myc/His. Mediante citometría de flujo se seleccionaron aquellas células positivas para anexina V (apoptosis temprana) y dobles positivas para anexina V y yoduro de propidio (apoptosis tardía) tras 24h de aplicación de los diferentes tratamientos. La **Figura 5A** muestra los histogramas obtenidos por citometría de flujo de un experimento representativo. Los transfectantes estables de AIRE<sup>+</sup> y AIRE-MycHis presentaron un incremento en la apoptosis espontánea e inducida comparados con las células control (**Figura 5B**).



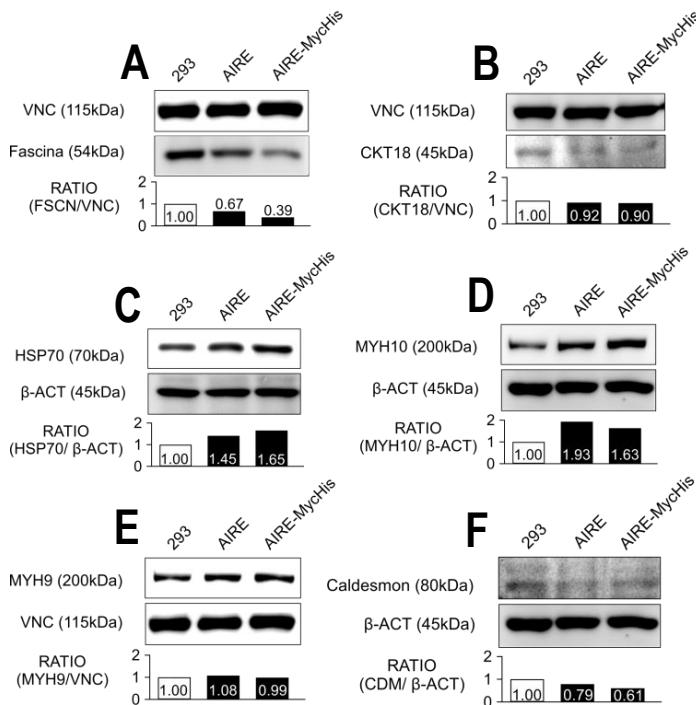
**Figura 5. La expresión de AIRE induce apoptosis.** Las células HEK293 y los diferentes transfectantes estables de AIRE fueron tratados con DMSO para medir apoptosis espontánea o con Etopósido (50uM) para medir apoptosis inducida. (A) Las células fueron teñidas con anexina V (FITC) and Yoduro de Propidio (PE) para analizar por citometría de flujo los niveles de apoptosis. (B) Ratio del incremento de apoptosis espontánea e inducida de transfectantes estables AIRE y AIRE-MychHis frente a HEK293 tras 24h tras el tratamiento con DMSO y etopósido. Valores de un experimento con n=4 réplicas. Análisis estadístico realizado con el test Wilcoxon, (\* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , \*\*\*\* $P < 0.0001$ ). (C) Ratio del incremento de apoptosis espontánea e inducida de diferentes transfectantes estables de AIRE frente a HEK293 tras 24h. Media de n=4 experimentos. Análisis estadístico realizado con el test Kruskal-Wallis, (\* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , \*\*\*\* $P < 0.0001$ ). Los valores representados de (B) y (C) muestran la apoptosis total (apoptosis temprana: células positivas para anexina V y apoptosis tardía: células positivas para anexina V y IP) normalizada frente a los valores de HEK293.

Posteriormente, se analizó el papel de los diferentes dominios en la inducción de apoptosis en las células HEK293. Para ello, se repitió el procedimiento anterior con los transfectantes estables de AIRE que expresaban proteínas truncadas en diferentes dominios de la proteína. Los datos mostraron un incremento de la apoptosis espontánea en aquellos transfectantes que expresaban, al menos, el dominio HSR/CARD y el motivo de importe nuclear NLS (transfectantes estables AIRE, AIRE-Myc/His, AIRE-(1-180), (1-280), (1-343), (1-475)-Myc/His) (**Figura 5B**). En relación a la apoptosis inducida por tratamiento con etopósido, se observa un incremento en los niveles de muerte celular en el caso de los transfectantes AIRE, AIRE-Myc/His y AIRE-(1-180)-Myc/His, la presencia del dominio HSR/CARD con el motivo NLS son suficientes para generar muerte celular (**Figura 5C**). El dominio HSR/CARD por sí solo no genera un incremento en la muerte celular programada puesto que el transfectante AIRE-(1-106)-Myc/His no muestra un incremento en apoptosis respecto a las células control. Asimismo, aunque la proteína pueda translocarse al núcleo por la presencia del motivo NLS, necesita el dominio HSR/CARD para inducir dicho efecto, puesto que el transfectante estable AIRE-(107-545)-Myc/His no presenta variaciones significativas en relación a niveles de apoptosis respecto a las células control (Figura 5B y C). El dominio HSR/CARD junto con el motivo NLS de AIRE son necesarios y suficientes para inducir apoptosis.

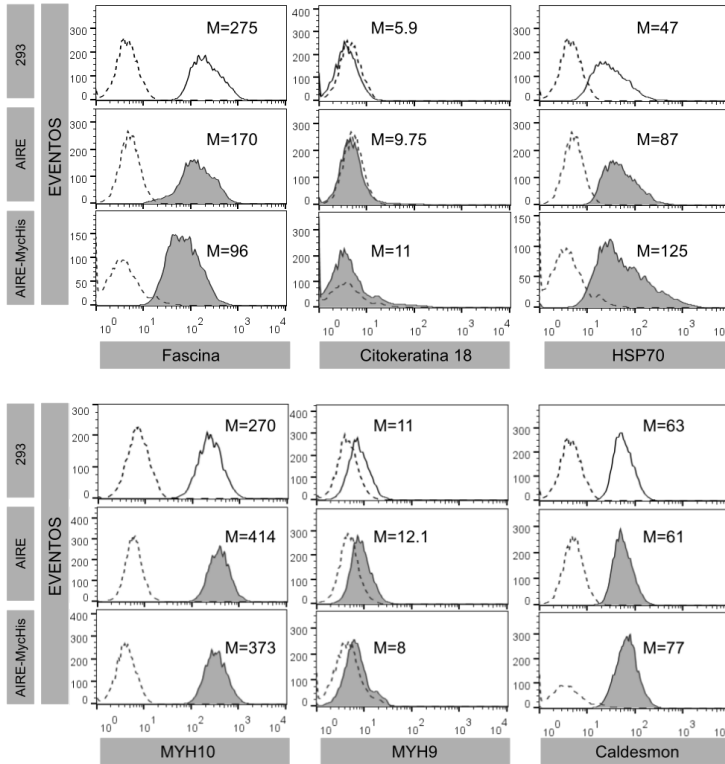
#### 1.4 La expresión de AIRE induce una mayor movilidad en células HEK293

Para confirmar los datos obtenidos en el análisis proteómico, se evaluó por *western blot* y citometría de flujo la abundancia de varias de las proteínas. Así, el análisis por *western blot* confirmó la disminución de las proteínas fascina y citokeratina 18 (**Figura 6A y B**). Asimismo, se vio incrementada la presencia de las proteínas HSP70 y la miosina 10 (miosina IIB, miosina II no muscular 10 o MYH10) (**Figura 6C y D**). Adicionalmente se analizaron dos proteínas de interés sin variación apreciable en el análisis proteómico: la miosina 9 (miosina IIA, miosina II no muscular 9 o MYH9) y caldesmón. Caldesmón se vio disminuida en ambos transfectantes de AIRE, en cambio, no se observó una

variación apreciable en MYH9 (Figura 6E y F). El análisis por citometría de flujo confirmó la variación en la proteína disminuida fascina, mientras que la citokeratina 18 no se detectó (Figura 7); el aumento de HSP70 y MYH10 se vio igualmente confirmado (Figura 7), en cambio, las proteínas MYH9 y caldesmón no se detectaron como modificadas en el mismo análisis (Figura 7). Por tanto, el análisis por *western blot* y citometría de flujo validaron los datos obtenidos por iTRAQ. Los datos de esta aproximación confirmaron los datos obtenidos previamente por nuestro grupo en la línea HT93 en los que se observó que la expresión de AIRE indujo la expresión de chaperonas, especialmente HSP70 y la disminución de varias proteínas relacionadas con el citoesqueleto<sup>110</sup>, incluida caldesmón, que ha dado resultados contradictorios entre el WB y la citometría de flujo. Asimismo, se observó un incremento en la miosina MYH10.

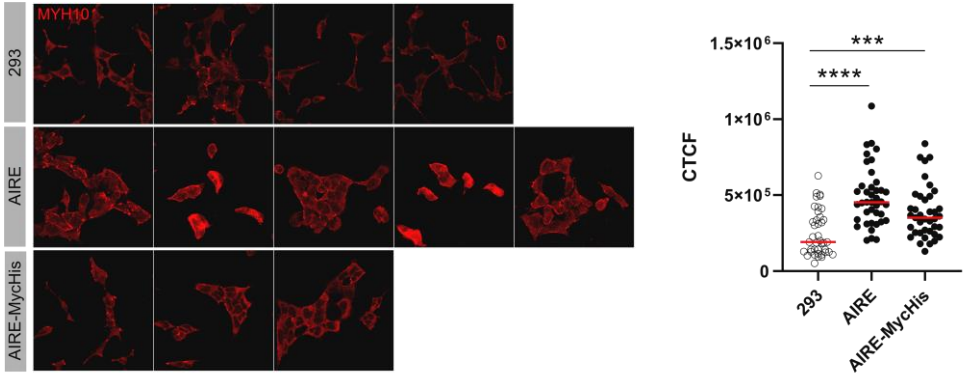


**Figura 6. Efecto de la expresión de AIRE de manera estable en diferentes transfectantes.** Análisis por *Western Blot* y Densitometría de lisados celulares de HEK293 y diferentes transfectantes estables de células HEK293-AIRE<sup>+</sup> y HEK293-AIRE-MychHis<sup>+</sup>. (A) WB y densitometría de fascina. (B) WB y densitometría de citokeratina 18. (C) WB y densitometría de HSP70. (D) WB y densitometría de MYH10. (E) WB y densitometría de miosina MYH9. (F) WB y densitometría de CALDESMON. RATIO indica la relación entre PROTEÍNA INDICADA:VINCULINA.

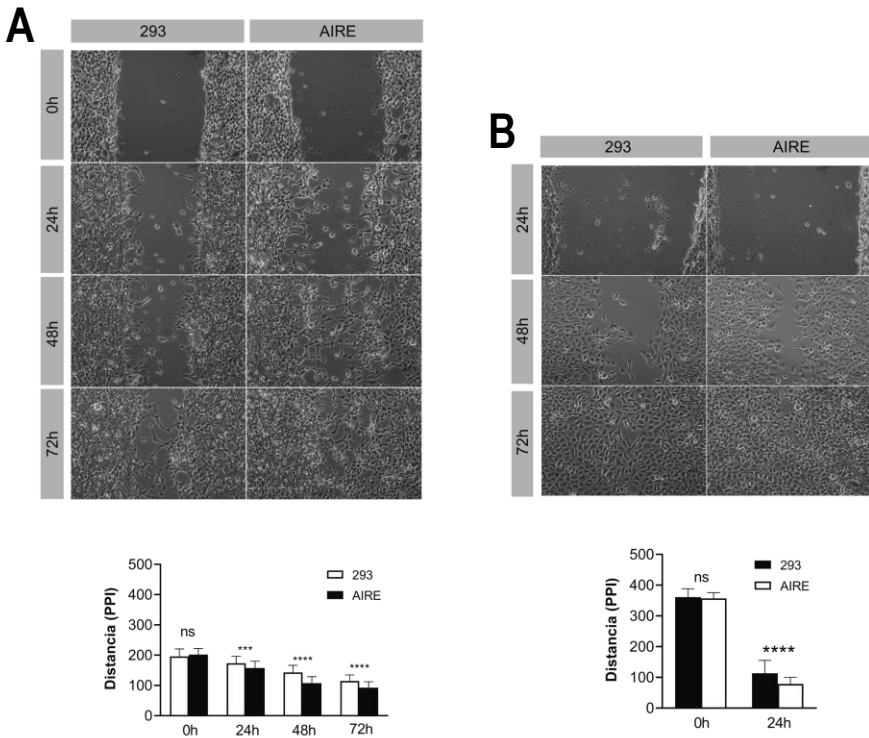


**Figura 7. Expresión de diferentes marcadores celulares y proteínas asociadas a los diferentes transfectantes estables de AIRE.** Análisis por citometría de flujo de la expresión de diferentes proteínas para transfectantes estables de células HEK293-AIRE<sup>+</sup> y HEK293-AIRE-MycHis<sup>+</sup>. (Nombre de la proteína=FITC), la línea discontinua indica la fluorescencia asociada a células HEK293 sin transfectar como control. M indica media de intensidad de fluorescencia asociada a cada marcador.

MYH10 ha sido definida como un elemento esencial para la compleción de la citocinesis y también para la adquisición de una correcta estructura celular, formación de adhesiones celulares y migración<sup>229 230</sup>. Por otro lado, la desregulación en la expresión de MYH10 ha sido atribuida a una mala prognosis en algunos tipos de cáncer, favoreciendo la invasividad y la migración<sup>231</sup>. Por tanto, adicionalmente a la comprobación del incremento en los niveles de MYH10 mediante *western blot* y citometría de flujo, se realizó una comprobación por microscopía confocal. MYH10 se vio incrementada en aquellas células que expresan tanto AIRE como AIRE-MycHis (Figura 8).



**Figura 8.** Las células HEK293-AIRE<sup>+</sup> ven incrementados los niveles de MYH10. (A) Inmunofluorescencia de células HEK293, HEK293-AIRE<sup>+</sup> y HEK293-AIRE-MycHis<sup>+</sup>. Las células fueron fijadas y teñidas con anticuerpos αMYH10 (Alexa568, rojo). Las imágenes fueron tomadas a 60X. (B) (CTCF) Fluorescencia Celular Total Corregida de n=50 células aleatorias. Análisis estadístico mediante Test de Mann-Whitney, (\**P*<0.05, \*\**P*<0.01, \*\*\**P*<0.001, \*\*\*\**P*<0.0001).



**Figura 9.** AIRE favorece la migración celular. (A) Panel izquierdo, parte superior: imágenes representativas de una zona de la herida sobre monocapa de células 293 y AIRE y su curación a lo largo de 72h con privación de suero. Panel izquierdo, parte inferior, media de distancias de 30 imágenes tomadas a lo largo de toda la herida. (B) Panel derecho, parte superior: imágenes representativas de una zona de la herida sobre monocapa de células 293 y AIRE y su curación a lo largo de 48h. Panel izquierdo, parte inferior, media de distancias de 30 imágenes tomadas a lo largo de toda la herida. Análisis estadístico realizado con t-test múltiple, (\**P*<0.05, \*\**P*<0.01, \*\*\**P*<0.001, \*\*\*\**P*<0.0001).

Se analizó la capacidad de migración de las células HEK293-AIRE<sup>+</sup> mediante experimentos de *wound healing*, tras realizar una herida con *cell scrapper* en cultivos celulares en monocapa bajo diferentes condiciones (medio suplementado con suero 10% FBS o sin suplemento) y a lo largo de 72h. El análisis de las imágenes mostró una recuperación de la herida más rápida en las células que expresaban AIRE en ambas condiciones experimentales (**Figura 9A y B**). Por tanto, nuestros datos sugieren que la expresión de AIRE favorece la capacidad de migración de las células.

### **1.5 Caracterización de los repertorios peptídicos asociados a las moléculas de HLA-I en células que expresan AIRE de forma diferencial**

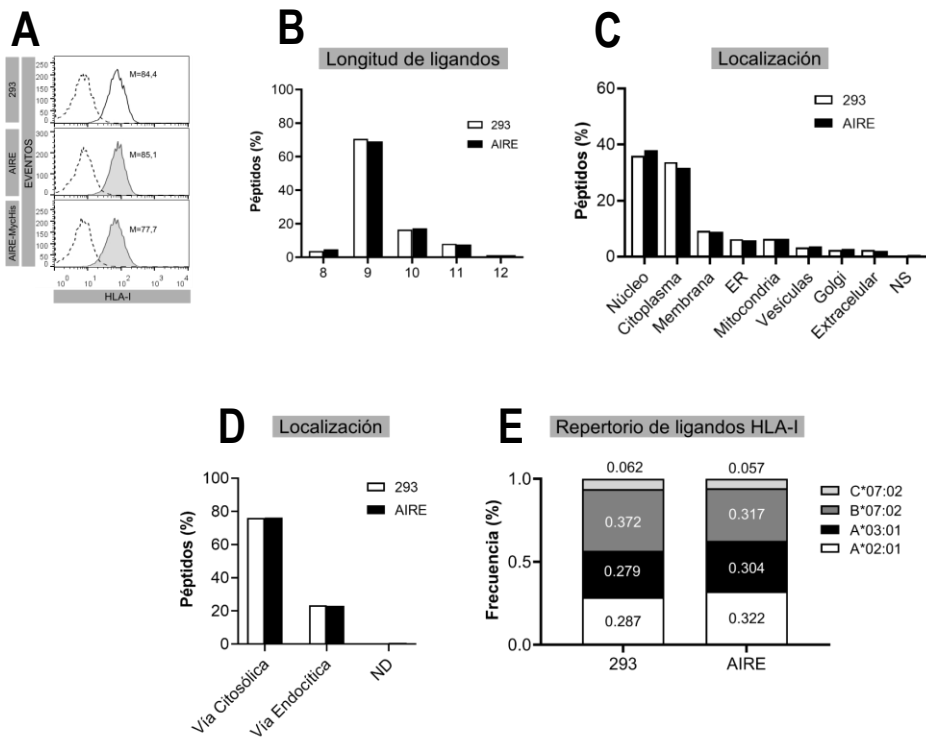
Se analizó por citometría la expresión de HLA-I en las células HEK293 y los transfectantes de AIRE y AIRE-MyHis. No se apreció una variación significativa en la expresión de las mismas (**Figura 10A**). Se purificaron por inmunoafinidad los péptidos unidos a complejos de HLA clase I. Para ello se utilizaron las líneas celulares HEK293 y HEK293 AIRE<sup>+</sup>. Dichas células expresan los alotipos de HLA de clase I: HLA-A\*02:01, HLA-A\*03:01, HLA-B\*07:02 y HLA-C\*07:02. Los inmunopeptidomas se analizaron por espectrometría de masas.

Se seleccionaron los péptidos entre 8 y 12 aminoácidos. Se identificaron un total de 1501 ligandos peptídicos de HLA-I únicos en las células de HEK293 y 1813 en las células HEK293-AIRE<sup>+</sup>. Estos péptidos derivaron de 1121 proteínas en HEK293 y 1275 en HEK293-AIRE<sup>+</sup>. La distribución por tamaño de los péptidos fue la estándar de los ligandos de clase I, con una mayoría de péptidos de 9 aminoácidos, seguida de péptidos de 10, 11, 8 y 12 aminoácidos (**Figura 10B**).

Posteriormente se determinó el origen de los péptidos unidos a las moléculas de HLA-I de ambas muestras. La localización de las proteínas parentales se definió utilizando las bases de datos: UniProt, Human Protein Reference Database y The Human



Protein Atlas. Cuando una proteína se encontraba localizada en diferentes compartimentos celulares, se eligió el más común asociado a su función biológica. Las proteínas parentales presentaron una localización principalmente nuclear y citoplasmática en ambas muestras (**Figura 10C**). Así, en torno a un 76% de los péptidos provenían de proteínas de la ruta endógena y aproximadamente un 23% provinieron de la vía endocítica, por lo que no se observaron diferencias en la vía de generación de los péptidos secuenciados (**Figura 10D**). Estos resultados son consistentes con los esperados para los inmunopeptidomas de HLA de clase I.

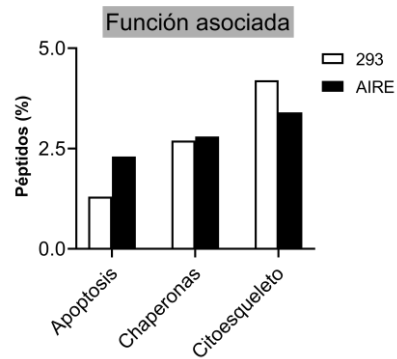


**Figura 10. Caracterización de los repertorios peptídicos asociados a las moléculas de HLA-I en células que expresan AIRE.** (A) Análisis por citometría de flujo de la expresión de HLA-I en transfectantes estables de AIRE y AIRE-MycHis, las células fueron teñidas con  $\alpha$ MHC-I (APC). La línea discontinua hace referencia al mismo tipo celular teñido con el control de isotipo correspondiente. (B) Distribución de tamaño por aminoácido del porcentaje de péptidos secuenciados comparados. (C) Porcentaje de péptidos secuenciados según la localización de su proteína parental. (D) Porcentaje de péptidos secuenciados según la vía de procesamiento antigénico. Vía citosólica comprende: citosol, núcleo y mitocondria. Vía endocítica comprende: membrana, compartimentos vesiculares, retículo endoplásmico, aparato de Golgi y proteínas extracelulares o secretadas. ER: Retículo Endoplásmico. ND: No determinado. (D) Porcentaje de péptidos secuenciados según la vía de procesamiento antigénico. Vía citosólica comprende: citosol, núcleo y mitocondria. Vía endocítica comprende: membrana, compartimentos vesiculares, retículo endoplásmico, aparato de Golgi y proteínas extracelulares o secretadas. ER: Retículo Endoplásmico. ND: No determinado. (E) Frecuencia de ligandos asociados a cada alelo de HLA-I.

Se analizó si la expresión de AIRE modificaba de alguna manera el procesamiento y la presentación antigénica por parte de los distintos alotipos expresados en la célula. Se analizó la afinidad de los péptidos secuenciados a las distintas moléculas de HLA-I expresadas y se asignó cada péptido a un alotipo específico. Tal y como se muestra en la **Figura 10E**, no se modificó de forma apreciable el porcentaje de ligandos presentados por cada alotipo.

Respecto a la función biológica de las proteínas parentales identificadas, utilizando las mismas bases de datos, se identificó un incremento destacable en el número de proteínas asociadas a procesos apoptóticos y de muerte celular programa en las muestras de HEK293-AIRE<sup>+</sup>. Comparando las mismas muestras se detectó un

ligero incremento en la identificación de proteínas de tipo chaperona y un descenso en el número de proteínas parentales asociadas a citoesqueleto (**Figura 11**). En general los resultados obtenidos refuerzan y confirman los datos obtenidos por iTRAQ y anteriormente en la línea celular HT93<sup>110</sup>. En las HEK293-AIRE se detectaron péptidos pertenecientes a proteínas que interaccionan con AIRE (Topoisomerasa 2A y 2B, RNA-pol II o HAUS6)<sup>63 83 52</sup> que no aparecieron en el análisis de las muestras de HEK293. Los diferentes péptidos identificados diferencialmente generados en células HEK293-AIRE<sup>+</sup> están recogidos en la **Tabla 4**.



**Figura 11. Distribución de la función asociada a las proteínas parentales de los ligandos secuenciados de HLA-I.** Comparación de porcentajes de péptidos provenientes de proteínas parentales con una función asociada: muerte celular o apoptosis, actividad de tipo chaperona o asociadas a citoesqueleto.

En conjunto, los datos obtenidos del análisis proteómico por iTRAQ así como la caracterización del repertorio peptídico asociado a HLA-I en células que expresan AIRE, sugieren que AIRE modificaría la expresión de proteínas relacionadas con el

citoesqueleto y proteínas de tipo chaperona, asimismo, un aumento en la susceptibilidad a apoptosis en células que expresan esta proteína y un aumento en la movilidad celular. La actividad de AIRE afecta, por tanto, a la composición del proteoma celular y a la fisiología celular.

**Tabla 4. Lista de ligandos diferencialmente expresados en las muestras de HEK293-AIRE<sup>+</sup>**

Función asociada	Secuencia peptídica	Identificador	Nombre de la proteína	Localización	Sublocalización
Apoptosis	APSPARLAP	Q43918 AIRE_HUMAN	Regulador Autoinmune	Núcleo	Citoplasma
	SUNVGLVSV	Q92688 AN32B_HUMAN	APRIL	Citoplasma	Núcleo
	KLLDAYLL	P61803 DAD1_HUMAN	Defensor contra muerte celular 1	Reticulo Endoplásmico	
	FQDPVPLTV	O15164 TIF1A_HUMAN	Factor 1-alfa intermediario de transcripción	Núcleo	
	FOGKPLRKK	Q9ULW0 TPX2_HUMAN	Factor de nucleación de microtúbulo TPX2	Núcleo	Citoesqueleto
	FLWKEPATM	Q9BQ70 TCF25_HUMAN	Factor de transcripción 25	Núcleo	
	AAPGAPAAV	Q8N726 ARF_HUMAN	Kinasa inhibidora 2A dependiente de ciclina	Núcleo	Mitocondria
	LTVEFKVQK	Q9Y5V3 MAGD1_HUMAN	Miembro D1 de familia MAGE	Citoplasma	Membrana
	ALSDGVHKI	Q9NVQ4 FAIM1_HUMAN	Molécula 1 inhibidora de apoptosis via FAS	Citoplasma	
	HIENINAV	Q76021 RL1D1_HUMAN	Proteína 1 contenedora de dominio L1 ribosomal	Núcleo	
	ATFTNVFGK	Q8IWZ3 ANKH1_HUMAN	Proteína 1 contenedora de repeticiones de ankirina y dominio KH	Citoplasma	
	ATAVAQLYK	Q14CZ0 CP072_HUMAN	Proteína 1 de respuesta a desgaste de telómeros y respuesta a p53	Núcleo	
	AVFDGQVTSK	O14683 P511_HUMAN	Proteína 11 inducible por p53	Reticulo Endoplásmico	
	ALDKATVLL	Q53E16 PDCD4_HUMAN	Proteína 4 de muerte celular programada	Núcleo	
	AAAYPHLTL	Q9P2R6 RERE_HUMAN	Proteína relacionada con atrofia 1	Núcleo	
	GPRLPWTRL	Q6NU04 TM214_HUMAN	Proteína Transmembrana 214	Reticulo Endoplásmico	
	ATWYSPLLK	O15350 P73_HUMAN	Proteína tumoral p73	Citoplasma	Núcleo
	APRALPGSAVL	Q8IUW5 RELL1_HUMAN	Receptor 1 similar a receptor expresado en tejido linfóide	Membrana	
	RVFTGVVTK	Q8IX12 CCAR1_HUMAN	Regulador 1 de ciclo celular y apoptosis	Citoplasma	
	SPAPLPHV	Q8N163 CCAR2_HUMAN	Regulador 2 de ciclo celular y apoptosis	Núcleo	Citoesqueleto
	SLGSGVATV	Q95197 RTN3_HUMAN	Reticulón 3	Reticulo Endoplásmico	
	SPARLSGL	Q14296 FASTK_HUMAN	Serina/Treonina Kinasa activada por FAS	Mitocondria	
	KLIDFGSGALLK	Q8QV86 PIM3_HUMAN	Serina/Treonina Kinasa PIM-3	Citoplasma	
SPVSNVSL	Q5TH74 STPG1_HUMAN	STPG1	Citoplasma	Núcleo	
VPRGQGSQF	Q75807 PR15A_HUMAN	Subunidad reguladora 15A de la proteína fosfatasa 1	Reticulo Endoplásmico	Mitocondria	
AIFGNADPAQK	Q6NU56 TECT3_HUMAN	Tectonic 3	Membrana		
Chaperona	SLDDPVEV	Q972616 GCN1_HUMAN	Activador GCN1 de EIF2AK4	Núcleo	
	GTLISGWLISK	P27824 CALX_HUMAN	Calnexina	Reticulo Endoplásmico	
	MIIAIPTVGK	P27797 CALR_HUMAN	Calreticulina	Reticulo Endoplásmico	
	ILGGPGTVQGV	O14618 CCS_HUMAN	Chaperona de cobre para la superóxido dismutasa	Citoplasma	
	KYGEVYTK	P10809 CH60_HUMAN	Miembro 1 de la familia de proteína de choque térmico D (Hsp60)	Mitocondria	
	GGPGGGGF	P0DMV9 HIS7B_HUMAN	Miembro 1B de la familia de proteína de choque térmico A (Hsp70)	Citoplasma	
	RILMNETTAV	P34932 HSP74_HUMAN	Miembro 4 de la familia de proteína de choque térmico A (Hsp70)	Citoplasma	
	KLDDTYKA	Q9P615 DNK7_HUMAN	Miembro 7 de la familia de proteínas de choque térmico DnaI (Hsp40)	Citoplasma	Golgi, Extracelular
	APSELLLRKF	Q9UKB3 DIC12_HUMAN	Miembro C12 de la familia de proteínas de choque térmico DnaI (Hsp40)	Citoplasma	Citoesqueleto, Núcleo
	LLLPGPSAA	P23284 PPIB_HUMAN	Peptidilprolil Isomerasa B	Reticulo Endoplásmico	
	FPAGPPSHSL	Q96AY3 FKBP10_HUMAN	Prolil isomerasa 10 FKBP	Reticulo Endoplásmico	
	ARADFGKVL	Q02790 FKBP4_HUMAN	Prolil isomerasa 4 FKBP	Citoplasma	
	IPAKPPVSF	Q6PCK3 TCP1_HUMAN	Proteína 11 contenedora de dominio de tioredoxina	Reticulo Endoplásmico	Membrana
	SAFGQAFSK	Q6NKR4 TTI2_HUMAN	Proteína 2 de interacción a TEO2	Núcleo	Citoesqueleto
	GLDSVSVSEV	Q9Y4R8 TELO2_HUMAN	Proteína 2 de mantenimiento de telómeros	Núcleo	
	YGPPEYGV	O00264 PGR1_HUMAN	Proteína de unión a progesterona	Reticulo Endoplásmico	Mitocondria
	FLGPWPAAS	P30533 AMRP_HUMAN	RAP	Reticulo Endoplásmico	
	ALLDGSNVVFK	O15212 PFD6_HUMAN	Subunidad 6 de Prefoldina	Citoplasma	
	IINSSITTK	P49368 TCPG_HUMAN	Subunidad 3 de la chaperonina contenedora TCP1	Citoplasma	
	HIETILGK	P48643 TCPE_HUMAN	Subunidad 5 de la chaperonina contenedora TCP1	Citoplasma	Citoesqueleto
	KTAEDAKDFK	Q99471 PFD5_HUMAN	Subunidad 5 de Prefoldina	Núcleo	
	RPAFPVKLL	Q13111 CAF1A_HUMAN	Subunidad A del factor 1 de ensamblaje de cromatina	Núcleo	
	QSFKEWVAK	P17987 TCPA_HUMAN	Subunidad alfa del Complejo T 1	Citoplasma	Citoesqueleto
TVIRDPSPI	Q13112 CAF1B_HUMAN	Subunidad B del factor 1 de ensamblaje de cromatina	Núcleo		
Interacción con AIRE	VLYNGFTGR	P30876 RPB2_HUMAN	RNA Polimerasa II	Núcleo	
	SAFGGSLPAK	Q7Z4H7 HAU56_HUMAN	Subunidad 6, complejo similar a augmina HAUS	Citoplasma	Citoesqueleto
	KLDETNLSL	P11388 TOP2A_HUMAN	Topoisomerasa 2A	Núcleo	
	AVSGEIPVW	Q02880 TOP2B_HUMAN	Topoisomerasa 2B	Núcleo	





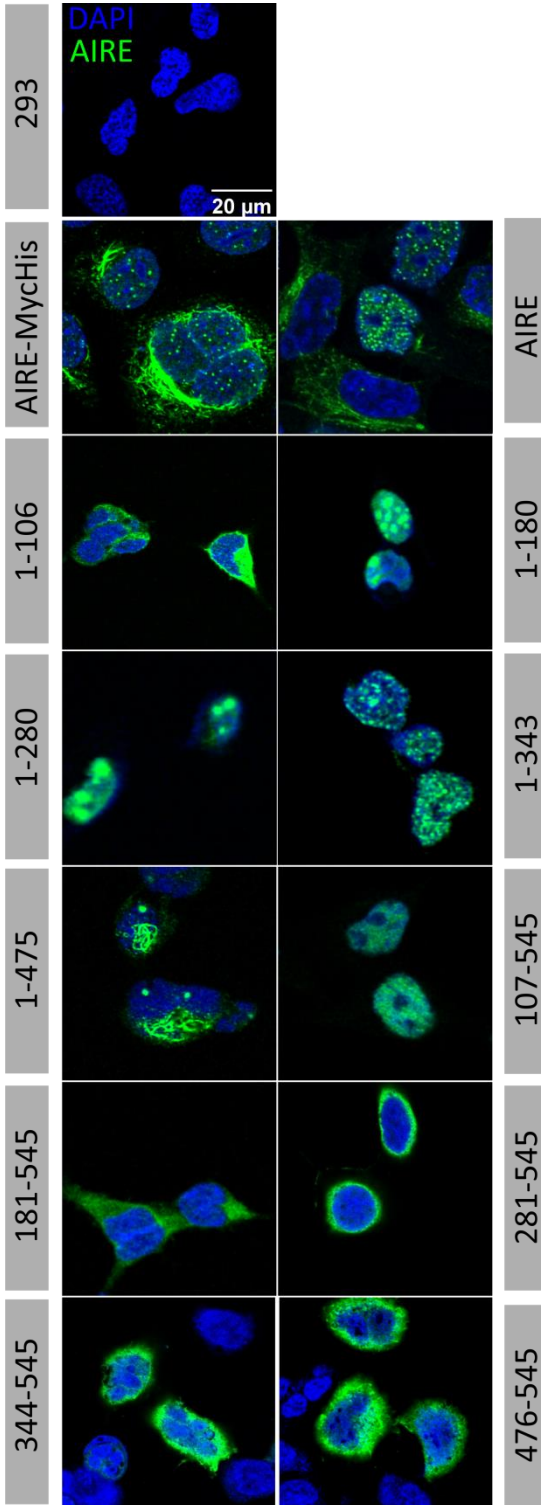


## CAPÍTULO II: AIRE y ciclo celular. Caracterización de una nueva localización de AIRE en células post-mitóticas

### 2.1 Localización celular de AIRE y de la proteína truncada conteniendo diversos dominios

Para analizar la influencia de los diferentes motivos de AIRE en la localización celular de la proteína se usaron los transfectantes estables generados y descritos previamente. Éstos son: AIRE, AIRE-Myc/His, AIRE-(1-106)-Myc/His, AIRE-(1-180)-Myc/His, AIRE-(1-280)-Myc/His, AIRE-(1-343)-Myc/His, AIRE-(1-475)-Myc/His, AIRE-(107-475)-Myc/His y AIRE-(181-545)-Myc/His (**Figuras 1 y 4**). Adicionalmente se usaron transfecciones transitorias con las construcciones AIRE-(281-545)-Myc/His, AIRE-(344-545)-Myc/His y AIRE-(476-545)-Myc/His, para los cuales no se pudo obtener transfectantes estables, a pesar de obtener múltiples transfecciones transitorias que no se lograron estabilizar con el tiempo. Se realizaron análisis de microscopía confocal sobre cada uno de los clones de transfectantes estables y de los transfectantes transitorios. Se confirmó que todos los clones de transfectantes estables generados expresaban la proteína AIRE completa o asociada al *tag* Myc/His, o las diferentes proteínas AIRE-Myc/His truncadas en diferentes dominios (1-106 a 181-545) correctamente (**Figura 12**).

Cada clon, dependiendo de los dominios que porte la proteína AIRE presenta una localización, morfología y distribución específicas. Así, las proteínas AIRE y AIRE-Myc/His tienen una localización celular similar, principalmente en el núcleo, distribuido de manera punteada en motas o *dots* nucleares, esta distribución es la habitual de la proteína y es la que predomina en las inmunofluorescencias realizadas sobre mTECs<sup>79</sup> (**Figura 12**). Asimismo, se aprecia que AIRE está localizado en fibras nucleares en el transfectante de AIRE-Myc/His y AIRE (**Figura 12**). La eliminación del motivo de importe nuclear de AIRE, situado entre los residuos 131-133, dirige la proteína completamente al citoplasma, distribuyéndose homogéneamente por todo el citosol de la célula. Así,



**Figura 12. Morfología y localización de AIRE en transfectantes estables y transitorios.** Inmunofluorescencia de células HEK293, HEK293-AIRE<sup>+</sup> (WT, AIRE-MycHis y AIRE-MycHis truncado en diferentes dominios). Las imágenes de AIRE-(281-545)-MycHis, AIRE-(344-545)-MycHis y AIRE-(476-545)-MycHis corresponden a transfectantes transitorios. Las células fueron fijadas y teñidas con anticuerpos  $\alpha$ Myc (Alexa488, verde) y DAPI para la visualización de los núcleos celulares. DAPI (azul), MERGE (combinación de canales azul y verde), *Brightfield* BF (negro). Las imágenes fueron tomadas a 60X con un aumento de de zoom 2X.



las proteínas: AIRE-(1-106)-Myc/His, AIRE-(181-545)-Myc/His, AIRE-(281-545)-Myc/His, AIRE-(344-545)-Myc/His y AIRE-(476-545)-Myc/His de los diferentes transfectantes la localizan únicamente fuera del núcleo (**Figura 12**). Todos los transfectantes que contienen el motivo de localización nuclear (AIRE-(1-180)-Myc/His, AIRE-(1-280)-Myc/His, AIRE-(1-343)-Myc/His, AIRE-(1-454)-Myc/His, AIRE-(107-475)-Myc/His) expresan la proteína en el núcleo, aunque los patrones varían. Así, la eliminación del dominio rico en prolinas (PRR) situado entre los residuos 350-430 y el dominio PHD2 (residuos 343-475) de la proteína AIRE no parece afectar a la localización y distribución habitual, como se muestran en las inmunofluorescencias del transfectante estable AIRE-(1-343)-Myc/His (**Figura 12**). Sin embargo, la eliminación del extremo C-terminal de la proteína a partir del dominio SAND (residuos 189-280; transfectante AIRE-(1-180)-Myc/His) como a partir del dominio PHD1 (residuos 299-340 (transfectante AIRE-(1-280)-Myc/His), provocan un engrosamiento de los *dots* nucleares pero su reducción en número, ambas proteínas truncadas se distribuyen principalmente en el núcleo. La eliminación del primer dominio de la proteína, HSR/CARD, (AIRE-(107-545)-Myc/His) dirige la proteína hacia el núcleo con una distribución difusa a lo largo de él, no se observa el punteado característico de AIRE-WT y además parecen evitarse los nucléolos o zonas ricas de heterocromatina. Por último, la eliminación de la última región rica en prolinas situada entre los residuos 518-524 (transfectante AIRE-(1-475)-Myc/His) de la proteína AIRE favorece la distribución y localización asociada a fibras o filamentos (**Figura 12**).

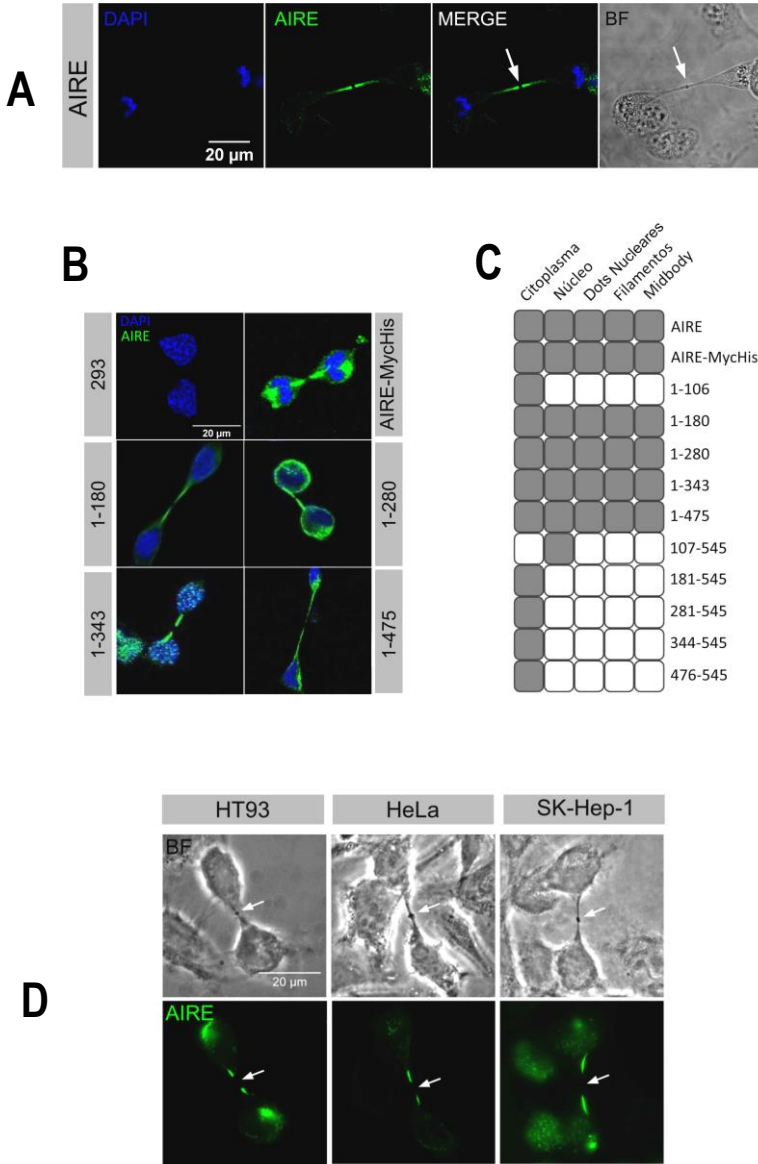
En conjunto, nuestros datos sugieren que el motivo principal para la oligomerización u homomultimerización de AIRE (dominio HSR/CARD)<sup>232</sup> y el motivo de importe nuclear (motivo NLS) son esenciales para la adquisición de una localización de AIRE similar al observado en las mTECs. Asimismo, las características bioquímicas de cada dominio de AIRE afectan a la sublocalización celular (**Figura 13C**).

## 2.2 AIRE se localiza en células postmitóticas cerca del *midbody*

La localización subcelular de AIRE dentro de la célula es dependiente del ciclo celular y la formación de los característicos punteados nucleares sería posterior a la división celular, desapareciendo poco antes de la entrada en mitosis y durante el proceso de división<sup>103</sup>. Este hecho sugeriría la imposibilidad de una implicación de AIRE en división celular. No obstante, estudios en células embrionarias de ratón identificaron a Aire como un regulador de proteínas del huso mitótico. Durante la caracterización de la localización de la proteína en diferentes transfectantes estables mediante microscopía confocal se identificó la presencia de AIRE en estructuras fibrilares en torno al *midbody* en células post-mitóticas, durante la fase final de división celular, la denominada citocinesis (**Figura 13A**).

Esta localización, identificada en células embrionarias de ratón<sup>52</sup>, también es adquirida por AIRE en células humanas. Observando la localización y distribución de la proteína completa o la proteína truncada de AIRE en los diferentes transfectantes estables generados en HEK293 y de las de las tres proteínas truncadas de las que no se obtuvieron transfecciones estables transfectadas transitoriamente, se pudo determinar que la presencia de los primeros 180 aminoácidos son necesarios para la localización de la proteína AIRE en o filamentos cerca del *midbody* (**Figura 13B** y **Figura 13C**). Como es de esperar, las proteínas truncadas de AIRE que no contienen los primeros 18 aminoácidos no logran generar las estructuras filamentosas en torno al *midbody* en células post mitóticas. Este hecho sugiere que esta nueva localización requiere de la correcta oligomerización de AIRE y que la región necesaria para la interacción con proteínas que intervienen en la citocinesis se localiza entre los residuos 1-180 de AIRE.

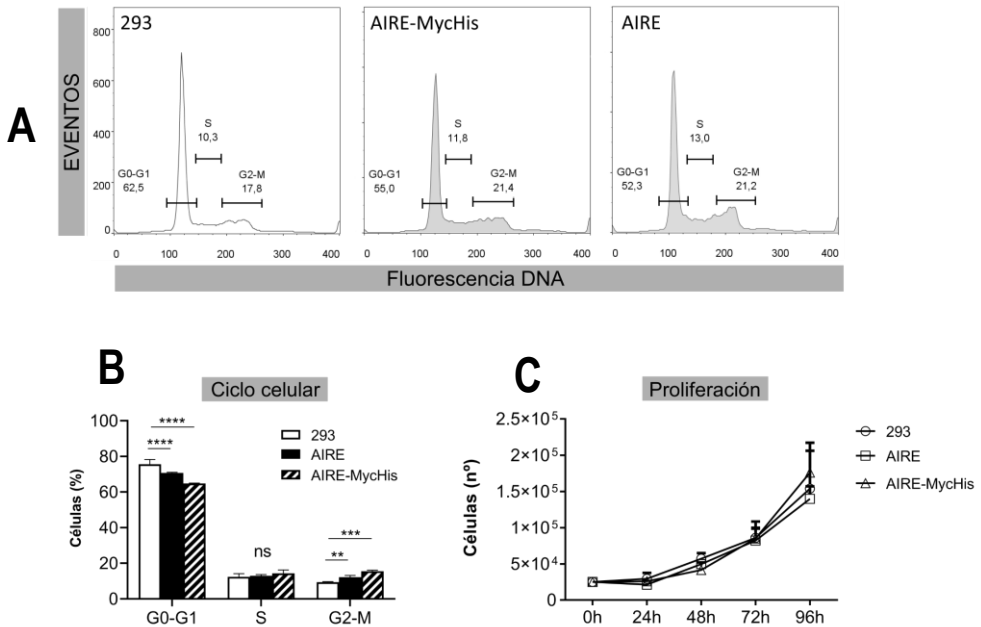
Para confirmar que esta localización no se debe a que las células HEK293 son células de origen embrionario, se transfectó AIRE en células HeLa y SK-Hep-1. Se analizó por microscopía de inmunofluorescencia si AIRE se encontraba en esta nueva localización celular en transfectantes de las líneas HT93, HeLa y SK-Hep-1, confirmando que esta localización se da en todos los transfectantes (**Figura 13D**).



**Figura 13. AIRE se localiza en torno al midbody en células post-mitóticas.** (A) y (B) Inmunofluorescencia de células HEK293, HEK293-AIRE<sup>+</sup> (WT, AIRE-MycHis y AIRE-MycHis truncado en diferentes dominios). Las células fueron fijadas y teñidas con anticuerpos  $\alpha$ Myc (Alexa488, verde) y DAPI para la visualización de los núcleos celulares. DAPI (azul), MERGE (combinación de canales azul y verde), *Brightfield* BF (negro). La flecha indica la visualización del *midbody* en campo claro. Las imágenes fueron tomadas a 60X con un aumento de zoom 2X. (C) Representación esquemática de la morfología y localización de la proteína AIRE en los diferentes transfectantes estables y transfectantes transitorios (AIRE-(281-545)-MycHis, AIRE-(344-545)-MycHis y AIRE-(476-545)-MycHis). (D) Inmunofluorescencia de células HT93, HeLa y SK-Hep-1 transfectadas transitoriamente con AIRE WT. Las células fueron fijadas y teñidas con anticuerpos  $\alpha$ AIRE (Alexa488, verde), *Brightfield* BF (negro). La flecha indica la visualización del *midbody* en campo claro. Las imágenes fueron tomadas a 60X con un aumento de de zoom 2X.

La expresión de AIRE es esencial para la correcta diferenciación y maduración de las mTECs y, por tanto, la adquisición de una arquitectura tímica óptima para llevar a cabo el proceso de selección tímica<sup>56 233 234</sup>. Sin embargo no se ha comprobado la importancia de AIRE respecto al ciclo celular en mTECs.

Se comprobó si la expresión de AIRE afectaba al ciclo celular. Se analizó la distribución poblacional de células HEK293, HEK293-AIRE<sup>+</sup> y HEK293-AIRE-MyHis<sup>+</sup> atendiendo a la fase del ciclo celular mediante citometría de flujo y tinción con yoduro de propidio (agente intercalante de doble hebra de DNA que se une de manera estequiométrica) (Figura 14A). Las células que expresan AIRE presentaban mayor población en fase G2-M



**Figura 14. La expresión de AIRE afecta al ciclo celular.** (A) Análisis de la distribución poblacional según la fase del ciclo celular de las células HEK293 y HEK293-AIRE<sup>+</sup> no sincronizadas mediante citometría de flujo, la distribución se realizó en base a la media de fluorescencia emitida por el DNA. Las células fueron fijadas previamente al marcaje con yoduro de propidio (rojo, FL3-A). (B) Media de la distribución poblacional según la fase del ciclo celular de las mismas células no sincronizadas analizadas en diferentes días (n=5). Análisis estadístico realizado con el método 2wayANOVA (\* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , \*\*\*\* $P < 0.0001$ ). (C) Ensayo de proliferación por contaje celular, se partió de un mismo número de células 25000/w (t=0h), cada 24h se realizó un contaje de hasta cuatro pocillos independientes por cada condición hasta las 96h.

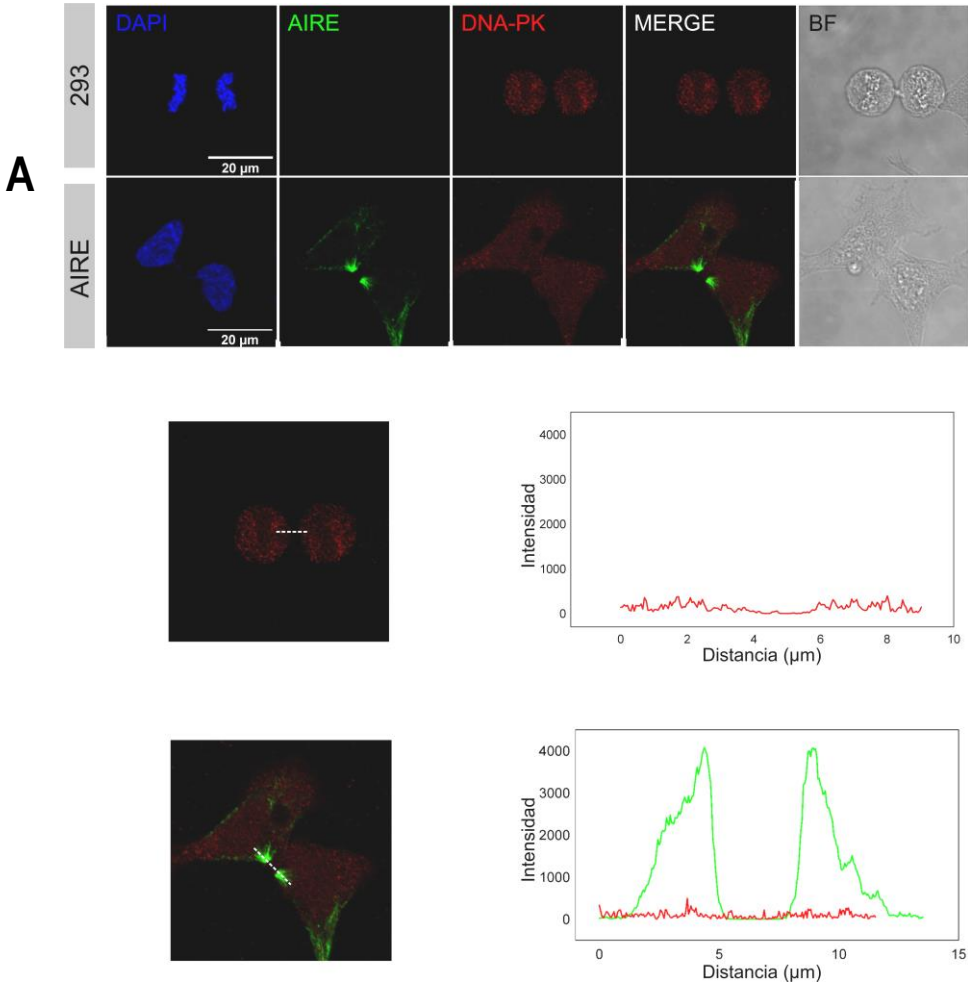
(fase previa al proceso de mitosis o división celular en la que las células han duplicado su material genético y sintetizan orgánulos y componentes necesarios para el reparto equitativo a las células hijas), por contrapartida ven reducido el número de células en fase G0-G1 (**Figura 14B**). Esto podría indicar una menor capacidad de proliferación respecto a las células HEK293, por tanto, se comparó la capacidad de proliferación de ambas mediante un conteo celular durante 96h en cultivo. No se observó ninguna disminución en la proliferación en aquellas células que expresan AIRE (**Figura 14C**). Por tanto, las células que expresan AIRE parecen presentar una desregulación del ciclo celular frente a las células que no lo expresan, sufren un incremento del número de células en fase G2-M, pero no se traduce en una variación de la capacidad proliferativa.

#### **2.4 Análisis de la distribución de AIRE, DNA-PK y Miosina 10 en el *midbody***

La citocinesis resulta el paso final y esencial para la escisión de la membrana celular tras el proceso de división celular, la generación del *midbody*, estructura transitoria donde se produce el corte por constricción de la membrana celular de las células hijas unidas, permite la separación física de ambas y resulta una zona activa con importante recambio y redistribución proteica<sup>235 96</sup>. Por ello, se procedió a analizar la localización de diversas proteínas en conjunción con AIRE y su distribución a lo largo del *midbody*.

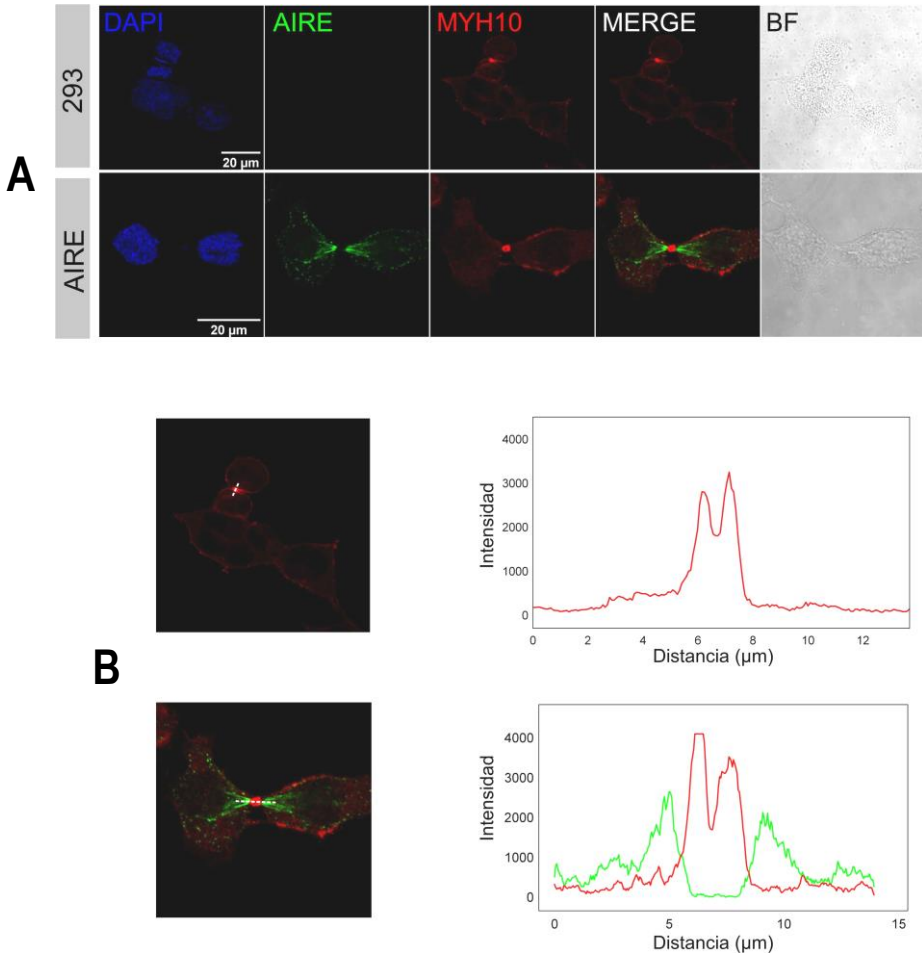
DNA-PK, además de su papel en el núcleo celular en la reparación del DNA, se encarga de regular enzimas claves en el proceso de citocinesis<sup>101</sup>, su interacción con AIRE dentro del núcleo nos hace suponer que es posible que la localización de AIRE en torno al *midbody* pueda deberse a esta interacción, por ello, se comprobó la localización de DNA-PK y AIRE en células post-mitóticas. En ningún caso se detectó la presencia de DNA-PK en torno al *midbody* (**Figura 15**). Por tanto, los datos sugieren que DNA-PK y AIRE no interaccionan durante este proceso, aunque la falta de detección de DNA-PK no permite confirmar este extremo. Posteriormente se analizó la distribución de MYH10 y AIRE. MYH10 se vio incrementada en células AIRE<sup>+</sup> (**Figuras 6D, 7 y 8**) y debido

a su implicación en la finalización de escisión durante la formación y desarrollo del surco de corte en células post-mitóticas, se comprobó la localización de MYH10 en el *midbody*. MYH10 se distribuye irregularmente a lo largo de la célula, en fase G1-G0 se



**Figura 15. Localización de AIRE y DNA-PK en procesos de citocinesis.** Inmunofluorescencias de células HEK293 y HEK293-AIRE<sup>+</sup>. Las células fueron fijadas y teñidas con anticuerpos  $\alpha$ AIRE (Alexa488, verde) y  $\alpha$ DNA-PK (Alexa568, rojo) y DAPI (azul) para la visualización de los núcleos celulares, MERGE (combinación de los canales verde y rojo) y *Brightfield* (BF, negro). (A) Inmunofluorescencias de DNA-PK y AIRE y (B) Perfil histográfico de la intensidad de fluorescencia media de una sección longitudinal correspondiente. Imágenes analizadas de (n=1) experimentos. Las imágenes fueron tomadas a 60X con diversos aumentos. Las líneas de puntos indican la sección longitudinal de corte.

en la zona donde se generará el surco de corte, paralelo a la placa ecuatorial tras el reparto cromosómico al finalizar la metafase y comenzar la telofase. MYH9 y MYH10 co-localizan en ambas situaciones<sup>229 100</sup>. Se procedió a analizar la distribución de MYH10 y AIRE conjuntamente.



**Figura 16. Localización de AIRE y MYH10 en procesos de citocinesis.** Inmunofluorescencias de células HEK293 y HEK293-AIRE<sup>+</sup>. Las células fueron fijadas y teñidas con anticuerpos  $\alpha$ AIRE (Alexa488, verde) y  $\alpha$ MYH10 (Alexa568, rojo) y DAPI (azul) para la visualización de los núcleos celulares, MERGE (combinación de los canales verde y rojo) y *Brightfield* (BF, negro). (A) Inmunofluorescencias de MYH10 y AIRE y (B) Perfil histográfico de la intensidad de fluorescencia media de una sección longitudinal correspondiente. Imágenes analizadas de (n=3) experimentos. Las imágenes fueron tomadas a 60X con diversos aumentos. Las líneas de puntos indican la sección longitudinal de corte.

En células post-mitóticas AIRE se localiza en la zona externa al *midbody* dejando en la parte más cercana a MYH10, no se identifica una co-localización específica (**Figura 16A**), de hecho, parecen localizarse de manera mutuamente excluyente como se puede observar en los perfiles histográficos de inmunofluorescencias (**Figura 16B**).

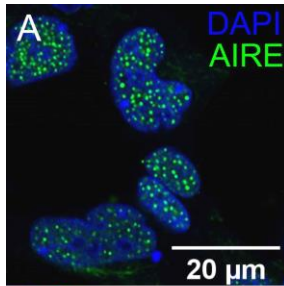
## 2.5 Distribución de AIRE a lo largo del ciclo celular

Durante los análisis por microscopía confocal de las diferentes localizaciones y distribuciones de AIRE en la célula, se identificaron distintas distribuciones de la proteína en distintas fases del ciclo celular. AIRE se presentó de manera habitual distribuido por todo el núcleo en forma de punteados nucleares durante la interfase celular (G0-G1), en algunos casos, apareció conjuntamente en el citoplasma formando fibras (**Figura 17A**). Respecto al ciclo celular, cuando las células HEK293-AIRE<sup>+</sup> comenzaban el proceso de división celular perdían la expresión de AIRE (**Figura 17B**) (Fase G2-M). Ya en mitosis, las células tornaron a expresar la proteína distribuyéndose por las cromátidas ya condensadas (**Figura 17C**) y asociada a los centriolos mientras se genera el huso acromático durante la profase y prometafase (**Figura 17D**). Una vez alineadas las cromátidas en el ecuador de la células, AIRE se observó más intensamente distribuido en los microtúbulos no-cinetocóricos y astrales (**Figura 17E**). Posteriormente, durante la anafase siguió localizada a lo largo de los microtúbulos no-cinetocóricos y astrales conforme las cromátidas se dirigieron a los polos celulares de las células hijas (**Figura 17F-H**). Finalmente AIRE siguió asociado a los microtúbulos durante la telofase, tanto al inicio de la ingresión del surco citocinético (**Figura 17H**), como durante la citocinesis (**Figura 17I**) y la escisión de membranas dando lugar a las células hijas (**Figura 17J**).

Nuestros datos sugieren que AIRE participa a lo largo del proceso de división celular en células humanas, desde su inicio y hasta la citocinesis. La presencia de los primeros 180 aminoácidos, que incluyen el dominio HSR/CARD y el motivo NLS, son esenciales para la

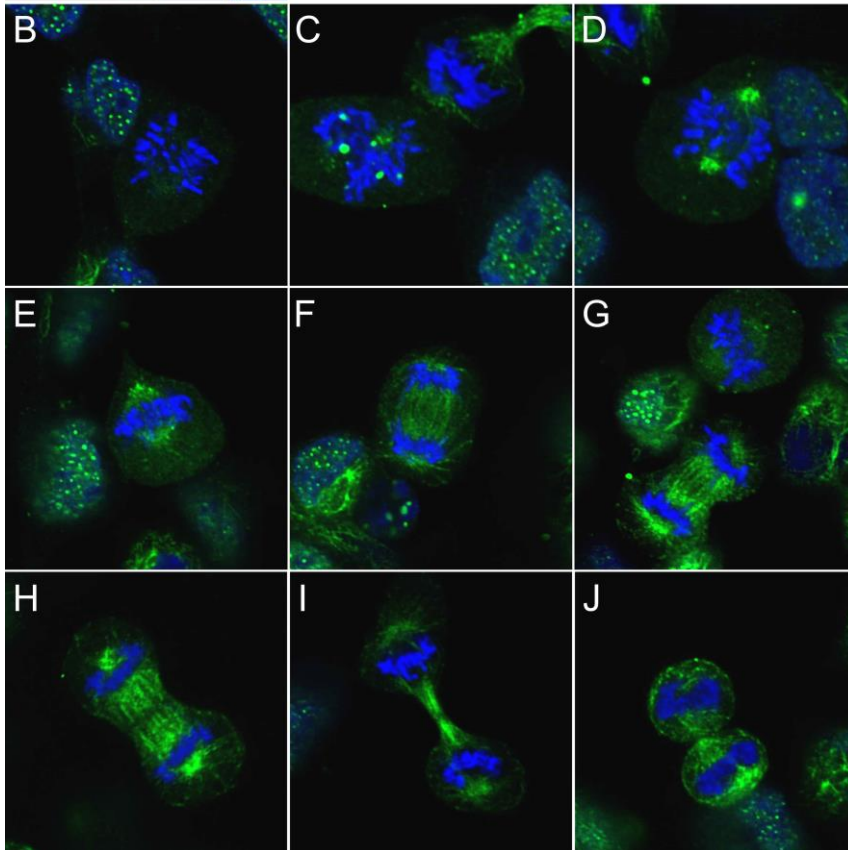


sublocalización de AIRE en la célula durante la mitosis y durante la citocinesis. La expresión de AIRE podría afectar al *checkpoint* G2-M puesto que el porcentaje de células que se encuentran en transfectantes estables de AIRE<sup>+</sup> y AIRE-Myc/His<sup>+</sup> respecto a células control se ve incrementado. Los dominios que pueden estar implicados en este efecto aún no están determinados.



**Figura 17. Localización de AIRE a lo largo de la división celular.**

Immunofluorescencia de células HEK293-AIRE<sup>+</sup>. Las células fueron fijadas y teñidas con anticuerpos  $\alpha$ AIRE (Alexa488, verde) y DAPI para la visualización de los núcleos celulares. DAPI (azul), MERGE (combinación de canales azul y verde), *Brightfield* BF (negro). Las imágenes fueron tomadas a 60X con un aumento de de zoom 4X. La progresión de las imágenes es de arriba abajo y de izquierda a derecha.









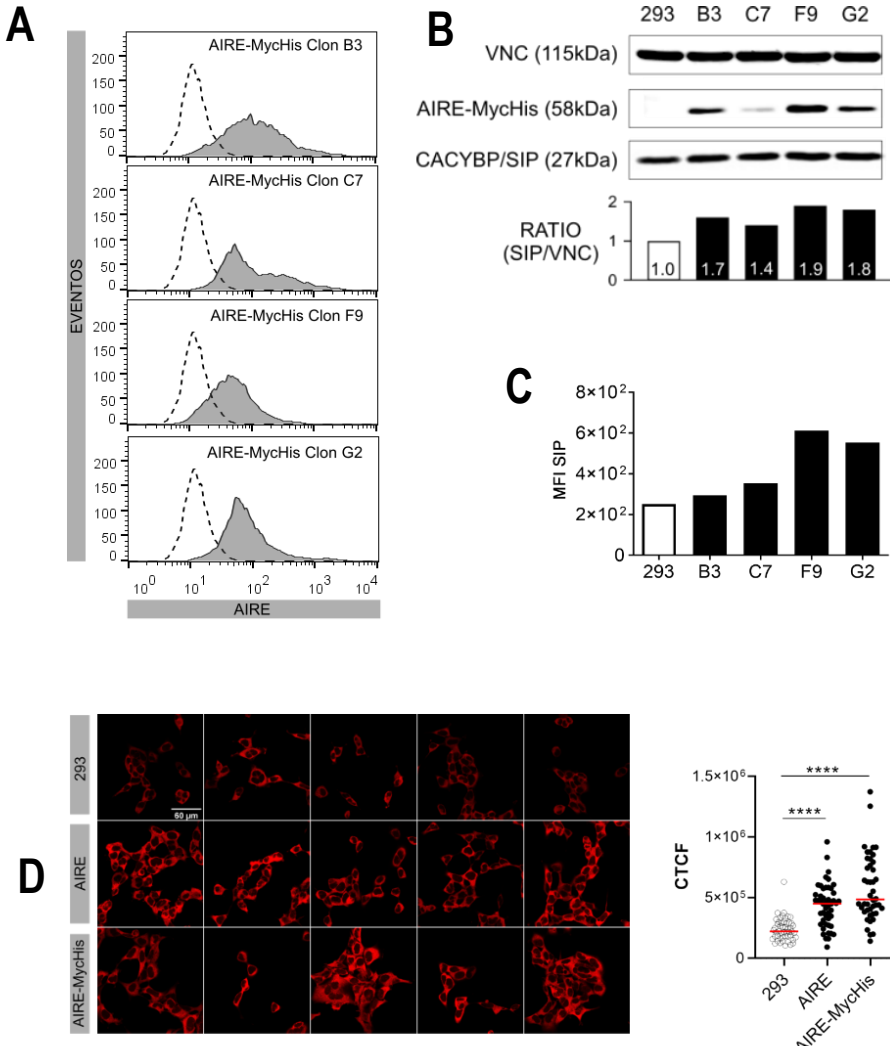
## CAPÍTULO III: Caracterización de la interacción de AIRE y proteínas pertenecientes a la familia de E3 Ubiquitín ligasas Seven in Absentia Homologue (SIAH)

### 3.1 CACYBP/SIP se ve incrementada en clones de transfectantes estables de AIRE-Myc/His

Como se ha comentado previamente, en un estudio previo de nuestro grupo de investigación se analizaron mediante 2 técnicas de proteómica cuantitativa (*2-dimensional difference gel electrophoresis* (2D-DIGE) e *Isotope-Coded Protein Label* (ICPL)) los proteomas de células de la línea celular de origen tiroideo (HT93) transfectadas con AIRE respecto a células no transfectadas. Se observó que aquellas células que expresaban AIRE presentaban un aumento en la expresión de algunas chaperonas como HSC70, HSP27 y de otras proteínas relacionadas con apoptosis como PCD5<sup>110</sup>. Una de las proteínas que se detectó incrementada por 2D-DIGE fue la CACYBP/SIP, proteína adaptadora de la E3-Ubiquitín ligasa SIAH1. Asimismo, se observó una disminución de proteínas relacionadas con el citoesqueleto. Este patrón era compatible con un aumento de muerte celular, que se confirmó mediante ensayos de apoptosis

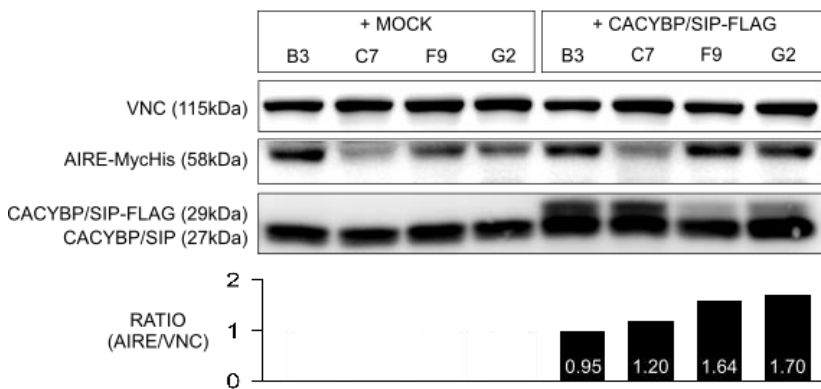
Para confirmar que el incremento en la abundancia de CACYBP/SIP es dependiente de AIRE se generó un transfectante estable en la línea HEK293 que expresaba AIRE con un tag compuesto de Myc/His (HEK293-AIRE-Myc/His). Se realizó un clonaje por dilución límite del *pool* transfectado inicial y se seleccionaron 4 clones donde el 100% de las células expresaban AIRE (**Figura 18A**). Se detectó que la abundancia de CACYBP/SIP aumentó en los transfectantes de AIRE, tanto mediante análisis por *western blot* (**Figura 18B**) como por citometría de flujo (**Figura 18C**). Además, el incremento de los niveles de CACYBP/SIP era dependiente de los niveles de expresión de la proteína AIRE (**Figura 18B**). Finalmente, se analizó por inmunofluorescencia uno de los clones descritos y un clon de un transfectante generado que expresaba la proteína AIRE sin ningún *tag*. Como se aprecia en la **Figura 18D**, CACYBP/SIP se ve incrementada en

ambos transfectantes de forma similar, lo que indica que este incremento es debido a la expresión de AIRE y el tag no ejerce un efecto significativo.



**Figura 19. CacyBP/SIP se ve incrementada en transfectantes estables de AIRE.** (A) Niveles de expresión de AIRE en cada clon transfectante estable AIRE-MyHis (AIRE=FITC), la línea discontinua indica la fluorescencia asociada a células HEK293 sin transfectar. (B) *Western blot* y densitometría del nivel de expresión de CacyBP/SIP de los diferentes clones. Imágenes representativas de (n=3) experimentos. RATIO indica la relación entre CacyBP/SIP:VINCULINA. (C) Representación gráfica de la media de intensidad de fluorescencia (MFI) de la proteína CacyBP/SIP en diferentes clones por Citometría de Flujo. Datos representativos de un experimento. (D) Inmunofluorescencia de células HEK293, HEK293-AIRE<sup>+</sup> y HEK293-AIRE-MyHis<sup>+</sup>. Las células fueron fijadas y teñidas con anticuerpos  $\alpha$ CacyBP/SIP (Alexa568, rojo). Las imágenes fueron tomadas a 60X. (CTCF) Fluorescencia Celular Total Corregida de n=50 células aleatorias entre las diferentes imágenes. Análisis estadístico mediante Test de Mann-Whitney (\* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , \*\*\*\* $P < 0.0001$ ).

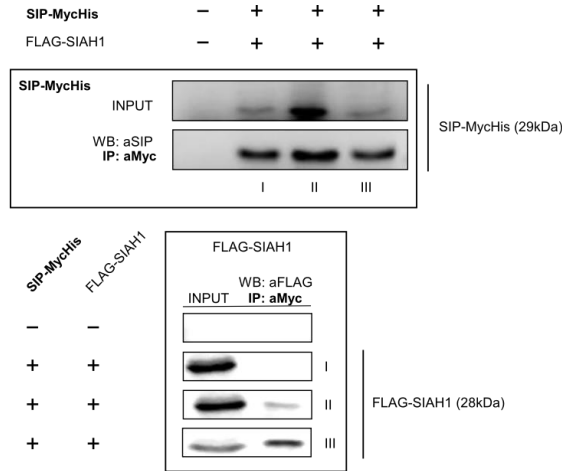
Se analizó asimismo el papel de la expresión de CACYBP/SIP sobre la abundancia de AIRE. Para ello se clonó el gen de CACYBP/SIP con un *tag* de FLAG (CACYBP/SIP-FLAG); los clones que expresaban AIRE de forma estable se transfectaron con dicha construcción, de manera que se sobreexpresó en ellos dicha proteína. Se midió la abundancia de AIRE en los transfectantes transitorios de CACYBP/SIP-FLAG mediante *western blot* y se observó un incremento en los niveles de AIRE en dichos transfectantes (Figura 19). Por tanto, el hecho de que la expresión de AIRE induzca un incremento de CACYBP/SIP y viceversa, sugiere que hay un mecanismo regulatorio entre dichas proteínas.



**Figura 19. El aumento en los niveles de CacyBP/SIP de manera exógena afecta a los niveles de expresión de AIRE.** Desde el panel superior al inferior: *Western blot* y densitometría del nivel de expresión de AIRE-MycHis de los diferentes clones tras una transfección transitoria con CacyBP/SIP (Imagen representativa de n=2). RATIO indica la relación entre AIRE:VINCULINA.

Se comprobó si ambas proteínas interaccionaban mediante experimentos de co-inmunoprecipitación. Inicialmente, se puso a punto un protocolo de co-inmunoprecipitación para optimizar las condiciones experimentales. Para ello, se usaron dos construcciones: SIAH1 con un *tag* de FLAG (FLAG-SIAH1), y se clonó el gen CACYBP/SIP con el *tag* Myc/His (CACYBP/SIP-Myc/His). La interacción entre ambas proteínas está bien establecida. Se transfectaron ambos genes en la línea HEK293 y a las 48h se capturó SIAH1 mediante una inmunoprecipitación con un anticuerpo anti-FLAG y se detectó CACYBP/SIP mediante *western blot* con un anticuerpo anti-

CACYBP/SIP. Se realizaron 3 protocolos diferentes de lisado celular e inmunoprecipitación siguiendo la bibliografía previa sobre el análisis de interacciones entre SIAH1 y CACYBP/SIP y AIRE y diferentes proteínas.



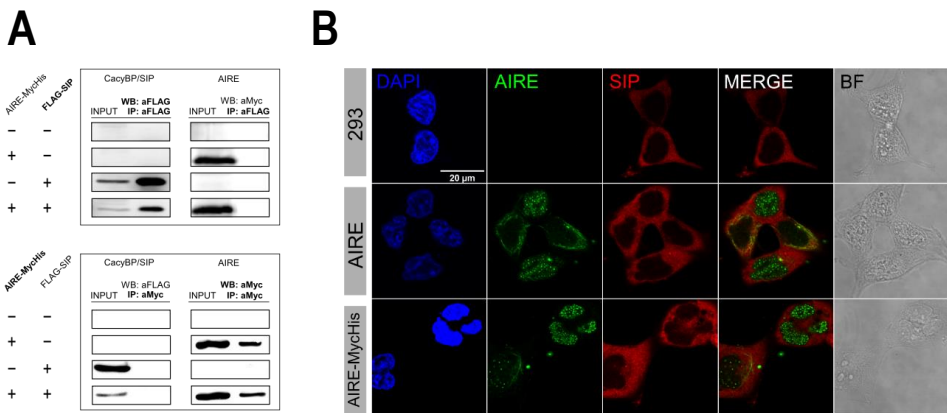
**Figura 20. La interacción entre CACYBP/SIP y SIAH1 es detectada con diferentes protocolos de inmunoprecipitación.** Células HEK293 que expresaban transitoriamente CacyBP/SIP y SIAH1 recombinantes con los *tags* MycHis y FLAG, respectivamente fueron lisadas, el lisado fue purificado con *beads* magnéticas  $\alpha$ Myc. El producto de la captura y toda su red de interacción fue procesado y analizado por *Western blot* usando los anticuerpos indicados. SIAH1 es capturada junto con CacyBP/SIP de manera eficiente bajo las condiciones óptimas de captura de proteínas que interaccionan con AIRE. Las condiciones I, II, III hacen referencia a las condiciones de lisis y co-inmunoprecipitación. Protocolo I. Lisado con *buffer* de lisis celular + *buffer* de extracción nuclear, co-inmunoprecipitación con *beads* de Sefarosa Protocolo II. Lisado con *buffer* de lisis celular, co-inmunoprecipitación con *beads* magnéticas,. Protocolo III. Fijación celular con PFA 1%, lisado celular con *buffer* de lisis celular + sonicación, co-inmunoprecipitación con *beads* magnéticas. Todos los *bufferes* de lisis contienen inhibidores de proteasas.

El protocolo que funcionó mejor, y que se seleccionó para experimentos futuros fue aquel que empleaba una breve fijación celular con PFA 1% previa a la lisis por sonicación (Figura 20).

Una vez se determinó la metodología óptima, se evaluó la interacción directa entre AIRE y CACYBP/SIP. Se transfectaron transitoriamente ambas proteínas (AIRE-Myc/His y CACYBP/SIP-FLAG) en HEK293 y se inmunoprecipitó el lisado celular resultante con los anticuerpos específicos de los *tags*. No se logró detectar una interacción aparente entre



ambas proteínas (**Figura 21A**). Por tanto, AIRE y CACYBP/SIP no parecen interactuar directamente entre ellas. Asimismo, se realizaron inmunofluorescencias sobre células HEK293, HEK293-AIRE<sup>+</sup> y HEK293-AIRE-MycHis<sup>+</sup> para determinar la existencia de co-localización entre ambas proteínas, AIRE mostró una localización típica, mayoritariamente en *dots* nucleares y, en algunas células, asociado a fibras en el citoplasma; CACYBP/SIP, en cambio, mostraba un patrón de tinción uniforme y distribuido por todo el citoplasma. En ningún caso se detectó co-localización entre AIRE y CACYBP/SIP (**Figura 21B**).



**Figura 21. AIRE no interacta con CACYBP/SIP.** (A) Co-inmunoprecipitación de AIRE-MycHis y FLAG-CacyBP/SIP. El panel superior hace referencia a la inmunoprecipitación frente al *tag* Myc. El panel inferior frente al *tag* FLAG. INPUT: lisado celular previo a la inmunoprecipitación. IP: lisado purificado por el anticuerpo correspondiente. WB: anticuerpo utilizado para el revelado de la membrana. La captura y detección de ambas proteínas es correcta, en la condición de co-transfección no se observa interacción específica. Imágenes representativas de (n=3) experimentos. (B) Inmunofluorescencia de células HEK293, HEK293-AIRE<sup>+</sup> y HEK293-AIRE-MycHis<sup>+</sup>. Las células fueron fijadas y teñidas con anticuerpos  $\alpha$ AIRE (Alexa488, verde) y  $\alpha$ CacyBP/SIP (Alexa568, rojo) y DAPI para la visualización de los núcleos celulares. DAPI (azul), SIP (rojo), AIRE (verde), MERGE (combinación de canales rojo y verde), *Brightfield* BF (negro). Las imágenes fueron tomadas a 60X con un aumento de zoom 4X.

### 3.2 AIRE interacta con las proteínas de la familia Seven in Absentia Homologue (SIAH): SIAH1 y SIAH2

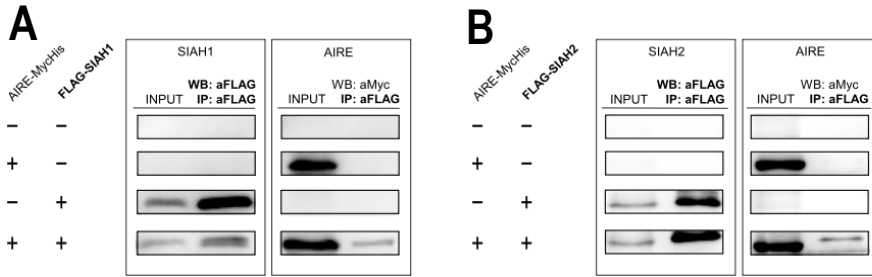
Las proteínas que interactúan con SIAH1 presentan una secuencia consenso. Dicho motivo está constituido por la secuencia PxAxVxP, siendo x cualquier aminoácido. Se

realizó un análisis de la secuencia de AIRE en busca de dicha secuencia. Se detectaron dos posibles motivos de interacción con SIAH: el primer motivo consenso de unión se encuentra entre los aminoácidos 19 y 125 y presenta la secuencia **PKALVPP**; el segundo motivo consiste en la secuencia **PLLCVGP** y se localiza entre los residuos 416 y 422 (Figura 22). El primer motivo queda situado entre los dominios HSR/CARD y SAND, mientras que el segundo motivo se encuentra solapado con uno de los denominados como motivos ricos en leucinas (motivos LxxLL).

	10	20	30	40	50
MATDAALRRL	LRLHRTEIAV	AVDSAPFLH	ALADHDVPE	DKFQETLHLK	
60	70	80	90	100	
EKEGCPQAFH	ALLSWLLTQD	STAILDFWRV	LFKDYNLERY	GRLQPILDSF	
110	120	130	140	150	
PKDVDLSQPR	KGRKPPAV <b>PK</b>	<b>LLLE</b> PRLPT	KRKASEEARA	AAPAALTFRG	
160	170	180	190	200	
TASPGSQLKA	KPPKPESSA	EQQRPLGNG	IQTMSASVQR	AVAMSSGDVP	
210	220	230	240	250	
GARGAVEGIL	IQQVFESGGS	KKCIQVGGEF	YTPSKFEDSG	SGKNKARSSS	
260	270	280	290	300	
GPKPLVRAKG	AQGAAPGGGE	ARLGQQGSVP	APLALPSDEQ	LHQKNEDECA	
310	320	330	340	350	
VCRDGGELIC	CDGCPRAFHL	ACLSPLPREI	PSGTWRCSSC	LQATVQEVQP	
360	370	380	390	400	
RAEEPRPQEP	FVETPLPPGL	RSAGEEVRGP	PGEPLAGMDT	TLVYKHLPA	
410	420	430	440	450	
PSAAPLPGLD	SSALH <b>LLC</b>	<b>GE</b> EGQNLAP	GARGVCGDG	TDVLRCTHCA	
460	470	480	490	500	
AAFHWCHFP	AGTSRPGTGL	RCRSCSGDVT	PAPVEGLVAP	SPARLAPGPA	
510	520	530	540		
KDDTASHEPA	LHRDDLESLL	SEHTFDGILQ	WAIQSMARPA	APFPS	

**Figura 22. AIRE presenta dos posibles motivos de unión a SIAH en su secuencia.** Secuencia aminoácida de la proteína AIRE. Subrayados en rojo los diferentes aminoácidos que conforman el motivo de unión. El primer motivo abarca desde el residuo P119 hasta el residuo P125. El segundo motivo abarca desde el residuo P416 hasta el residuo P422.

Para demostrar una interacción entre las proteínas AIRE y SIAH1, se transfectaron transitoriamente células HEK293 con AIRE-Myc/His y FLAG-SIAH1 y se realizó una inmunoprecipitación con un anticuerpo anti-FLAG y la detección por *western blot* con anticuerpos anti-FLAG (como control) y anti-Myc (para estudiar la posible interacción). Los datos mostraron que había una interacción específica de AIRE con SIAH1 (Figura 23A). Debido a la alta similitud de secuencia en dominios críticos entre SIAH1 y SIAH2 y la existencia de numerosos sustratos que comparten ambos miembros de la familia de E3-Ubiquitin ligasas, se procedió a determinar si AIRE también interaccionaba con SIAH2. Para ello se usó una construcción similar a la de SIAH1 con un *tag* de FLAG (FLAG-SIAH2). Los resultados fueron similares a los obtenidos con SIAH1 (Figura 23B). Por tanto, AIRE interacciona específicamente con SIAH1 y SIAH2.

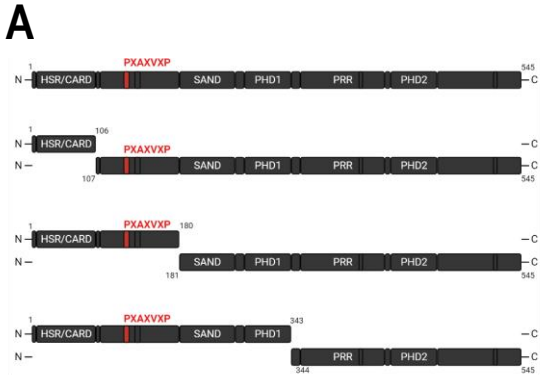


**Figura 23. AIRE-MycHis interacciona con FLAG-SIAH1 y FLAG-SIAH2.** (A) Co-inmunoprecipitación FLAG-SIAH1. (B) Co-inmunoprecipitación FLAG-SIAH2. INPUT: lisado celular previo a la inmunoprecipitación. IP: lisado purificado por el anticuerpo  $\alpha$ FLAG. WB: anticuerpo utilizado para el revelado de la membrana. La captura y detección de ambas proteínas es correcta, en la condición de co-transfección se observa interacción específica.

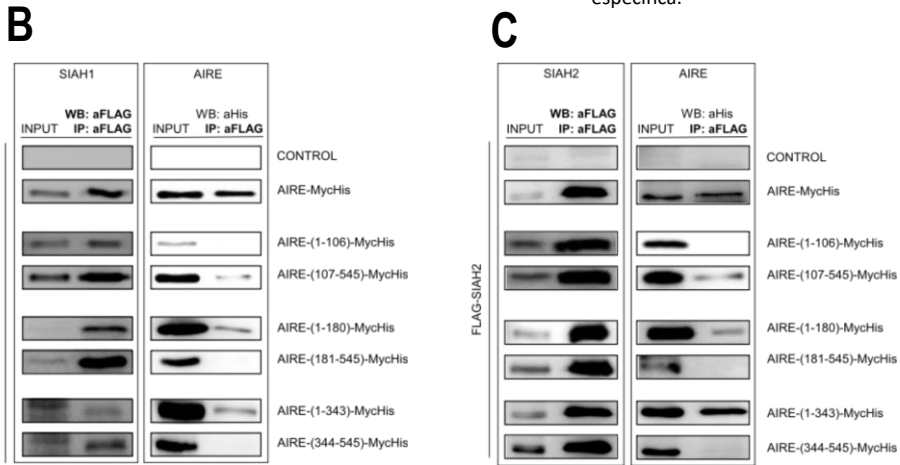
### 3.3 AIRE requiere la región que contiene la primera secuencia consenso (PKALVPP) para la interacción con SIAH1 y SIAH2

Para dilucidar si esta interacción es debida a la presencia de alguno de los motivos de unión a SIAH1 presentes en AIRE se usaron varias de las construcciones que contienen diversas formas de AIRE truncadas con diferentes motivos de la proteína con el *tag* Myc/His. Se repitieron los experimentos de co-inmunoprecipitación utilizando las proteínas AIRE truncadas, conteniendo algunas de ellas el primer motivo de unión a SIAH1 (con la secuencia canónica completa) y la contra-parte que no lo contiene (**Figura 24A**). SIAH1 interaccionó con las proteínas AIRE truncadas que contenían los residuos 1-180, 1-343 y 107-545, que portan el primer motivo de unión a SIAH1 pero no con la que contiene únicamente el segundo. Las proteínas de AIRE truncadas (1-106, 181-545 y 344-545) que no contenían ningún motivo de unión a SIAH1 o solamente el segundo posible motivo de unión no lograron capturarse tras la inmunoprecipitación (**Figura 24B**). Por tanto, los datos mostraron que la interacción con dicha E3 Ubiquitin ligasa se realiza a través de la región que contiene el primer motivo de unión **PKALVPP** (residuos 119-125) y no con el segundo motivo **PLLCVGP** (residuos 416-422). Los mismos experimentos se repitieron con SIAH2, mostrando resultados similares. Así, solamente las proteínas truncadas que contenían el primer motivo consenso de unión co-precipitaron con SIAH2 (**Figura 24C**). En definitiva, la región que contiene el motivo

PKALVPP localizado entre los residuos 119 y 125 de AIRE es necesaria para la interacción con las proteínas SIAH.

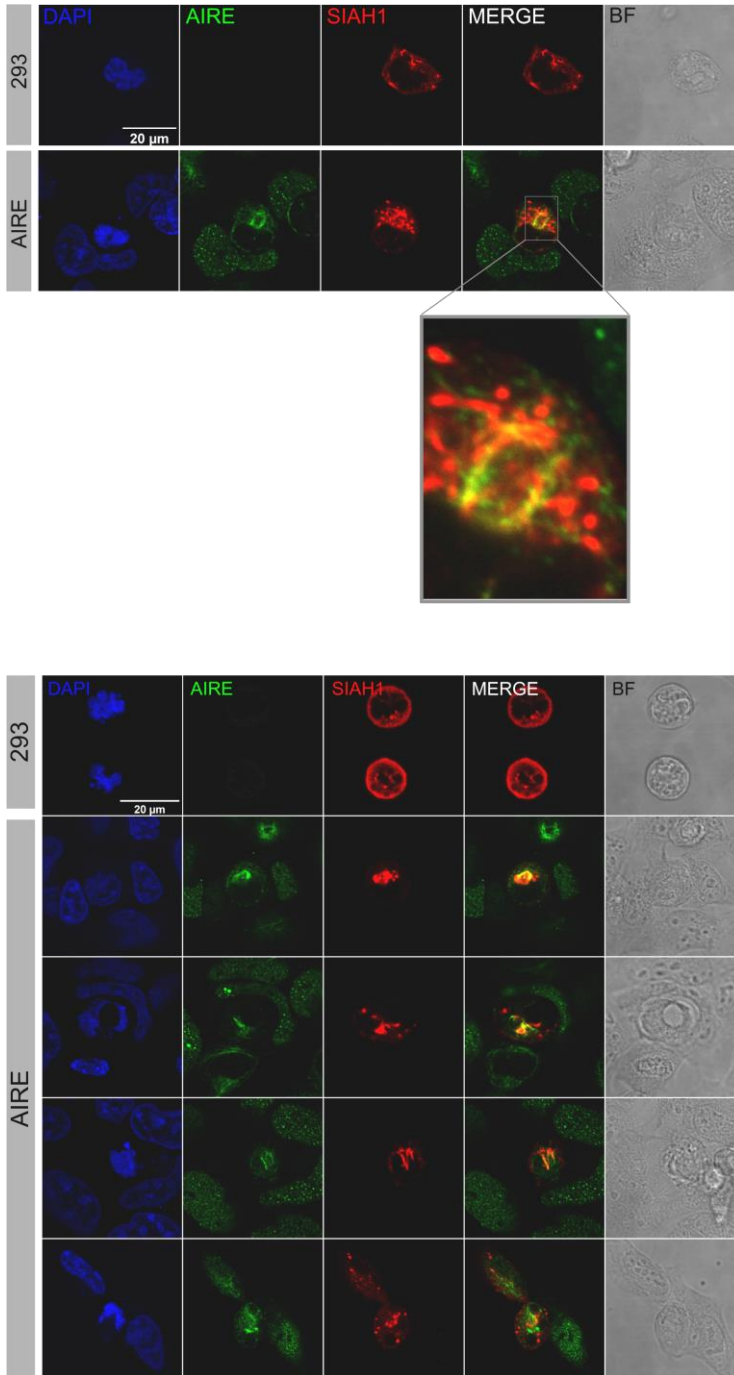


**Figura 24. AIRE interacciona con SIAH1 y SIAH2 y lo hace a través de dominios que contengan el primer motivo de unión.** (A) Esquema representativo de las proteínas de AIRE completa y las proteínas con los diferentes dominios truncados. (B) Co-inmunoprecipitación con FLAG-SIAH1. (C) Co-Inmunoprecipitación con FLAG-SIAH2. INPUT: lisado celular previo a la inmunoprecipitación. IP: lisado purificado por el anticuerpo aFLAG. WB: anticuerpo utilizado para el revelado de la membrana. La captura y detección de ambas proteínas es correcta, en la condición de co-transfección se observa interacción específica.

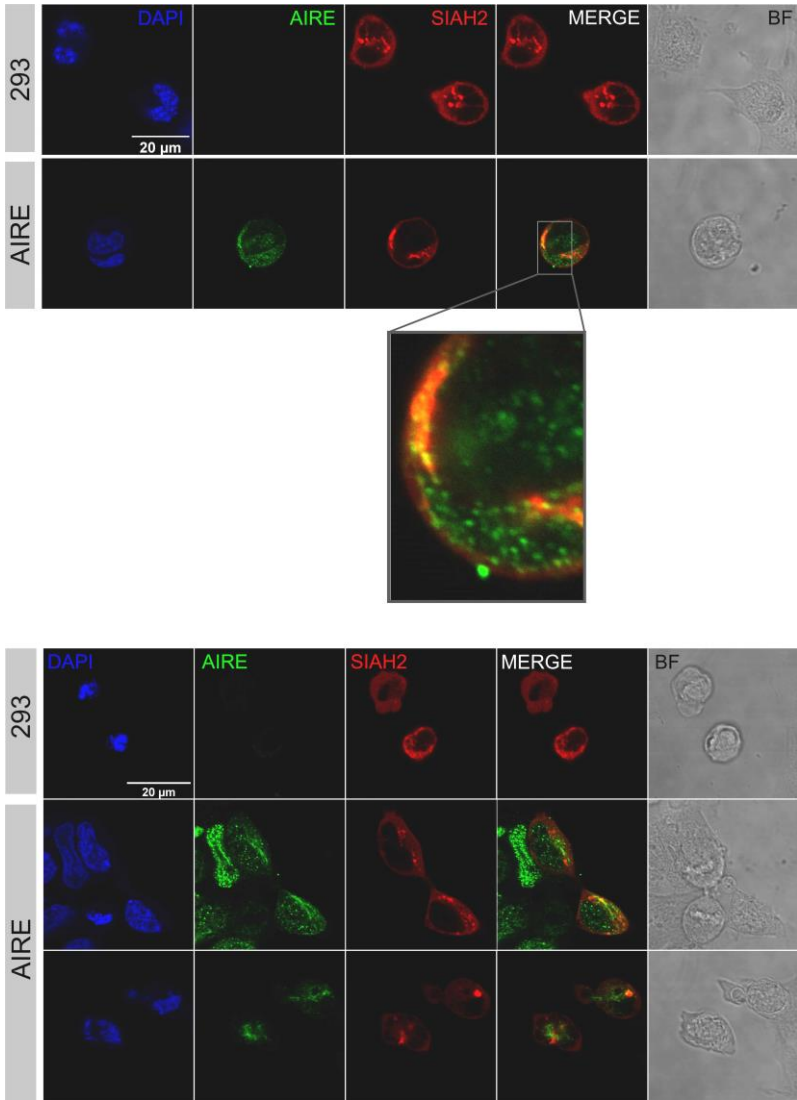


### 3.4 AIRE co-localiza con las proteínas SIAH1 y SIAH2.

Se analizó mediante microscopía confocal la co-localización entre AIRE, SIAH1 y SIAH2. Mediante microscopía confocal e inmunofluorescencia tras una transfección transitoria de ambas E3-Ubiquitín ligasas en células que expresan AIRE de manera estable se detectó co-localización entre AIRE y SIAH1 (Figura 25) y AIRE y SIAH2 (Figura 26). Las co-localizaciones detectadas se concentraban en el citoplasma, en la zona perinuclear. No se logró detectar presencia de SIAH1 y SIAH2 en el núcleo.



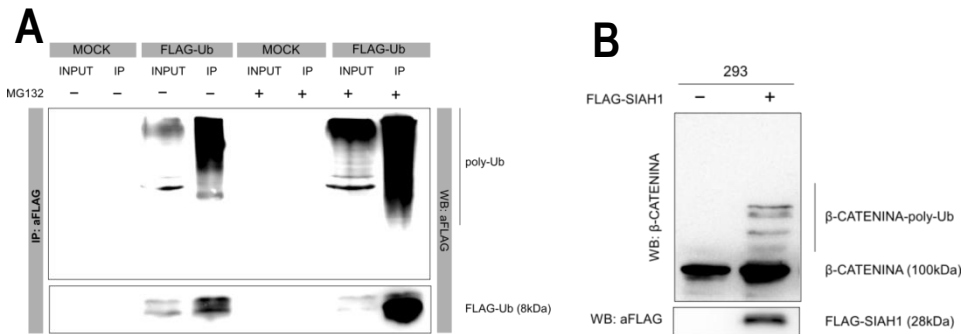
**Figura 25. AIRE co-localiza con SIAH1 en células HEK293-AIRE<sup>+</sup>.** Inmunofluorescencia de células HEK293 y HEK293-AIRE<sup>+</sup> transfectadas transitoriamente con FLAG-SIAH1. Las células fueron fijadas y teñidas con anticuerpos  $\alpha$ AIRE (Alexa488, verde) y  $\alpha$ FLAG (Alexa568, rojo) y DAPI para la visualización de los núcleos celulares. DAPI (azul), FLAG-SIAH1 (rojo), AIRE (verde), MERGE (combinación de canales rojo y verde), *Brightfield* BF (negro). Las imágenes fueron tomadas a 60X o a 60X con un aumento de zoom 2X. Imágenes representativas de (n=3) experimentos.



**Figura 26. AIRE co-localiza con SIAH1 en células HEK293-AIRE<sup>+</sup>**. Inmunofluorescencia de células HEK293 y HEK293-AIRE<sup>+</sup> transfectadas transitoriamente con FLAG-SIAH2. Las células fueron fijadas y teñidas con anticuerpos  $\alpha$ AIRE (Alexa488, verde) y  $\alpha$ FLAG (Alexa568, rojo) y DAPI para la visualización de los núcleos celulares. DAPI (azul), FLAG-SIAH2 (rojo), AIRE (verde), MERGE (combinación de canales rojo y verde), *Brightfield* BF (negro). Las imágenes fueron tomadas a 60X o a 60X con un aumento de zoom 2X. Imágenes representativas de (n=3) experimentos.

### 3.5 SIAH1 incrementa los niveles de AIRE poli-ubiquitinado

Como la función definida de las proteínas SIAH son la ubiquitinación de sus sustratos, se evaluó si la interacción de AIRE con dichas proteínas producía su ubiquitinación. Para ello, se realizaron una serie de experimentos control para determinar que las condiciones experimentales para la detección de ubiquitinación específica eran correctas. Se clonó el gen de ubiquitina humana con un *tag* de FLAG (FLAG-Ub). Se comprobó que la acción del inhibidor del proteasoma MG-132 sobre el sistema ubiquitín-proteasoma era eficiente y lograba incrementar los niveles de proteínas ubiquitinadas al no poder degradarse vía proteasoma, detectando FLAG-Ub en células HEK293 transfectadas transitoriamente con el vector pcDNA3.1 vacío o FLAG-Ub y el tratamiento con MG-132 o DMSO (Figura 27A). Los niveles tanto de Ubiquitina como de proteínas ubiquitinadas aumentaron al inhibir la vía ubiquitín-proteasoma.

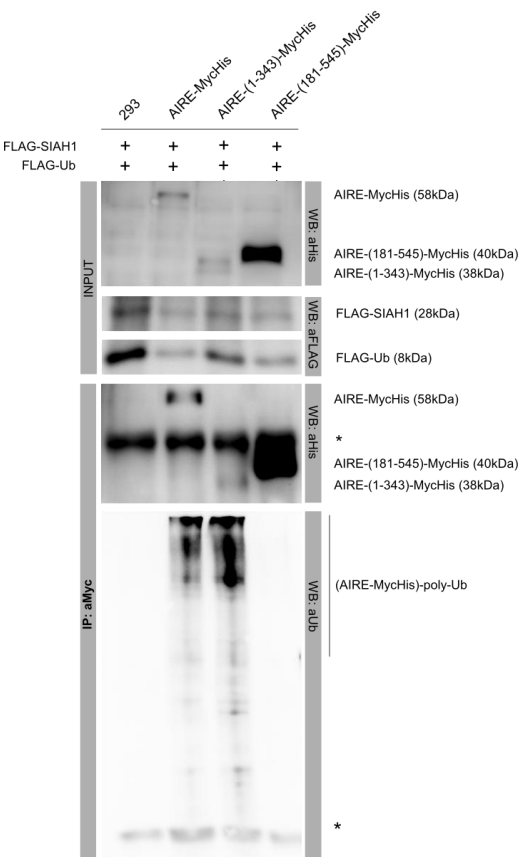


**Figura 27. Optimización de la detección de ubiquitinación específica.** Análisis por *Western Blot* de lisados celulares de transfectantes transitorios de células HEK293 y HEK293-FLAG-SIAH1<sup>+</sup>. (A) Efecto de la aplicación de MG-132 20uM 4h en el bloqueo del proceso de degradación del sistema ubiquitín-proteasoma. MOCK hace referencia a la transfección del vector pcDNA3.1 vacío. (B) Efecto de co-transfección de SIAH1 en los niveles de  $\beta$ -Catenina endógena. WB: anticuerpo utilizado para el revelado de la membrana.

La proteína  $\beta$ -Catenina es un sustrato específico de SIAH1 y es degradada por el proteasoma a través de la interacción de ésta con la proteína adaptadora CACYBP/SIP<sup>209</sup>. Por tanto, se estudió la ubiquitinación de  $\beta$ -Catenina endógena en células HEK293 tras la transfección transitoria de FLAG-SIAH1, como puesta a punto de

la técnica. La transfección transitoria de FLAG-SIAH1 en células HEK293 incrementó la poliubiquitinación de  $\beta$ -catenina (**Figura 27B**).

Se determinó, por un lado, si SIAH1 ubiquitinaba específicamente AIRE y, por otro, si la presencia del motivo de unión a SIAH1 en la secuencia de AIRE era necesaria para que la E3 Ubiquitin Ligasa fuese capaz de reconocer y ubiquitinar a AIRE. Se realizó un experimento de ubiquitinación utilizando la proteína completa de AIRE-Myc/His (Clon B3) y proteínas de AIRE-Myc/His que contenían diversos dominios con el motivo de

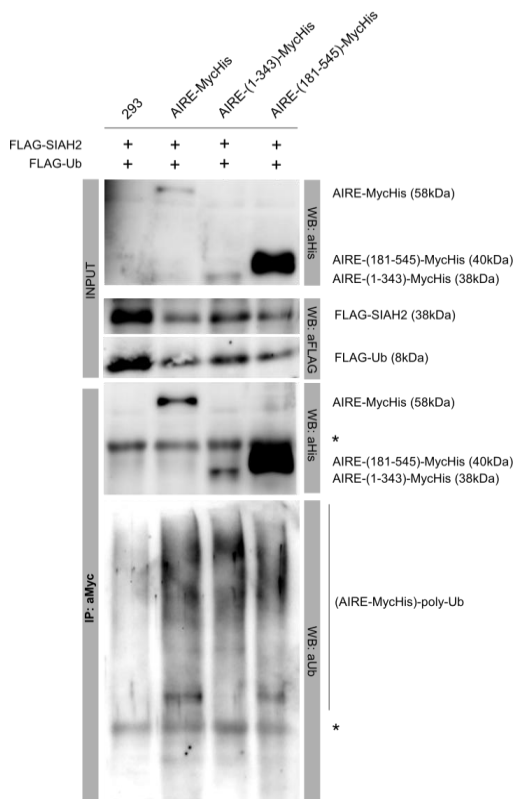


unión a SIAH1 (AIRE-(1-343)-MycHis) o sin el mismo (AIRE-(181-545)-MycHis). Se realizó una transfección transitoria con FLAG-SIAH1 y FLAG-Ub, seguida de una inmunoprecipitación de los lisados celulares con el anticuerpo anti-Myc. Se observó una mayor recuperación de AIRE ubiquitinado en los lisados celulares de AIRE-Myc/His y AIRE-(1-343)-Myc/His pero no en lisados celulares control y de AIRE-(181-545)-MycHis (**Figura 28**). Por tanto, los datos indican que SIAH1 ubiquitina AIRE y se requiere la región que contiene el motivo consenso de unión a SIAH1 en la secuencia de AIRE para su ubiquitinación específica.

**Figura 28. El motivo consenso de unión a SIAH1 es necesario para la ubiquitinación específica de AIRE.** Análisis por *Western Blot* de lisados celulares de transfectantes transitorios de células HEK293, HEK293-AIRE-MycHis<sup>+</sup>, HEK293-AIRE-(1-343)-MycHis<sup>+</sup> y HEK293-AIRE-(181-545)-MycHis<sup>+</sup>. INPUT: lisado celular previo a la inmunoprecipitación. IP: lisado celular purificado con el anticuerpo  $\alpha$ FLAG. WB: membrana revelada con el anticuerpo indicado. (\*) Bandas inespecíficas de la cadena pesada del anticuerpo utilizado para la inmunoprecipitación.



En el caso de SIAH2, cuando inmunoprecipitamos con el anticuerpo anti-Myc se observó un ligero incremento en la recuperación de AIRE ubiquitinado en los lisados celulares de AIRE-MycHis y AIRE-(1-343)-MycHis comparado con los otros lisados celulares (control y AIRE-(181-545)-MycHis) (Figura 29). Sin embargo, los datos no son tan concluyentes como con SIAH1. Aún así, los datos obtenidos sugieren que SIAH2 es también capaz de ubiquitinar AIRE y la presencia del motivo consenso de unión a SIAH es necesario para dicho proceso (Figura 29).

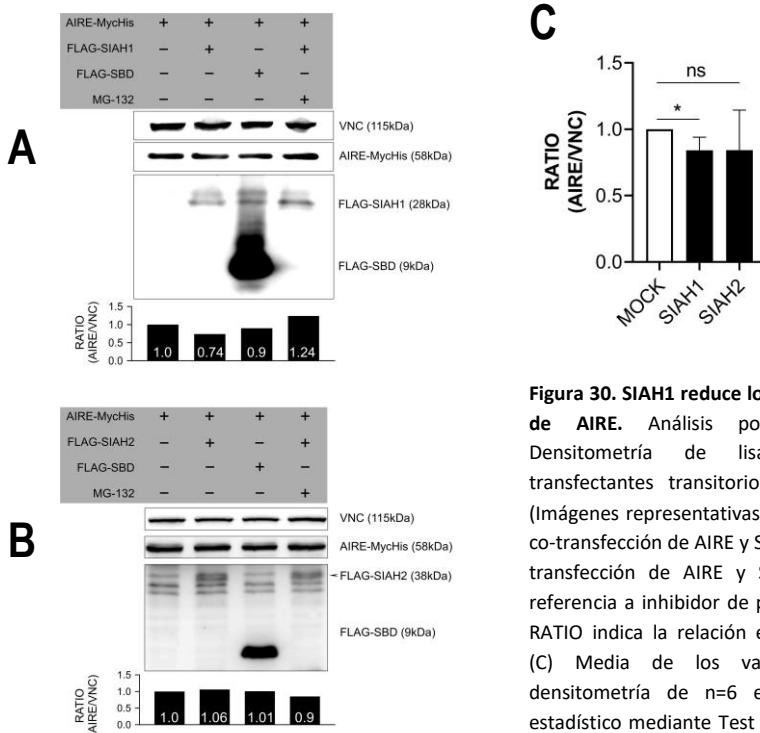


**Figura 29. El motivo consenso de unión a SIAH1 es necesario para la ubiquitinación específica de AIRE.** Análisis por *Western Blot* de lisados celulares de transfectantes transitorios de células HEK293, HEK293-AIRE-MycHis<sup>+</sup>, HEK293-AIRE-(1-343)-MycHis<sup>+</sup> y HEK293-AIRE-(181-545)-MycHis<sup>+</sup>. INPUT: lisado celular previo a la inmunoprecipitación. IP: lisado celular purificado con el anticuerpo αFLAG. WB: membrana revelada con el anticuerpo indicado. (\*) Bandas inespecíficas de la cadena pesada del anticuerpo utilizado para la inmunoprecipitación.

### 3.6 SIAH1 reduce los niveles intracelulares de AIRE

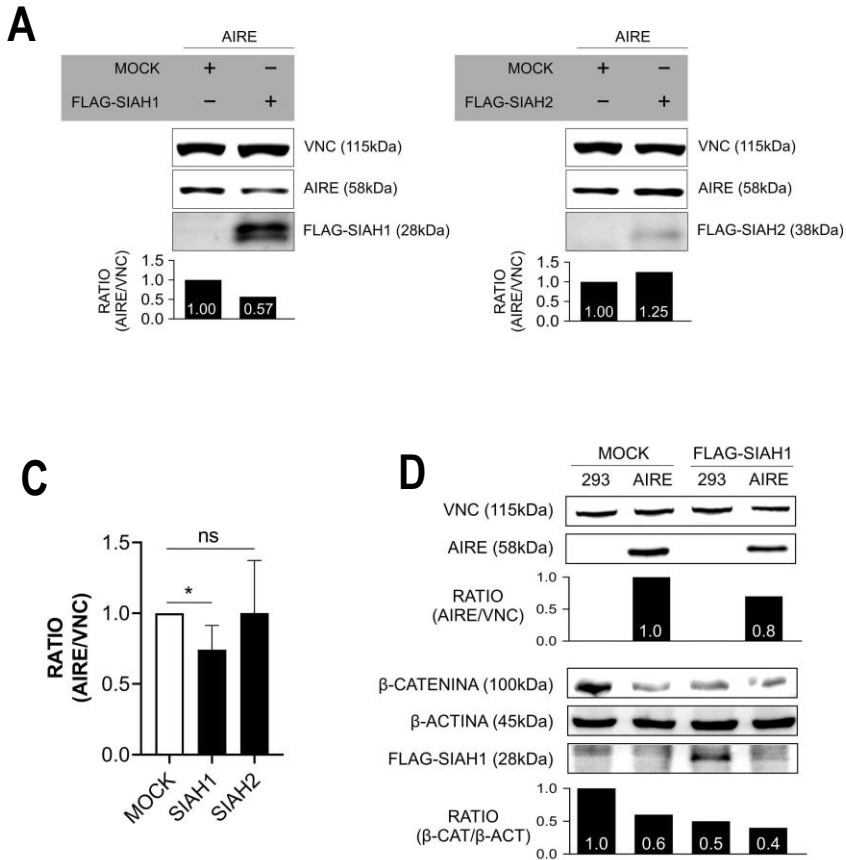
La interacción de AIRE con las E3 Ubiquitin Ligasas de la familia SIAH podría sugerir la existencia de una nueva vía de regulación en los niveles de AIRE intracelular. Por ello, se estudió si la abundancia de AIRE disminuye por la expresión de SIAH1 y SIAH2. Se transfectaron transitoriamente células HEK293 las construcciones AIRE-Myc/His, FLAG-SIAH1, FLAG-SBD y FLAG-SIAH2. El FLAG-SBD consiste en el gen de SIAH1 sin los

primeros 89 residuos (dominio RING), lo que le permite interactuar con las proteínas diana pero sin la capacidad de ubiquitinarlas. La cantidad de AIRE disminuyó al transfectarse con SIAH1 respecto de la transfección en solitario, mientras que la co-transfección con FLAG-SBD no produjo una disminución significativa de AIRE. Igualmente, la inhibición de la ruta ubiquitín-proteasoma con el tratamiento de MG-132 se tradujo en el mantenimiento de la abundancia de AIRE intracelular en presencia de SIAH1 (**Figura 30A**). Este proceso no pudo ser reproducido con SIAH2 (**Figura 30B**). La reducción de los niveles de AIRE al co-transfectar SIAH1 aunque leve, resulta consistente, específica y reproducible en múltiples experimentos, mientras que los resultados con SIAH2 son más variables y no se aprecia una reducción consistente (**Figura 30C**).



**Figura 30. SIAH1 reduce los niveles de expresión de AIRE.** Análisis por *Western Blot* y Densitometría de lisados celulares de transfectantes transitorios de células HEK293 (Imágenes representativas de n=6). (A) Efecto de co-transfección de AIRE y SIAH1. (B) Efecto de co-transfección de AIRE y SIAH2. (MG-132 hace referencia a inhibidor de proteasoma 20uM 4h). RATIO indica la relación entre AIRE:VINCULINA. (C) Media de los valores de AIRE por densitometría de n=6 experimentos. Análisis estadístico mediante Test de Wilcoxon (\* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , \*\*\*\* $P < 0.0001$ ).

Se llevó a cabo una aproximación similar con SIAH1 sobre un clon del transfectante estable de AIRE. Así, los niveles de expresión inicial de AIRE fueron los mismos en todas las condiciones y no dependieron de la eficiencia de transfección. Los niveles de AIRE se vieron reducidos cuando se expresó transitoriamente FLAG-SIAH1 (Figura 31A y 31B). Una vez más, esta reducción no se observó con SIAH2 (Figura 31A y 31B).



**Figura 31. Los niveles de expresión de AIRE en transfectantes estables de HEK293 se ven reducidos al transfectar transitoriamente SIAH1.** Análisis por *Western Blot* y Densitometría de lisados celulares de HEK293 y transfectantes estables de células HEK293-AIRE<sup>+</sup> (Imágenes representativas de n=6 experimentos). (A) Efecto de transfección transitoria de SIAH1 y SIAH2 sobre los niveles de AIRE. RATIO indica la relación entre AIRE:VINCULINA. (B) Media de los valores de AIRE por densitometría de n=6 experimentos. Análisis estadístico mediante Test Wilcoxon (\* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , \*\*\*\* $P < 0.0001$ ). (C) Efecto de transfección transitoria de SIAH1 sobre los niveles de AIRE y  $\beta$ -CATENINA. RATIO indica la relación entre AIRE:VINCULINA y  $\beta$ -CATENINA: $\beta$ -ACTINA. MOCK hace referencia a una transfección transitoria con el vector pCDNA3.1 vacío

Dado que la reducción fue moderada, se analizó el descenso de  $\beta$ -catenina endógena (diana bien establecida de SIAH1) como control. La disminución de la  $\beta$ -catenina fue similar a la obtenida con AIRE (**Figura 31C**). Por tanto, nuestros datos muestran que AIRE es una diana de SIAH1 para su degradación vía ubiquitín-proteasoma.





# DISCUSIÓN

Las TECs presentan una capacidad inusual de expresión promiscua de genes (PGE) independientemente del tipo de localización y estado de maduración dentro del tejido tímico<sup>236</sup>. Este hecho resulta esencial para la generación de un repertorio de linfocitos T inmunocompetente, suficientemente diverso pero tolerante a lo propio, de tal manera que en última instancia se eviten posibles respuestas autoinmunes. En las mTECs esta característica se encuentra potenciada por la expresión de AIRE, un regulador transcripcional único que actúa reconociendo marcas epigenéticas genéricas de sitios transcripcionalmente inactivos como la Lisina 4 no metilada de la Histona 3 (H3K4me0), la hipometilación de la Lisina 4 de la Histona 3 (H3K4me1), la acetilación de la Lisina 27 de la Histona 3 (H3K27ac) o promotores con la RNA polimerasa II (RNA-PolIII) encallada permitiendo la transcripción de prácticamente todo el genoma<sup>86 87 237</sup>.

A nivel molecular, AIRE se localizaría en regiones amplias de cromatina donde se aloja un elevado número de factores de transcripción generales y específicos de cada tipo celular, los denominados *super-enhancers*. Estas estructuras fomentan la transcripción de regiones intra-cromosómicas o inter-cromosómicas de manera simultánea al organizar proteínas y factores de transcripción a los sitios de inicio de transcripción (TSSs) de dichas regiones<sup>84</sup>. En los *super-enhancers* AIRE induciría la expresión de genes atraídos por marcas epigenéticas de represión, actuaría formando parte de un complejo multiproteico que incluye diversas proteínas, como DNA-PK, PARP-1, TOP1, TOP2, FACT, Ku80, Ku70,  $\gamma$ H2AX, CBP, BRD4 y RNA-PolIII. AIRE interaccionaría con TOP1 y TOP2a que favorecerían DSBs en el DNA, se induciría el reclutamiento a estas regiones de la maquinaria de respuesta a daño genotóxico (DNA-PK, PARP-1, FACT, Ku80, Ku70 y  $\gamma$ H2AX), todo ello, junto con el reclutamiento de otros factores como BRD4 y pTEFb permitiría, en última instancia, la activación del proceso de elongación de la RNA-PolIII<sup>63 83 85</sup>. Este modelo explicaría el efecto de la expresión de AIRE sobre diferentes tipos celulares: AIRE genera un impacto en la expresión de amplios sets de

genes pero los *loci* afectados es diferente en cada tipo celular<sup>107</sup>. Cada tipo de célula y su estado de maduración presenta un conjunto de diferentes promotores actuando sobre los *super-enhancers*, asimismo, los *super-enhancers* suelen distribuirse regulando genes que definen el tipo y el estado de diferenciación celular y la distribución de la RNA-PolIII no productiva a lo largo de la cromatina también es variable<sup>108</sup>. En líneas celulares monocíticas, AIRE logra inducir la expresión de, por un lado, genes propios de procesos de maduración de células dendríticas<sup>109</sup>. Estos datos están de acuerdo con datos previos de nuestro laboratorio donde se vio que células epiteliales transfectadas con AIRE vieron incrementadas la expresión de genes relacionados con la expresión de proteínas pro-apoptóticas y chaperonas mientras que algunas proteínas asociadas al citoesqueleto se ven disminuidas<sup>110</sup>, nuestro modelo celular también mostró este patrón de modificaciones.

En la presente tesis, se evaluó el impacto de la expresión estable de la proteína AIRE en otro tipo celular diferente a los anteriormente citados. El análisis se realizó por otra técnica de proteómica cuantitativa, iTRAQ, que se está usando en la actualidad con preferencia sobre las técnicas usadas en nuestro trabajo previo. Los resultados obtenidos por iTRAQ fueron similares a los obtenidos por DIGE e ICPL en el trabajo previo de nuestro laboratorio<sup>110</sup>. Así, Las células HEK293-AIRE<sup>+</sup> ven incrementada la expresión de proteínas chaperonas (calnexina, HSP74, peroxiredoxina o disulfuro isomerasa) mientras que la expresión de proteínas asociadas al citoesqueleto (transgelina, plectina 1, fascina o keratinas 18 y 19) se encuentra reducida respecto a las células HEK293 WT. También se observan diferencias en proteínas relacionadas con la apoptosis o ciclo celular. Los resultados de WB y citometría resultaron similares en los transfectantes de AIRE y de AIRE-Myc/His, lo que da mayor consistencia a que el efecto observado es debido intrínsecamente a AIRE y no a otro posible efecto derivado de la transfección, como una posible inserción del gen exógeno en alguna localización del genoma que afecte a la fisiología celular.



El incremento de la apoptosis en las células AIRE<sup>+</sup> confirma el trabajo previo de nuestro grupo. Además, se había descrito un arresto en el ciclo celular de las mTECs AIRE<sup>+</sup>, el cual se postuló como una de las funciones de AIRE para el incremento de la presentación de TRAs por parte de las DCs a través de la captación de cuerpos apoptóticos de mTECs y la consecuente *cross*-presentación (en caso de MHC-I)<sup>56</sup>. Además, en esta tesis, se han caracterizado los motivos de AIRE requeridos para su función como inductor de apoptosis. Los datos concuerdan con estudios previos<sup>228 238</sup>, donde también se concluyó que es necesario el dominio HSR/CARD y el motivo de importe nuclear (NLS) para que AIRE induzca la apoptosis. En la presente tesis, se ha mapeado de una forma más exhaustiva los diferentes dominios de AIRE en esta función. En el análisis proteómico se observó que algunas proteínas relacionadas con apoptosis estaban incrementadas y otras reducidas, como es el caso de prohibitina o plastina 3, la down-regulación de las mismas induciría o generaría un estado celular de susceptibilidad a apoptosis<sup>239 240</sup>.

Aunque la capacidad de regulación a nivel transcripcional ejercida por AIRE puede ser trasladada y reproducida en otras líneas celulares diferentes a las mTECs, resulta dependiente del tipo celular, del entorno y estado celular, de la disponibilidad y expresión de proteínas colaboradoras que definen los genes susceptibles de ser regulados<sup>107</sup>, podría definirse una impronta regulatoria reproducible y dependiente de la expresión de AIRE basada en un incremento de proteínas relacionadas con el control de plegamiento y traducción de proteínas y una intensa remodelación del citoesqueleto que conllevaría un aumento de la susceptibilidad a apoptosis<sup>110</sup>. Este hecho resultaría consistente con los estudios de secuenciación de RNA de células únicas (*single cell RNA-seq*) realizado sobre poblaciones de células maduras mTEC-AIRE<sup>+</sup> y mTEC-AIRE-KO, de la totalidad de los genes up-regulados por AIRE, solamente un 15% serían enteramente dependientes de su actividad, la expresión del 85% restante se encontraría incrementada de manera indirecta<sup>31</sup>. Los transfectantes estables que expresaban AIRE WT o la proteína truncada que portaba el dominio HSR/CARD y el motivo NLS veían incrementada su apoptosis espontánea y su susceptibilidad a apoptosis inducida

(Figura 7). La región HSR/CARD junto con el motivo NLS son necesarias para la interacción de AIRE con la proteína DAXX<sup>227</sup>. DAXX localiza en los denominados cuerpos nucleares de la Leucemia Promielocítica o PML-NBs. Estas estructuras favorecerían una inducción de vías de señalización de apoptosis desde el núcleo, puesto que contienen una variedad de proteínas relacionadas con la muerte celular programada como p53, CNP, Sp100 o caspasa 2<sup>241</sup>. La apoptosis inducida por AIRE, por tanto, podría estar mediada en concomitancia con DAXX. Los resultados analizados de los ligandos peptídicos asociados a HLA-I de células HEK293-AIRE<sup>+</sup> así como los obtenidos de proteómica cuantitativa comparativa por iTRAQ sugirieron un incremento en la expresión de proteínas relacionadas con la maquinaria de reparación del DNA (TOP2A, TOP1 o DNA-PK) así como proteínas relacionadas con el plegamiento (HSP74, STIP1, calnexina o calreticulina) y estrés oxidativo (peroxiredoxina 4 o proteína 1 up-regulada por hipoxia) (Tablas 1 a 4). La apoptosis inducida por AIRE podría estar igualmente relacionada con una respuesta a daño genotóxico y estrés oxidativo. AIRE induce la expresión de proteínas como HSC70, HSP72 o superóxido dismutasa (SOD)<sup>110</sup> relacionadas con este último evento. La generación de radicales libre derivados de un plegamiento citoplasmático aberrante o de proteínas inducidas por AIRE podría favorecer la traslocación del sensor de estrés oxidativo GAPDH al núcleo activando vías de señalización apoptótica<sup>218</sup>. El motivo HSR/CARD por sí solo no induciría apoptosis como se observó en los ensayos de apoptosis con el transfectante estable HEK293-AIRE-(1-106)-MycHis<sup>+</sup> (Figura 5). La apoptosis inducida tampoco se vio incrementada cuando la proteína truncada de AIRE presentaba el dominio SAND, el dominio PHD1 o ambos dominios PHD, aún portando el dominio HSR/CARD y el motivo NLS. Estos resultados podrían deberse, por un lado, a la necesidad de una correcta oligomerización de AIRE y una localización nuclear para ejercer su función como regulador transcripcional<sup>232 80 242</sup>, por otro, la capacidad de regulación transcripcional podría variar de un transfectante a otro, es decir, las proteínas up-reguladas por AIRE-WT podrían diferir de aquellas up-reguladas por proteínas AIRE sin algunos dominios funcionales generando, en última instancia diferencias en la susceptibilidad a tratamiento con etopósido. Por otro lado, puede ser que la inducción de apoptosis en las células HEK293 sin transfectar active otras rutas de apoptosis que son

suficientemente importantes para que enmascaren las generadas por los diferentes dominios de AIRE. Se ha de remarcar que, en los ensayos de apoptosis, la sobreexpresión de AIRE-(1-106)-Myc/His redujo tanto la apoptosis espontánea como la inducida y, aunque no somos capaces de explicar este efecto, un estudio ha determinado la capacidad inhibitoria de polipéptidos recombinantes que contienen el motivo CARD sobre cascadas de activación de la procaspasa-9 o procaspasa-1<sup>243</sup>. El polipéptido que expresa el transfectante estable HEK293-AIRE-(1-106)-MycHis podría estar ejerciendo esta actividad inhibitoria sobre la vía de señalización de apoptosis reduciendo la susceptibilidad a muerte celular.

Adicionalmente al aumento de apoptosis, los transfectantes de AIRE mostraron una capacidad de migración más elevada. Este dato puede estar asociado al hecho de que las células que expresan AIRE ven modificado el patrón de expresión de proteínas relacionadas con el citoesqueleto, lo que afecta a su capacidad de desplazamiento. El incremento en la movilidad de las células AIRE<sup>+</sup> no había sido descrito previamente. El papel de este efecto en las mTECs es discutible, ya que, a priori, son células con poco movimiento en la médula tímica, pero puede tener una importancia en la expresión extratímica de AIRE<sup>244</sup> o en células tumorales que expresan AIRE<sup>245</sup>.

Respecto a proteínas relacionadas con el citoesqueleto, de nuevo, algunas se ven disminuidas, como keratina 18 y keratina 19, testina o calumenina. La disminución en la expresión de estos marcadores han sido relacionados con una mayor capacidad de invasividad en algunos tipos de cáncer<sup>246 247 248 249</sup>. El incremento en los niveles de MYH10 en los análisis de iTRAQ junto con los resultados obtenidos de los ensayos de *wound healing*, sugirieron que la expresión de AIRE induciría el incremento de los niveles de MYH10 y esta situación favorecería, probablemente junto con otras proteínas no identificadas, la capacidad de migración de las células (**Figura 9**). Se ha comprobado que la desregulación de MYH10 en algunos tipos de cáncer, como glioma, incrementa la capacidad de migración de líneas celulares similares<sup>231</sup>. La expresión extra-tímica del Regulador Autoinmune ha sido asociada en procesos de malignización

celular, así, ha sido detectado en células de cáncer de mama y cáncer oral con conclusiones dispares, por un lado, la expresión de AIRE en cáncer de mama sería indicativo de una mejor prognosis asociada a un incremento en la apoptosis e inhibición de progresión de ciclo celular<sup>250</sup>. Por otro lado, en carcinoma de células escamosas orales la expresión de AIRE se asociaba con una regulación positiva de un set de genes asociados a malignización epidérmica<sup>245</sup>.

El incremento de apoptosis y de movilidad en células AIRE<sup>+</sup> no se correspondió con un aumento de proliferación celular (**Figura 16**). Aquellas células que expresaban AIRE sufrían un incremento de células en fase G2, sugiriendo alguna desregulación en ciclo celular (**Figura 14**). Como se ha comentado previamente, sí se había observado previamente que las células mTEC-AIRE<sup>+</sup> presentan un arresto en la progresión normal del ciclo celular y son susceptibles de sufrir apoptosis, favoreciendo un alto recambio celular en el timo que favorecería el proceso de *cross*-presentación de diversos TRAs a células dendríticas residentes en el timo durante la selección tímica<sup>56</sup>. No obstante no se ha definido el mecanismo por el cual AIRE controla la progresión del ciclo celular.

El inhibidor 1 de la kinasa dependiente de ciclina o más conocido como p21 (CDKN1A) (**Tabla 1**) se vio aumentado en células AIRE<sup>+</sup>, es una proteína relacionada principalmente con la regulación de ciclo celular impidiendo la progresión de fase G2 a fase M<sup>251</sup> pudiendo explicar las diferencias observadas en la distribución poblacional según la fase del ciclo celular de los transfectantes estables respecto a las células control (**Figura 14B**).

Aunque la función e importancia de AIRE desde su descubrimiento se centró en su capacidad de regulación del proceso de selección tímica y adquisición de tolerancia central, su implicación en ciclo celular y división celular quedó relegada. Pacientes con mutaciones en el gen de AIRE sufren a menudo procesos de infertilidad, también estudiado en ratones y resulta esencial para la correcta compleción del ciclo celular. En

estudios con células embrionarias de ratón se comprobó la importancia de papel de AIRE en la regulación y fidelidad del proceso de división celular, así como la implantación de los embriones<sup>52</sup>.

La distribución y localización de AIRE en mTECs solo se ha detectado en el núcleo en forma de *dots* nucleares<sup>79</sup>. La dificultad para el estudio del mismo a nivel celular *in situ* en el tejido tímico o para el aislamiento de mTECs-AIRE<sup>+</sup> debido a la baja proporción de células en comparación con otras poblaciones celulares del tejido, llevó al estudio de la proteína en líneas celulares, transfectando AIRE WT, AIRE truncado en diversos dominios o portando mutaciones comunes en pacientes de APECED. La distribución de AIRE en líneas celulares recordaba a la distribución natural observada en las mTECs, con la diferencia de la presencia también en el citoplasma formando estructuras fibrilares<sup>79</sup><sup>80</sup>. En los diferentes clones de transfectantes estables generados sobre las HEK293 que expresaron de manera estable AIRE WT y AIRE truncado en los diferentes dominios se observó una localización de AIRE habitual en forma de *dots* nucleares distribuidas por el núcleo y en estructuras fibrilares en el citoplasma. La presencia del motivo de importe nuclear (NLS) es necesaria y suficiente para dirigir la proteína al núcleo, su ausencia relega la proteína al citoplasma. Este hecho concuerda con las observaciones anteriores, la modificación de los residuos básicos de Arginina o Lisina en la parte C-terminal del motivo NLS provocaban la traslocación de AIRE al citoplasma<sup>57</sup>. Además, la presencia del dominio HSR/CARD es necesaria para la adquisición de estructuras definidas (*dots* nucleares y/o fibras) como se observó en las inmunofluorescencias de los transfectantes estables que expresaban AIRE-(1-106), AIRE-(107-545), AIRE-(181-545), AIRE-(281-545), AIRE-(344-545) o AIRE-(476-545). En todos ellos la proteína presentaba una distribución uniforme, dentro o fuera del núcleo, dependiendo de si el polipéptido truncado presentaba, además, el motivo NLS. El dominio HSR/CARD concentra, junto con los dominios SAND y PHD, las mutaciones predominantes en AIRE que provocan APECED<sup>44</sup>, la eliminación del dominio HSR/CARD supone la pérdida en la capacidad de activación de transcripción necesaria para la función de la proteína puesto que la generación de homo-multímeros queda impedida<sup>53</sup><sup>252</sup>. Igualmente, solo

aquellos transfectantes que expresaban AIRE con el dominio HSR/CARD y la región que contenía el motivo NLS (al menos los aminoácidos 1-180) fueron capaces de localizar la proteína conformando estructuras fibrilares y/o asociadas al *midbody* durante la división celular y citocinesis; el transfectante que expresaba AIRE-(1-106) no presentó un patrón de tinción definido en estructuras.

Por otro lado, la ausencia de dominios SAND, PDH1 o PHD2 modificaban igualmente la distribución de AIRE en la célula. Cuando los dominios SAND, PHD1 y PHD2 no estaban presentes (AIRE-(1-180)), AIRE se distribuyó principalmente en el núcleo de manera difusa y en conjunto con un menor número de *dots* nucleares siendo éstos últimos de mayor tamaño. La presencia del dominio SAND (AIRE-1-280) exhibió el mismo patrón de localización anterior. La adición a la secuencia del dominio PHD1 (AIRE-(1-343)) restableció la localización celular de la proteína prácticamente al mismo nivel que AIRE WT con menor presencia en el citoplasma. La eliminación de la última región rica en leucinas AIRE-(1-475) provocó una distribución predominantemente citoplasmática en forma de estructuras fibrilares con algunos *dots* nucleares en el núcleo (**Figura 13**).

La modificación de la secuencia de AIRE afecta directamente a la localización subcelular de la proteína<sup>80 242</sup>, no obstante nuestros resultados difieren con algunas observaciones de estudios anteriores, mientras que la proteína AIRE-(1-257) (construcción similar a la mutación más común entre finlandeses R257X) parece observarse únicamente en el citoplasma formando estructuras fibrilares<sup>253 254</sup>, en nuestras manos, los transfectantes AIRE (1-280) mostraron un patrón de tinción principalmente nuclear, esto puede deberse a que nuestra secuencia codifica para una proteína con un dominio SAND intacto, con la consecuente capacidad para permitir la compartimentalización nuclear. Igualmente nuestro transfectante AIRE-(1-180), sin la presencia de un dominio SAND funcional, mostró una localización principalmente nuclear. Respecto a la eliminación de la parte C-terminal de la proteína (AIRE-(1-475)) la traslocación sustancial de AIRE podría explicarse por la pérdida del cuarto motivo LXXLL. Los motivos LXXLL son secuencias aminoacídicas muy conservadas que median

interacciones proteína-proteína entre receptores nucleares y sus co-activadores<sup>255</sup>. Algunas mutaciones en pacientes APECED son localizadas precisamente en este motivo<sup>256</sup> provocando la completa disrupción de la actividad activadora transcripcional de AIRE<sup>257</sup>. Esta situación resulta esperable, puesto que la localización predominantemente citoplasmática mostrada en nuestro transfectante estable impediría la interacción de AIRE con la cromatina y la activación transcripcional de los genes dependientes del mismo. Estos resultados dispares podrían deberse a que la mayoría de estudios en líneas celulares se realizaron a través de transfecciones transitorias de diferentes líneas celulares, no en transfectantes estables. Las características de cada línea celular podrían suponer diferencias en los resultados obtenidos<sup>254 103</sup>. No obstante, podemos establecer que el motivo HSR/CARD es necesario para la generación de estructuras definidas necesarias para la correcta función de AIRE. El motivo de importe nuclear NLS determina y permite su localización nuclear. El motivo SAND por sí solo no favorece la correcta localización de la proteína formando *dots* nucleares a lo largo del núcleo, es necesaria la presencia del motivo PHD1, favoreciendo la interacción con componentes esenciales del mantenimiento y estabilidad del genoma como la DNA-PK<sup>225</sup> y reconociendo y estableciendo contacto con marcas epigenéticas<sup>258</sup>. El motivo PHD2 parece no tener implicación alguna en la localización subcelular de AIRE<sup>66</sup> pero la presencia del motivo rico en prolinas podría favorecer la interacción con proteínas que afectan a la estabilidad de AIRE en el núcleo, siendo imprescindible la presencia del extremo C-terminal que contiene el motivo LXXLL para una correcta localización, sugiriendo que esta parte final de la secuencia permite la interacción con algún regulador de la proteína AIRE a nivel de estabilización de la proteína o la modificación postraduccional de la misma favoreciendo su actividad en el núcleo. Todos estos análisis refuerzan la importancia no sólo de aquellos elementos estructurales dentro de la propia secuencia del Regulador Autoinmune, sino también de la adquisición de una correcta localización subcelular para ejercer la activación transcripcional en las mTECs.

La localización subcelular de AIRE dentro de la célula es dependiente del ciclo celular y la formación de los característicos *dots* nucleares sería posterior a la división celular, desapareciendo poco antes de la entrada en mitosis y durante el proceso de división<sup>103</sup>. Este hecho sugeriría la imposibilidad de una implicación de AIRE en división celular. No obstante, estudios en células embrionarias de ratón identificaron AIRE como regulador de proteínas del huso mitótico<sup>52</sup>. Durante la caracterización de la localización de la proteína en diferentes transfectantes estables mediante microscopía confocal se identificó la presencia de estructuras fibrilares de AIRE en torno al *midbody* en células post-mitóticas, durante la fase final de división celular, la denominada citocinesis. La presencia de AIRE en esta localización no fue fortuita y fue reproducible en todos los transfectantes estables que expresaban AIRE cuando el dominio HSR/CARD y el motivo NLS estaban presentes en la proteína (**Figura 14**). Además, esta localización se detectó en otras líneas celulares transfectadas con AIRE, como HT93, HeLa o SK-Hep-1 (**Figura 14D**), lo que indica que no es únicamente debido a características intrínsecas de la línea HEK293. A lo largo del ciclo celular AIRE dejó de detectarse en momentos previos a la entrada en ciclo celular, durante la profase mitótica. Posteriormente se localizó en agregados distribuidos entre las cromátidas y más tarde asociado a los centriolos, en pro-metafase. Conforme las cromátidas se establecieron en el plano ecuatorial de las células en división, AIRE siguió localizándose en los centriolos y en estructuras fibrilares similares a los microtúbulos no cinetocóricos y microtúbulos astrales durante la metafase y a lo largo de la anafase y telofase celulares, para quedar relegado a las estructuras fibrilares en torno al *midbody* anteriormente descritas. Durante todo el proceso de división celular AIRE no generó los característicos *dots* nucleares (**Figura 18**). En células embrionarias de ratón esta situación es similar. A lo largo del ciclo celular AIRE colocaliza con tubulina- $\alpha$  y tubulina- $\beta$ , componentes esenciales de los microtúbulos que conforman el citoesqueleto del huso acromático, también se distribuye de manera discreta en los husos polares de las células en división. Los dominios HSR/CARD, SAND y el cuarto motivo LXXLL en la secuencia de AIRE resultan esenciales para ejercer su función asociada al huso mitótico<sup>52</sup>. En nuestro modelo celular de HEK293-AIRE<sup>+</sup> la sobreexpresión de AIRE no parece afectar a la proliferación, no obstante, parece generar algún tipo de perturbación en el progreso normal del ciclo



celular ya que se observó un enriquecimiento en el número de células en fase G2-M (Figura 15). Este hecho es identificado en estudios sobre espermatogénesis en ratones, donde la expresión de AIRE es fundamental para generar procesos apoptóticos indispensables para el mantenimiento de la estabilidad de las células de línea germinal<sup>50</sup>, el proceso de expresión promiscua de genes en el tejido testicular<sup>49</sup> facilitaría la inducción de apoptosis debido a respuesta por mal plegamiento de proteínas, manteniéndose aquellas células de línea germinal con menor probabilidad de carga mutacional<sup>50</sup>. *A priori*, la ausencia del motivo LXXLL no parece tener efecto en el progreso del ciclo celular en células embrionarias humanas, el transfectante estable HEK293-AIRE-(1-475)<sup>+</sup> completaba el ciclo celular y localizaba la proteína AIRE en estructuras en torno al *midbody* similares a las que fueron observadas en AIRE WT (Figura 14). No obstante, quedaría por comprobar el impacto global en la distribución poblacional en diferentes fases del ciclo celular y la implicación de la ausencia de los diferentes motivos de AIRE en nuestros transfectantes estables.

La citocinesis y escisión constituyen los procesos finales que completan la división celular. El huso mitótico sufre una intensa remodelación que genera una estructura denominada huso central, localizada en la placa ecuatorial de la célula en división. El huso central es determinante para la progresión de la citocinesis puesto que establece la posición del surco de corte que separará físicamente la membrana celular de las células hijas<sup>96</sup>. La actividad del huso central es iniciada gracias a la actividad de diversas proteínas (el Complejo Centralspindlina, diversas GTPasas y el Complejo Pasajero del Cromosoma) permitiendo la cascada de señalización que favorece la ingresión del surco de corte para la escisión gracias a la reorganización citoesquelética y la activación de proteínas motoras<sup>97</sup>. Una vez que el surco de corte ha ingresado se genera el *midbody*, una estructura centralizada en el puente citoplasmático que sigue uniendo las células hijas, compuesto por microtúbulos empaquetados junto con material amorfo electrodenso. La Centralspindlina queda altamente concentrada en las regiones próximas al *midbody*<sup>98</sup>, mientras que en las zonas más distales se localizan microtúbulos y proteínas del complejo ESCRT que permitirán la escisión<sup>99</sup>. Algunas proteínas como

MYH10 y DNA-PK han sido descritas como moduladoras del proceso de citocinesis<sup>100 101</sup><sup>102</sup>. Aunque AIRE y DNA-PK interactúan<sup>225</sup> no identificamos co-localización de DNA-PK en torno al *midbody* en células post-mitóticas, la presencia de DNA-PK en torno al *midbody* podría ser muy transitoria y su detección fue menor en términos de intensidad de fluorescencia respecto a las células en interfase (**Figura 16**). El incremento en los niveles de MYH10 en transfectantes estables de AIRE nos llevó a postular una posible relación entre ambas proteínas durante este proceso. MYH10 efectivamente localiza en el surco de ingresión durante el proceso de división celular quedando relegada al anillo de escisión tras la citocinesis. MYH10 y AIRE no co-localizaron, AIRE localizó en el exterior de los anillos de MYH10 que flanquean el *midbody* (**Figura 17**). No obstante la importancia de MYH10 en la compleción de la mitosis y también de la meiosis en células de línea germinal masculina<sup>229</sup> refuerza la implicación de AIRE en la regulación de ciclo celular, el incremento en la expresión de MYH10 podría asegurar la correcta división meiótica y la correcta estabilidad genómica en línea germinal.

Probablemente la localización en estructuras en torno al *midbody* pueda deberse a la interacción de AIRE con los microtúbulos y otros componentes esenciales para la citocinesis incluyendo Aurora Kinasa B, proteínas centrosómicas como CEP55, CNTROB, HAUS5 y HAUS8 y proteínas reguladoras de la dinámica de microtúbulos como CLASP1 y CLASP2<sup>52 259</sup>. No obstante, queda por discernir los motivos y dominios de AIRE implicados en esta localización e interacción con el resto de proteínas asociadas a citocinesis. Esta localización relacionada con el ciclo celular no hace más que acrecentar la plétora de actividades que ejerce este factor de transcripción tan inusual más allá de la regulación de la Tolerancia Central.

Una de las proteínas de tipo chaperona que también se vio incrementada en células HEK293-AIRE<sup>+</sup> fue CACYBP/SIP (**Figura 19**), una proteína adaptadora que interviene en múltiples procesos biológicos como modificación postraduccional de proteínas, ubiquitinación, dinámica del citoesqueleto, regulación de expresión génica y diferenciación así como regulación de la proliferación celular y la tumorigénesis<sup>213</sup>.

CACYBP/SIP también se vio incrementada en la proteómica comparativa de células HT93-AIRE<sup>+</sup> <sup>110</sup>, sugiriendo una posible relación funcional entre esta proteína y AIRE. CACYBP/SIP se vio incrementada en clones de transfectantes estables de HEK293-AIRE-MycHis<sup>+</sup>, debido a la principal actividad como proteína mediadora en procesos de ubiquitinación, cabe pensar que este incremento se podría deber a una acción reguladora sobre los niveles de AIRE. CACYBP/SIP podría mediar en la degradación de AIRE, sin embargo, la transfección transitoria de CACYBP/SIP sobre los mismos transfectantes estables HEK293-AIRE-MycHis<sup>+</sup> no logró reducir los niveles de AIRE sino que incluso llegaron a incrementarse. La relación entre ambas proteínas tampoco es directa, experimentos de co-inmunoprecipitación tras una co-transfección de ambas proteínas determinaron que ambas proteínas no interaccionaban. Este hecho sugirió que la posible relación entre ambas proteínas podría deberse a un mecanismo de respuesta biológica de compensación: CACYBP/SIP y AIRE podrían competir por el mismo sustrato, lo que se traduce en un incremento cuando la otra proteína se sobreexpresa. Una de las características más llamativas de CACYBP/SIP es la presencia en su secuencia de un motivo consenso de unión a SIAH1, compuesto por (PxAxVxP)<sup>208</sup>. CACYBP/SIP interacciona a través de dicho motivo con la E3 Ubiquitin Ligasa SIAH1, facilitando la ubiquitinación específica de  $\beta$ -Catenina en respuesta a daño genotóxico<sup>209</sup> <sup>260</sup>. Este motivo aparece hasta dos veces a lo largo de la secuencia de AIRE: el primer motivo de unión, completo, situado entre los 119-125aa (PKALVPP) y el segundo motivo de unión, con la Ala3 sustituida por una Leu, situado entre los 416-422aa (PLLCVGP). La presencia de ambos motivos sugiere una posible relación directa entre proteínas de la familia SIAH y la proteína AIRE. La captura específica de SIAH1 o SIAH2 en lisados celulares tras una co-transfección transitoria con AIRE-MycHis y FLAG-SIAH1 o AIRE-MycHis y FLAG-SIAH2 permitió la recuperación, en ambos casos, de AIRE-MycHis, confirmando la existencia de una interacción entre ambas proteínas. La región que contiene el motivo de unión de SIAH1 con el que CACYBP/SIP interacciona con SIAH1 no presenta ninguna estructura definida ni dominio específico, es una zona de unión o *linker* libre<sup>211</sup>. El primer motivo de unión a SIAH1 de la secuencia de AIRE también se sitúa en una región no definida estructuralmente (119-125aa PKALVPP) entre los dominios HSR/CARD y SAND, fácilmente accesible para otras proteínas. El

segundo motivo de unión, en cambio, recae dentro de la región rica en prolinas (PRR) de la secuencia de AIRE y se incluye dentro del tercer motivo de Leucinas (LXXLL). Aunque ambas son regiones muy conservadas en secuencia y que participan en múltiples aspectos de interacciones proteína-proteína<sup>70 261</sup>, AIRE establece interacción con varias proteínas a través de dichas regiones, como CREBBP, P-TEFb o DNA-PK entre otras<sup>262 257 225 63</sup>. Es factible pensar, por tanto, que ese mismo hecho dificulte la interacción de SIAH1 a través de ese motivo o sea más inaccesible, además, la secuencia del segundo motivo no cumple los requisitos del motivo consenso al presentar una Leucina en vez de una Alanina en la posición número dos. La importancia del primer motivo de unión a SIAH1 en la secuencia de AIRE quedó corroborada al detectarse en la fracción purificada de los lisados celulares aquellas proteínas de AIRE truncadas que portaban la región que contenía el primer motivo de unión (AIRE-(107-545)-MycHis, AIRE-(1-180)-MycHis y AIRE-(1-343)-MycHis) cuando se inmunoprecipitaron específicamente SIAH1 y SIAH2 (**Figura 25**).

El motivo consenso de unión a SIAH1 ha sido ampliamente estudiado, más de la mitad de las proteínas que se unen a SIAH1 lo contienen, resultando esencial para su interacción y funcionalidad<sup>208 209 263 264 265</sup>. Estudios de afinidad con diversas secuencias han demostrado una importancia clave en las posiciones V3 y P4 del motivo: las cadenas laterales de ambos aminoácidos apuntan a un surco hidrofóbico formado por los residuos F165, V179, V176, A167, W178 y M180 de SIAH1<sup>203</sup>. No obstante, varias proteínas que establecen interacción con dicha E3 ubiquitín ligasa difieren en los aminoácidos de dichas posiciones<sup>208 207</sup>. También se ha recalcado la implicación de aminoácidos adyacentes al motivo, como por ejemplo la Lisina previa a la posición P1 (o aminoácidos de la misma naturaleza, como Histidina o Arginina), que interacciona favorablemente con el residuo D162 de la proteína SIAH1. El residuo conservado P1 formaría un clúster hidrofóbico con las cadenas laterales de L158 y I163 para mantener la Lisina precedente en una conformación favorable para la interacción con el residuo D162 de SIAH1<sup>203</sup>.

Respecto a la funcionalidad de la interacción, AIRE se veía eficientemente reducido en transfecciones transitorias con SIAH1 y parece que en alguna medida con SIAH2, aunque era más claro con SIAH1 que con SIAH2. La expresión transitoria de AIRE-MyHis y FLAG-SIAH1 o AIRE-MyHis y FLAG-SIAH2 provocaba una disminución en los niveles totales de AIRE comparado con la misma condición de transfección en presencia de inhibidor de proteasoma y con la transfección de AIRE-MyHis y el Dominio de Unión a Sustrato (FLAG-SBD) que no porta el dominio catalítico de SIAH1. Este descenso era más acusado con SIAH1 que con SIAH2. Con SIAH2 los resultados no fueron suficientemente consistentes ni estadísticamente significativos. Este resultado se repetía en el caso de transfecciones transitorias sobre el clon transfectante estable HEK293-AIRE<sup>+</sup>. SIAH1 produjo una ubiquitinación específica que dependía de la presencia del motivo de unión a SIAH1 en la secuencia de AIRE, solamente se detectaba un incremento en AIRE poliubiquitinado en los transfectantes estables AIRE-MyHis y AIRE-(1-343)-MyHis que sí portaban el motivo consenso en su secuencia, mientras que el transfectante estable AIRE-(181-545)-MyHis no lo mostraba (**Figura 29**). Este resultado sugiere fuertemente que la interacción de AIRE con SIAH1 y, en menor medida, con SIAH2 lleva a su degradación vía proteasoma. AIRE ha sido relacionado con eventos de modificación post-traduccionales como la ubiquitinación o con el sistema Ubiquitín-Proteasoma debido a la presencia de dos dominios PHD en su secuencia<sup>67</sup>. De manera general, los dominios PHD se encuentran conservados entre proteínas cuya actividad está asociada a la regulación de la transcripción mediada por interacción con cromatina<sup>68</sup>. Los dominios PHD de AIRE presentan la estructura canónica de dominio PHD<sup>64</sup> y ésta, resulta claramente distinta al plegamiento de un dedo RING típicamente presente en enzimas E3 Ubiquitín Ligasa<sup>69</sup>. Algunos autores afirman, sin embargo, que la proteína es capaz de exhibir una actividad E3 Ubiquitín Ligasa a través de su dominio PHD1<sup>266</sup>, ya que se han descrito varias proteínas con dominio PHD capaces de realizar procesos de ubiquitinación específica<sup>267 268</sup>. Otros, en cambio, han confirmado que AIRE-PHD1 no presenta una actividad E3 Ubiquitín Ligasa intrínseca, ni establece interacción con su enzima E2 de Conjugación de Ubiquitina putativa para completar el proceso de ubiquitinación específica<sup>64</sup>.

Inequívocamente, AIRE es modificado post-traduccionalmente influyendo en su capacidad de regulación de transcripción<sup>225 227 269</sup>. De manera general, AIRE puede ser fosforilado favoreciendo su dimerización u homomultimerización, estado esencial para ejercer su actividad como regulador de transcripción<sup>232 225</sup>. Esta situación permite la ubiquitinación específica de AIRE al unirse a la E3 Ubiquitín Ligasa Proteína 3 F-Box (FBOX3) generando el complejo SCF<sup>FBOX3</sup> (SKP1-CUL1-F-box) que incrementa la interacción entre AIRE y P-TEFb, potenciando la actividad transcripcional<sup>269</sup>. La regulación transcripcional ejercida por AIRE favorece la activación de enzimas reguladoras de cromatina como metiltransferasas o deacetilasas de histonas, asociadas a complejos proteicos como el compuesto por DAXX, que interacciona con AIRE, regulando negativamente su actividad<sup>227</sup>. En esta tesis no se ha abordado, pero sería de gran interés, la caracterización de la influencia de modificaciones postraduccionales en la interacción de AIRE con las proteínas SIAH y su efecto en la fisiología celular.

Todo ello sugiere un escenario en el que la expresión y actividad de AIRE ha de ser regulada cuidadosamente. La homomultimerización de AIRE para ejercer su actividad<sup>54</sup> puede hacerlo susceptible a sufrir un “control de calidad” mediada por PML-NBs. Esta susceptibilidad presenta una relación inversa al grado de interacción de AIRE con la cromatina. Afectaciones en el dominio CARD impedirían la correcta estabilización de los homomultímeros, favoreciendo la generación de agregados y alteraciones en el dominio PHD1 inhabilitarían a AIRE para interactuar con la cromatina, derivando a su traslocación hacia los PMLs<sup>270</sup>. Los PML-NBs están implicados en una amplia variedad de procesos celulares: actividad supresora tumoral, defensa viral, regulación de la transcripción, entre otros. Estas estructuras parecen localizarse adyacentes a regiones transcripcionalmente activas y puede presuponerse que puedan modular el proceso de transcripción sobre *loci* específicos<sup>271</sup>. Se ha demostrado que los PML-NBs también se localizan en zonas de daño del DNA asociados a la maquinaria celular de reparación de DNA y DSBs y también parecen estar relacionados con procesos de apoptosis debido a su relación con p53<sup>272</sup>.

La S-nitrosilación de GAPDH en respuesta a estrés oxidativo induce la interacción con SIAH1 y su estabilización. La traslocación posterior al núcleo favorecería la ubiquitinación y degradación de proteínas nucleares<sup>218</sup>. Los niveles de SIAH1 también se ven incrementados en respuesta a daño genotóxico y apoptosis<sup>273 202</sup>. Con ello, podríamos suponer que la expresión de SIAH1 podría ser inducida en respuesta a daño genotóxico derivado de la actividad de TOP1 y TOP2a al generar DSBs sobre zonas del genoma reguladas por AIRE al mismo tiempo que el estrés oxidativo derivado de un plegamiento aberrante de proteínas favorecería la activación de GAPDH. Esta situación favorecería la actividad de SIAH1 permitiendo, por un lado la ubiquitinación de AIRE pudiendo inducir la traslocación de AIRE a PML-NBs donde sería regulado por otras proteínas presentes. La implicación de AIRE en apoptosis también podría verse relacionada con SIAH1, esta E3 Ubiquitín ligasa también participa en la inducción de apoptosis vía p53/TP53 y podría explicar el incremento de muerte celular tanto en líneas celulares como las mTECs.

Mutaciones en el gen *AIRE* provocan APS-1 APECED, una enfermedad rara autoinmune monogénica de carácter autosómico recesivo<sup>41</sup>. Este síndrome provoca respuestas autoinmunes órgano-específicas con el consecuente deterioro y destrucción de multitud de tejidos, principalmente, tejidos de tipo endocrino. Las manifestaciones principales de esta enfermedad se dividen en tres componentes: candidiasis crónica, hipoparatiroidismo autoinmune y enfermedad de Addison, la identificación de al menos dos de ellos son necesarios para el diagnóstico<sup>40 42 43</sup>. Los factores etiológicos que provocan la enfermedad se producen por la pérdida de la tolerancia central debido a fallos en el proceso de selección tímica. La pérdida en la correcta expresión de AIRE ya sea por la producción de una proteína aberrante o una proteína con dominios no funcionales, convergen en el impedimento de la regulación en la expresión de TRAs y, por tanto, de la adquisición de una correcta antigenicidad en la médula del timo, provocando en última instancia la liberación a periferia de timocitos potencialmente autorreactivos y afectaciones a las poblaciones de linfocitos Treg para el control de la

tolerancia en periferia<sup>43 44</sup>. Todo ello provoca la generación de una plétora de autoanticuerpos dirigidos contra autoantígenos comunes entre pacientes de este síndrome, así como infiltrados linfocitarios en diversos órganos y tejidos. Es una enfermedad rara pero especialmente prevalente en algunas poblaciones, como población finlandesa (1/25000), población sarda (1/14000) y judíos iraníes (1/9000)<sup>40</sup>.

La mayoría de las mutaciones identificadas en el gen de AIRE se concentran en los dominios funcionales de la proteína, entre ellas, las predominantes aparecen sobre la región codificante para el dominio HSR/CARD y los dominios PHD<sup>44 73</sup>. Algunas de ellas consisten en mutaciones *nonsense* o de cambio de marco de lectura, generando polipéptidos truncados; otras son mutaciones de cambio de sentido en un solo aminoácido, generando la proteína completa pero no funcional<sup>73</sup>. Hasta la fecha se han identificado muy pocas mutaciones que operan de manera dominante negativa, entre ellas: algunas que afectan al dominio SAND (p.G228W o p.R247C) y otras que recaen sobre los dominios PHD1 (p.C311Y) y PHD2 (p.C446G)<sup>73 74</sup>. En estudios *in vitro* se comprobó el efecto dominante negativo sobre genes regulados por AIRE tras la co-transfección de AIRE WT y AIRE con diversas mutaciones sobre el dominio PHD1, incluida p.C311Y sobre líneas celulares. Otras mutaciones relativas al dominio HSR/CARD y la mutación p.C311\* mostraron un efecto recesivo. La mutación p.R471C referente al dominio PHD2 no ofrecía variaciones a nivel de regulación sobre genes dependientes comparado con AIRE WT<sup>73 74 75</sup>. En estudios *in vivo* en modelos animales de ratón se confirmaron estos efectos y características: ratones con mutaciones monoalélicas (p.C311Y) y (p.C446G) presentaban un impedimento para la expresión de TRAs dependientes de AIRE en mTECs y una reducción en el número de células Treg en el timo, así como respuestas autoinmunes de carácter leve<sup>76</sup>.

Las mutaciones dominantes negativas muestran una penetrancia incompleta y los pacientes portadores presentan cuadros clínicos alejados de los comunes establecidos para APS-1 de carácter recesivo<sup>77</sup> esto es debido a que, de manera general, las mutaciones dominantes-negativas generan proteínas disfuncionales, actuando como



inhibidores competitivos de la función de la proteína WT, la formación de hetero-oligómeros (MT/WT) u homo-oligómeros (MT/MT)<sup>54</sup> provocaría, por una parte, la incapacidad de AIRE para formar oligómeros funcionales y una localización celular óptima para ejercer su actividad como factor de transcripción, por otra, la inhibición de la interacción con otras proteínas que colaboran con AIRE regulando su función<sup>73</sup>. Algunas mutaciones sin sentido p.R257\* o p.C311\* generarían proteínas truncadas que establecerían hetero-oligómeros con proteína AIRE WT, manteniendo el *core* funcional (dímeros de AIRE WT activo), ello explicaría el carácter recesivo de éstas mutaciones<sup>74</sup>. Afectaciones a los dominios PHD1 o PHD2 en las mutaciones p.C311Y y p.C446G generarían homo-oligómeros o hetero-oligómeros incapaces de ejercer su actividad sobre la cromatina pese a la correcta localización subcelular<sup>76</sup>.

El descubrimiento de una nueva vía regulatoria sobre AIRE gracias a la actividad de SIAH1 podría suponer la identificación de una nueva diana terapéutica. El potenciamiento de la actividad de SIAH1 o el incremento en su expresión, más que la inhibición de su actividad con compuestos químicos o péptidos competidores podría suponer la degradación de las proteínas aberrantes de estas mutaciones dominantes negativas (p.C311Y y p.C446G) permitiría regular los niveles de las mismas, ya sea por control desde los PML-NBs o al verse incrementada su capacidad de expresión respecto a la proteína WT permitiendo en última instancia la generación de homo-oligómeros (WT/WT) funcionales pudiendo rescatar o, en todo caso, incrementar la actividad de AIRE WT en mTECs de pacientes heterocigóticos para éstas mutaciones dominantes negativas. En otro tipo de mutaciones recesivas, donde la funcionalidad de AIRE se ve disminuida, pero no se produce una proteína aberrante, el bloqueo de la degradación vía SIAH1 (con péptidos bloqueantes de la interacción con AIRE u otra terapia similar) podría alargar la vida de la proteína AIRE defectiva, lo que podría facilitar que ejerciese su función y paliar los efectos de la mutación.







# CONCLUSIONES

1. Los proteomas de células que expresan AIRE muestran, en general, un incremento de proteínas proapoptóticas y de diversas chaperonas y una disminución de diversas proteínas relacionadas con el citoesqueleto. Este patrón proteico se ve reflejado en un incremento de la apoptosis y en una mayor movilidad celular.
2. El análisis de los peptidomas asociados a HLA de clase I se correlaciona con los datos obtenidos en el análisis de los proteomas. La expresión de AIRE no modifica de forma apreciable la generación de ligandos asociados a los diferentes alotipos de HLA de clase I.
3. Se ha caracterizado una nueva localización de AIRE cerca del *midbody* en células post-mitóticas. Para la localización de AIRE en esta nueva localización, hace falta al menos la región 1-180, que es la región necesaria para la oligomerización de la proteína.
4. Se ha identificado por primera vez la presencia de AIRE en el huso acromático durante las diferentes fases de la mitosis en células humanas. Hasta el momento, solo se había descrito en células embrionarias de ratón.
5. AIRE interacciona con las E3-Ubiquitín ligasas SIAH1 y SIAH2 a través del primer motivo de interacción presente en su secuencia. Dicha interacción permite la ubiquitinación de AIRE y su degradación por SIAH1, mientras que con SIAH2 este efecto es menor.









# BIBLIOGRAFÍA

1. Miller JF. ROLE OF THE THYMUS IN IMMUNITY. *Br Med J.* 1963 Aug 24;2(5355):459-64. doi: 10.1136/bmj.2.5355.459. PMID: 14043707; PMCID: PMC1874012.
2. Romagnani, S. Immunological tolerance and autoimmunity. *Intern Emerg Med* **1**, 10 (2006).
3. Thapa, P. & Farber, D. L. The Role of the Thymus in the Immune Response. *Thorac. Surg. Clin.* **29**, 123–131 (2019).
4. Turvey, S. E. & Broide, D. H. Innate immunity. *J. Allergy Clin. Immunol.* **125**, S24–S32 (2010).
5. Bonilla, F. A. & Oettgen, H. C. Adaptive immunity. *J. Allergy Clin. Immunol.* **125**, S33–S40 (2010).
6. *Pathogen-Host Interactions: Antigenic Variation v. Somatic Adaptations*. vol. 57 (Springer International Publishing, 2015).
7. Schatz, D. G. & Ji, Y. Recombination centres and the orchestration of V(D)J recombination. *Nat. Rev. Immunol.* **11**, 251–263 (2011).
8. Nemazee, D. Receptor editing in lymphocyte development and central tolerance. *Nat. Rev. Immunol.* **6**, 728–740 (2006).
9. Gray, D. H. *et al.* Controlling the thymic microenvironment. *Curr. Opin. Immunol.* **17**, 137–143 (2005).
10. Gameiro, J., Nagib, P. & Verinaud, L. The thymus microenvironment in regulating thymocyte differentiation. *Cell Adhes. Migr.* **4**, 382–390 (2010).
11. Yamano, T., Steinert, M. & Klein, L. Thymic B Cells and Central T Cell Tolerance. *Front. Immunol.* **6**, (2015).
12. Starr, T. K., Jameson, S. C. & Hogquist, K. A. Positive and Negative Selection of T Cells. *Annu. Rev. Immunol.* **21**, 139–176 (2003).

13. Rodewald, H.-R. Thymus Organogenesis. *Annu. Rev. Immunol.* **26**, 355–388 (2008).
14. Bennett, A. R. *et al.* Identification and Characterization of Thymic Epithelial Progenitor Cells. *Immunity* **16**, 803–814 (2002).
15. Rossi, S. W., Jenkinson, W. E., Anderson, G. & Jenkinson, E. J. Clonal analysis reveals a common progenitor for thymic cortical and medullary epithelium. *Nature* **441**, 988–991 (2006).
16. Nowell, C. S. *et al.* Foxn1 Regulates Lineage Progression in Cortical and Medullary Thymic Epithelial Cells But Is Dispensable for Medullary Sublineage Divergence. *PLoS Genet.* **7**, e1002348 (2011).
17. Klein, L. & Kyewski, B. Self-antigen presentation by thymic stromal cells: a subtle division of labor. *Curr. Opin. Immunol.* **12**, 179–186 (2000).
18. Takahama, Y., Ohigashi, I., Baik, S. & Anderson, G. Generation of diversity in thymic epithelial cells. *Nat. Rev. Immunol.* **17**, 295–305 (2017).
19. Alawam, A. S., Anderson, G. & Lucas, B. Generation and Regeneration of Thymic Epithelial Cells. *Front. Immunol.* **11**, 858 (2020).
20. Ribeiro, A. R., Rodrigues, P. M., Meireles, C., Di Santo, J. P. & Alves, N. L. Thymocyte Selection Regulates the Homeostasis of IL-7–Expressing Thymic Cortical Epithelial Cells In Vivo. *J. Immunol.* **191**, 1200–1209 (2013).
21. Tomaru, U. *et al.* Exclusive expression of proteasome subunit  $\beta 5t$  in the human thymic cortex. *Blood* **113**, 5186–5191 (2009).
22. Guerder, S., Hassel, C. & Carrier, A. Thymus-specific serine protease, a protease that shapes the CD4 T cell repertoire. *Immunogenetics* **71**, 223–232 (2019).
23. Akiyama, T. *et al.* The Tumor Necrosis Factor Family Receptors RANK and CD40 Cooperatively Establish the Thymic Medullary Microenvironment and Self-Tolerance. *Immunity* **29**, 423–437 (2008).

24. Surh, C. D. *et al.* Two subsets of epithelial cells in the thymic medulla. *J. Exp. Med.* **176**, 495–505 (1992).
25. Kadouri, N., Nevo, S., Goldfarb, Y. & Abramson, J. Thymic epithelial cell heterogeneity: TEC by TEC. *Nat. Rev. Immunol.* **20**, 239–253 (2020).
26. Takaba, H. *et al.* *Fezf2* Orchestrates a Thymic Program of Self-Antigen Expression for Immune Tolerance. *Cell* **163**, 975–987 (2015).
27. Nishikawa, Y. *et al.* Temporal Lineage Tracing of Aire-Expressing Cells Reveals a Requirement for Aire in Their Maturation Program. *J. Immunol.* **192**, 2585–2592 (2014).
28. Laan, M. *et al.* Post-Aire Medullary Thymic Epithelial Cells and Hassall's Corpuscles as Inducers of Tonic Pro-Inflammatory Microenvironment. *Front. Immunol.* **12**, 635569 (2021).
29. Watanabe, N. *et al.* Hassall's corpuscles instruct dendritic cells to induce CD4+CD25+ regulatory T cells in human thymus. *Nature* **436**, 1181–1185 (2005).
30. Derbinski, J., Schulte, A., Kyewski, B. & Klein, L. Promiscuous gene expression in medullary thymic epithelial cells mirrors the peripheral self. *Nat. Immunol.* **2**, 1032–1039 (2001).
31. Sansom, S. N. *et al.* Population and single-cell genomics reveal the *Aire* dependency, relief from Polycomb silencing, and distribution of self-antigen expression in thymic epithelia. *Genome Res.* **24**, 1918–1931 (2014).
32. Kyewski, B. & Klein, L. A CENTRAL ROLE FOR CENTRAL TOLERANCE. *Annu. Rev. Immunol.* **24**, 571–606 (2006).
33. Bunting, M. D., Comerford, I. & McColl, S. R. Finding their niche: chemokines directing cell migration in the thymus. *Immunol. Cell Biol.* **89**, 185–196 (2011).
34. Klein, L., Kyewski, B., Allen, P. M. & Hogquist, K. A. Positive and negative selection of the T cell repertoire: what thymocytes see (and don't see). *Nat. Rev. Immunol.* **14**, 377–391 (2014).
35. Anderson, M. S. *et al.* Projection of an Immunological Self Shadow Within the Thymus by the Aire Protein. *Science* **298**, 1395–1401 (2002).

36. Liston, A., Lesage, S., Wilson, J., Peltonen, L. & Goodnow, C. C. Aire regulates negative selection of organ-specific T cells. *Nat. Immunol.* **4**, 350–354 (2003).
37. Aschenbrenner, K. *et al.* Selection of Foxp3+ regulatory T cells specific for self antigen expressed and presented by Aire+ medullary thymic epithelial cells. *Nat. Immunol.* **8**, 351–358 (2007).
38. Herbin, O. *et al.* Medullary thymic epithelial cells and CD8 $\alpha$  + dendritic cells coordinately regulate central tolerance but CD8 $\alpha$  + cells are dispensable for thymic regulatory T cell production. *J. Autoimmun.* **75**, 141–149 (2016).
39. Weinreich, M. A. & Hogquist, K. A. Thymic Emigration: When and How T Cells Leave Home. *J. Immunol.* **181**, 2265–2270 (2008).
40. Peterson, P., Pitkänen, J., Sillanpää, N. & Krohn, K. Autoimmune polyendocrinopathy candidiasis ectodermal dystrophy (APECED): a model disease to study molecular aspects of endocrine autoimmunity. *Clin. Exp. Immunol.* **135**, 348–357 (2004).
41. Aaltonen, J. *et al.* An autoimmune disease, APECED, caused by mutations in a novel gene featuring two PHD-type zinc-finger domains. *Nat. Genet.* **17**, 399–403 (1997).
42. Capalbo, D. *et al.* Autoimmune Polyendocrinopathy Candidiasis Ectodermal Dystrophy: Insights into Genotype-Phenotype Correlation. *Int. J. Endocrinol.* **2012**, 1–9 (2012).
43. Ferré, E. M. N., Schmitt, M. M. & Lionakis, M. S. Autoimmune Polyendocrinopathy-Candidiasis-Ectodermal Dystrophy. *Front. Pediatr.* **9**, 723532 (2021).
44. De Martino, L. *et al.* APECED: A Paradigm of Complex Interactions between Genetic Background and Susceptibility Factors. *Front. Immunol.* **4**, (2013).
45. Venanzi, E. S., Melamed, R., Mathis, D. & Benoist, C. The variable immunological self: Genetic variation and nongenetic noise in Aire-regulated transcription. *Proc. Natl. Acad. Sci.* **105**, 15860–15865 (2008).

46. Hikosaka, Y. *et al.* The Cytokine RANKL Produced by Positively Selected Thymocytes Fosters Medullary Thymic Epithelial Cells that Express Autoimmune Regulator. *Immunity* **29**, 438–450 (2008).
47. Irla, M. *et al.* Autoantigen-Specific Interactions with CD4<sup>+</sup> Thymocytes Control Mature Medullary Thymic Epithelial Cell Cellularity. *Immunity* **29**, 451–463 (2008).
48. Gardner, J. M. *et al.* Deletional Tolerance Mediated by Extrathymic Aire-Expressing Cells. *Science* **321**, 843–847 (2008).
49. Radhakrishnan, K. *et al.* Autoimmune Regulator (AIRE) Is Expressed in Spermatogenic Cells, and It Altered the Expression of Several Nucleic-Acid-Binding and Cytoskeletal Proteins in Germ Cell 1 Spermatogonial (GC1-spg) Cells. *Mol. Cell. Proteomics* **15**, 2686–2698 (2016).
50. Schaller, C. E. *et al.* Expression of Aire and the Early Wave of Apoptosis in Spermatogenesis. *J. Immunol.* **180**, 1338–1343 (2008).
51. Nishikawa, Y. *et al.* Biphasic Aire expression in early embryos and in medullary thymic epithelial cells before end-stage terminal differentiation. *J. Exp. Med.* **207**, 963–971 (2010).
52. Gu, B., Lambert, J.-P., Cockburn, K., Gingras, A.-C. & Rossant, J. AIRE is a critical spindle-associated protein in embryonic stem cells. *eLife* **6**, e28131 (2017).
53. Ferguson, B. J. *et al.* AIRE's CARD Revealed, a New Structure for Central Tolerance Provokes Transcriptional Plasticity. *J. Biol. Chem.* **283**, 1723–1731 (2008).
54. Kumar, P. G. *et al.* The Autoimmune Regulator (AIRE) Is a DNA-binding Protein. *J. Biol. Chem.* **276**, 41357–41364 (2001).
55. Halonen, M. *et al.* APECED-causing mutations in AIRE reveal the functional domains of the protein. *Hum. Mutat.* **23**, 245–257 (2004).
56. Gray, D., Abramson, J., Benoist, C. & Mathis, D. Proliferative arrest and rapid turnover of thymic epithelial cells expressing Aire. *J. Exp. Med.* **204**, 2521–2528 (2007).

57. Ilmarinen, T. *et al.* The monopartite nuclear localization signal of autoimmune regulator mediates its nuclear import and interaction with multiple importin alpha molecules. *FEBS J.* **273**, 315–324 (2006).
58. Saltis, M. *et al.* Evolutionarily conserved and divergent regions of the Autoimmune Regulator (Aire) gene: a comparative analysis. *Immunogenetics* **60**, 105–114 (2008).
59. Isaac, A., Wilcox, K. W. & Taylor, J. L. SP100B, a repressor of gene expression preferentially binds to DNA with unmethylated CpGs. *J. Cell. Biochem.* **98**, 1106–1122 (2006).
60. Burnett, E., Christensen, J. & Tattersall, P. A consensus DNA recognition motif for two KDWK transcription factors identifies flexible-length, CpG-methylation sensitive cognate binding sites in the majority of human promoters. *J. Mol. Biol.* **314**, 1029–1039 (2001).
61. Mathis, D. & Benoist, C. Aire. *Annu. Rev. Immunol.* **27**, 287–312 (2009).
62. Saare, M., Rebane, A., Rajashekar, B., Vilo, J. & Peterson, P. Autoimmune regulator is acetylated by transcription coactivator CBP/p300. *Exp. Cell Res.* **318**, 1767–1778 (2012).
63. Abramson, J., Giraud, M., Benoist, C. & Mathis, D. Aire's Partners in the Molecular Control of Immunological Tolerance. *Cell* **140**, 123–135 (2010).
64. Bottomley, M. J. *et al.* NMR Structure of the First PHD Finger of Autoimmune Regulator Protein (AIRE1). *J. Biol. Chem.* **280**, 11505–11512 (2005).
65. Yang, S., Bansal, K., Lopes, J., Benoist, C. & Mathis, D. Aire's plant homeodomain(PHD)-2 is critical for induction of immunological tolerance. *Proc. Natl. Acad. Sci.* **110**, 1833–1838 (2013).
66. Yang, S., Bansal, K., Lopes, J., Benoist, C. & Mathis, D. Aire's plant homeodomain(PHD)-2 is critical for induction of immunological tolerance. *Proc. Natl. Acad. Sci.* **110**, 1833–1838 (2013).
67. Musco, G. & Peterson, P. PHD finger of autoimmune regulator: An epigenetic link between the histone modifications and tissue-specific antigen expression in thymus. *Epigenetics* **3**, 310–314 (2008).

68. Musselman, C. A. & Kutateladze, T. G. Handpicking epigenetic marks with PHD fingers. *Nucleic Acids Res.* **39**, 9061–9071 (2011).
69. Chasapis, C. & Spyroulias, G. RING Finger E3 Ubiquitin Ligases: Structure and Drug Discovery. *Curr. Pharm. Des.* **15**, 3716–3731 (2009).
70. Plevin, M. J., Mills, M. M. & Ikura, M. The LxxLL motif: a multifunctional binding sequence in transcriptional regulation. *Trends Biochem. Sci.* **30**, 66–69 (2005).
71. Meloni, A., Incani, F., Corda, D., Cao, A. & Rosatelli, M. C. Role of PHD fingers and COOH-terminal 30 amino acids in AIRE transactivation activity. *Mol. Immunol.* **45**, 805–809 (2008).
72. Bruserud, Ø., Oftedal, B. E., Wolff, A. B. & Husebye, E. S. AIRE-mutations and autoimmune disease. *Curr. Opin. Immunol.* **43**, 8–15 (2016).
73. Oftedal, B. E. *et al.* Dominant Mutations in the Autoimmune Regulator AIRE Are Associated with Common Organ-Specific Autoimmune Diseases. *Immunity* **42**, 1185–1196 (2015).
74. Abbott, J. K. *et al.* Dominant-negative loss of function arises from a second, more frequent variant within the SAND domain of autoimmune regulator ( AIRE ). *J. Autoimmun.* **88**, 114–120 (2018).
75. Björnses, P. *et al.* Mutations in the AIRE Gene: Effects on Subcellular Location and Transactivation Function of the Autoimmune Polyendocrinopathy-Candidiasis–Ectodermal Dystrophy Protein. *Am. J. Hum. Genet.* **66**, 378–392 (2000).
76. Goldfarb, Y. *et al.* Mechanistic dissection of dominant AIRE mutations in mouse models reveals AIRE autoregulation. *J. Exp. Med.* **218**, e20201076 (2021).
77. The Norwegian Addison Registry Study Group *et al.* GWAS for autoimmune Addison’s disease identifies multiple risk loci and highlights AIRE in disease susceptibility. *Nat. Commun.* **12**, 959 (2021).
78. Björnses, P. Localization of the APECED protein in distinct nuclear structures. *Hum. Mol. Genet.* **8**, 259–266 (1999).

79. Heino, M. *et al.* Autoimmune Regulator Is Expressed in the Cells Regulating Immune Tolerance in Thymus Medulla. *Biochem. Biophys. Res. Commun.* **257**, 821–825 (1999).
80. Pitkänen, J., Vähämurto, P., Krohn, K. & Peterson, P. Subcellular Localization of the Autoimmune Regulator Protein. *J. Biol. Chem.* **276**, 19597–19602 (2001).
81. Rinderle, C. AIRE encodes a nuclear protein co-localizing with cytoskeletal filaments: altered sub-cellular distribution of mutants lacking the PHD zinc fingers. *Hum. Mol. Genet.* **8**, 277–290 (1999).
82. Gaetani, M. *et al.* AIRE-PHD fingers are structural hubs to maintain the integrity of chromatin-associated interactome. *Nucleic Acids Res.* **40**, 11756–11768 (2012).
83. Bansal, K., Yoshida, H., Benoist, C. & Mathis, D. The transcriptional regulator Aire binds to and activates super-enhancers. *Nat. Immunol.* **18**, 263–273 (2017).
84. Hnisz, D. *et al.* Super-Enhancers in the Control of Cell Identity and Disease. *Cell* **155**, 934–947 (2013).
85. Bansal, K. *et al.* Aire regulates chromatin looping by evicting CTCF from domain boundaries and favoring accumulation of cohesin on superenhancers. *Proc. Natl. Acad. Sci.* **118**, e2110991118 (2021).
86. Oven, I. *et al.* AIRE Recruits P-TEFb for Transcriptional Elongation of Target Genes in Medullary Thymic Epithelial Cells. *Mol. Cell. Biol.* **27**, 8815–8823 (2007).
87. Giraud, M. *et al.* Aire unleashes stalled RNA polymerase to induce ectopic gene expression in thymic epithelial cells. *Proc. Natl. Acad. Sci.* **109**, 535–540 (2012).
88. Meredith, M., Zemmour, D., Mathis, D. & Benoist, C. Aire controls gene expression in the thymic epithelium with ordered stochasticity. *Nat. Immunol.* **16**, 942–949 (2015).
89. Murray A. Cell cycle checkpoints. *Curr Opin Cell Biol.* 1994.
90. Barnum, K. J. & O'Connell, M. J. Cell Cycle Regulation by Checkpoints. in *Cell Cycle Control* (eds. Noguchi, E. & Gadaleta, M. C.) vol. 1170 29–40 (Springer New York, 2014).



91. McIntosh, J. R. Mitosis. *Cold Spring Harb. Perspect. Biol.* **8**, a023218 (2016).
92. Bertoli, C., Skotheim, J. M. & de Bruin, R. A. M. Control of cell cycle transcription during G1 and S phases. *Nat. Rev. Mol. Cell Biol.* **14**, 518–528 (2013).
93. van den Heuvel, S. Cell-cycle regulation. *WormBook* (2005) doi:10.1895/wormbook.1.28.1.
94. Schafer KA. The cell cycle: a review. *Vet Pathol.* 1998.
95. Blackford AN, Jackson SP. ATM, ATR, and DNA-PK: The Trinity at the Heart of the DNA Damage Response. *Mol Cell.* 2017.
96. White, E. A. & Glotzer, M. Centralspindlin: At the heart of cytokinesis. *Cytoskeleton* **69**, 882–892 (2012).
97. Pollard, T. D. & O’Shaughnessy, B. Molecular Mechanism of Cytokinesis. *Annu. Rev. Biochem.* **88**, 661–689 (2019).
98. Hu, C.-K., Coughlin, M. & Mitchison, T. J. Midbody assembly and its regulation during cytokinesis. *Mol. Biol. Cell* **23**, 1024–1034 (2012).
99. Peterman, E. & Prekeris, R. The postmitotic midbody: Regulating polarity, stemness, and proliferation. *J. Cell Biol.* **218**, 3903–3911 (2019).
100. Roy, A. *et al.* Activity of nonmuscle myosin II isoforms determines localization at the cleavage furrow of megakaryocytes. *Blood* **128**, 3137–3145 (2016).
101. Huang, B. *et al.* DNA-PKcs Associates With PLK1 and Is Involved in Proper Chromosome Segregation and Cytokinesis: DNA-PKcs Associates With PLK1 in Mitosis. *J. Cell. Biochem.* **115**, 1077–1088 (2014).
102. Jones, M. C., Zha, J. & Humphries, M. J. Connections between the cell cycle, cell adhesion and the cytoskeleton. *Philos. Trans. R. Soc. B Biol. Sci.* **374**, 20180227 (2019).
103. Akiyoshi, H. *et al.* Subcellular Expression of Autoimmune Regulator Is Organized in a Spatiotemporal Manner. *J. Biol. Chem.* **279**, 33984–33991 (2004).

104. Laan, M. *et al.* Autoimmune Regulator Deficiency Results in Decreased Expression of CCR4 and CCR7 Ligands and in Delayed Migration of CD4<sup>+</sup> Thymocytes. *J. Immunol.* **183**, 7682–7691 (2009).
105. Metzger, T. C. & Anderson, M. S. Control of central and peripheral tolerance by Aire: Control of central and peripheral tolerance by Aire. *Immunol. Rev.* **241**, 89–103 (2011).
106. Gu, B. *et al.* Aire regulates the expression of differentiation-associated genes and self-renewal of embryonic stem cells. *Biochem. Biophys. Res. Commun.* **394**, 418–423 (2010).
107. Guerau-de-Arellano, M., Mathis, D. & Benoist, C. Transcriptional impact of Aire varies with cell type. *Proc. Natl. Acad. Sci.* **105**, 14011–14016 (2008).
108. Whyte, W. A. *et al.* Master Transcription Factors and Mediator Establish Super-Enhancers at Key Cell Identity Genes. *Cell* **153**, 307–319 (2013).
109. Sillanpää, N. *et al.* Autoimmune regulator induced changes in the gene expression profile of human monocyte-dendritic cell-lineage. *Mol. Immunol.* **41**, 1185–1198 (2004).
110. Colomé, N. *et al.* Increased Apoptosis after Autoimmune Regulator Expression in Epithelial Cells Revealed by a Combined Quantitative Proteomics Approach. *J. Proteome Res.* **9**, 2600–2609 (2010).
111. Gutmann, D. H. & Niederhuber, J. E. Major histocompatibility complex regulation of the immune response. *J. Surg. Res.* **39**, 172–181 (1985).
112. Kumánovics, A., Takada, T. & Lindahl, K. F. Genomic Organization of the Mammalian *Mhc*. *Annu. Rev. Immunol.* **21**, 629–657 (2003).
113. Milner, C. M. & Campbell, R. D. GENETIC ORGANIZATION OF THE HUMAN MHC CLASS III REGION. 13.
114. Simpson E. Function of the MHC. *Immunol Suppl.* 1988;1:27-30. PMID: 3133311.
115. Unanue, E. R., Turk, V. & Neefjes, J. Variations in MHC Class II Antigen Processing and Presentation in Health and Disease. *Annu. Rev. Immunol.* **34**, 265–297 (2016).

116. Radwan, J., Babik, W., Kaufman, J., Lenz, T. L. & Winternitz, J. Advances in the Evolutionary Understanding of MHC Polymorphism. *Trends Genet.* **36**, 298–311 (2020).
117. Klein & Sato. Birth of the Major Histocompatibility Complex. *Scand. J. Immunol.* **47**, 199–209 (1998).
118. Goodfellow PN, Jones EA, Van Heyningen V, Solomon E, Bobrow M, Miggiano V, et al. The beta2-microglobulin gene is on chromosome 15 and not in the HL-A region. *Nature*. 20 de marzo de 1975;254(5497):267-9.
119. Bjorkman PJ, Saper MA, Samraoui B, Bennett WS, Strominger JL, Wiley DC. Structure of the human class I histocompatibility antigen, HLA-A2. *Nature*. 8 de octubre de 1987;329(6139):506-12.
120. Cresswell, P., Turner, M. J. & Strominger, J. L. Papain-Solubilized HL-A Antigens from Cultured Human Lymphocytes Contain Two Peptide Fragments. *Proc. Natl. Acad. Sci.* **70**, 1603–1607 (1973).
121. Brown JH, Jardetzky TS, Gorga JC, Stern LJ, Urban RG, Strominger JL, et al. Pillars article: three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. *Nature* (1993).
122. Cresswell, P. Assembly, Transport, and Function of MHC Class II Molecules. 35.
123. Engelhard, V. H. Structure of Peptides Associated with Class I and Class II MHC Molecules. 27.
124. Wooldridge, L. *et al.* A Single Autoimmune T Cell Receptor Recognizes More Than a Million Different Peptides. *J. Biol. Chem.* **287**, 1168–1177 (2012).
125. Wieczorek, M. *et al.* Major Histocompatibility Complex (MHC) Class I and MHC Class II Proteins: Conformational Plasticity in Antigen Presentation. *Front. Immunol.* **8**, (2017).
126. Jardetzky TS, Lane WS, Robinson RA, Madden DR, Wiley DC. Identification of self peptides bound to purified HLA-B27. *Nature*.1991;353(6342):326-9.

127. Schumacher, T. N. M. *et al.* Peptide selection by MHC class I molecules. *Nature* **350**, 703–706 (1991).
128. Madden DR, Gorga JC, Strominger JL, Wiley DC. The structure of HLA-B27 reveals nonamer self-peptides bound in an extended conformation. *Nature*. 1991;353(6342):321-5.
129. Guo HC, Jardetzky TS, Garrett TP, Lane WS, Strominger JL, Wiley DC. Different length peptides bind to HLA-Aw68 similarly at their ends but bulge out in the middle. *Nature*.1992;360(6402):364-6.
130. DiBrino, M. *et al.* Endogenous peptides bound to HLA-A3 possess a specific combination of anchor residues that permit identification of potential antigenic peptides. *Proc. Natl. Acad. Sci.* **90**, 1508–1512 (1993).
131. Corr, M., Boyd, L. F. & Frankel, S. R. Endogenous Peptides of a Soluble Major Histocompatibility Complex Class I Molecule, H-2Lds: Sequence Motif, Quantitative Binding, and Molecular Modeling of the Complex. 12.
132. Engelhard VH *et al.* Mass spectrometric analysis of peptides associated with the human class I MHC molecules HLA-A2.1 and HLA-B7 and identification of structural features that determine binding. *Chem Immunol*. 1993.
133. Fremont DH, Matsumura M, Stura EA, Peterson PA, Wilson IA. Crystal structures of two viral peptides in complex with murine MHC class I H-2Kb. *Science*. 1992;257(5072):919-27.
134. Silver ML, Guo HC, Strominger JL, Wiley DC. Atomic structure of a human MHC molecule presenting an influenza virus peptide. *Nature*. 1992;360(6402):367-9.
135. Rammensee, H. G., Falk, K. & Rötzschke, O. Peptides Naturally Presented by MHC Class I Molecules. 32.
136. Rudensky AY, Preston-Hurlburt P, Hong S-C, Barlow A, Janeway CA. Pillars Article: Sequence Analysis of Peptides Bound to MHC Class II Molecules. *Nature*. 1991. 353: 622-627. *J Immunol Baltim Md* 1950. 2016;196(3):941- 6.

137. Sette A et al. HLA DR4w4-binding motifs illustrate the biochemical basis of degeneracy and specificity in peptide-DR interactions. *J Immunol Baltim Md* 1950. 1993;151(6):3163-70.
138. Chiczc RM, Urban RG, Gorga JC, Vignali DA, Lane WS, Strominger JL. Specificity and promiscuity among naturally processed peptides bound to HLA-DR alleles. *J Exp Med*. 1993;178(1):27-47.
139. Rudensky AY, Preston-Hurlburt P, al-Ramadi BK, Rothbard J, Janeway CA. Truncation variants of peptides isolated from MHC class II molecules suggest sequence motifs. *Nature*. 1992;359(6394):429-31.
140. Vyas, J. M., Van der Veen, A. G. & Ploegh, H. L. The known unknowns of antigen processing and presentation. *Nat. Rev. Immunol.* **8**, 607–618 (2008).
141. Yewdell, J. W., Schubert, U. & Bennink, J. R. DRiPs and other sources of peptide ligands for MHC class I molecules. **7**.
142. Reits, E. *et al.* A Major Role for TPPII in Trimming Proteasomal Degradation Products for MHC Class I Antigen Presentation. *Immunity* **20**, 495–506 (2004).
143. Geier, E. *et al.* A Giant Protease with Potential to Substitute for Some Functions of the Proteasome. *Science* **283**, 978–981 (1999).
144. Shen, X. Z., Lukacher, A. E., Billet, S., Williams, I. R. & Bernstein, K. E. Expression of Angiotensin-converting Enzyme Changes Major Histocompatibility Complex Class I Peptide Presentation by Modifying C Termini of Peptide Precursors. *J. Biol. Chem.* **283**, 9957–9965 (2008).
145. York, I. A. *et al.* The Cytosolic Endopeptidase, Thimet Oligopeptidase, Destroys Antigenic Peptides and Limits the Extent of MHC Class I Antigen Presentation. *Immunity* **18**, 429–440 (2003).
146. Howard, J. C. Supply and transport of peptides presented by class I MHC molecules. *Curr. Opin. Immunol.* **7**, 69–76 (1995).

147. Parcej, D. & Tampé, R. ABC proteins in antigen translocation and viral inhibition. *Nat. Chem. Biol.* **6**, 572–580 (2010).
148. York, I. A. *et al.* The ER aminopeptidase ERAP1 enhances or limits antigen presentation by trimming epitopes to 8–9 residues. *Nat. Immunol.* **3**, 1177–1184 (2002).
149. Serwold, T., Gonzalez, F., Kim, J., Jacob, R. & Shastri, N. ERAAP customizes peptides for MHC class I molecules in the endoplasmic reticulum. *Nature* **419**, 480–483 (2002).
150. Saric, T. *et al.* An IFN- $\gamma$ -induced aminopeptidase in the ER, ERAP1, trims precursors to MHC class I-presented peptides. *Nat. Immunol.* **3**, 1169–1176 (2002).
151. Saveanu, L. *et al.* Concerted peptide trimming by human ERAP1 and ERAP2 aminopeptidase complexes in the endoplasmic reticulum. *Nat. Immunol.* **6**, 689–697 (2005).
152. Rajagopalan S, Brenner MB. Calnexin retains unassembled major histocompatibility complex class I free heavy chains in the endoplasmic reticulum. *J Exp Med.* 1994;180(1):407-12.
153. Sadasivan, B., Lehner, P. J., Ortmann, B., Spies, T. & Cresswell, P. Roles for Calreticulin and a Novel Glycoprotein, Tapasin, in the Interaction of MHC Class I Molecules with TAP. *Immunity* **5**, 103–114 (1996).
154. Wearsch, P. A., Peaper, D. R. & Cresswell, P. Essential glycan-dependent interactions optimize MHC class I peptide loading. *Proc. Natl. Acad. Sci.* **108**, 4950–4955 (2011).
155. Ortmann B, Copeman J, Lehner PJ *et al.* A critical role for tapasin in the assembly and function of multimeric MHC class I complexes. *Science.* 1997;277(5330):1306-9.
156. Morrice, N. A. & Powis, S. J. A role for the thiol-dependent reductase ERp57 in the assembly of MHC class I molecules. *Curr. Biol.* **8**, 713–716 (1998).
157. Hughes, E. A. & Cresswell, P. The thiol oxidoreductase ERp57 is a component of the MHC class I peptide-loading complex. *Curr. Biol.* **8**, 709–713 (1998).

158. Fisetto, O., Wingbermühle, S., Tampé, R. & Schäfer, L. V. Molecular mechanism of peptide editing in the tapasin–MHC I complex. *Sci. Rep.* **6**, 19085 (2016).
159. Montserrat, V., Galocha, B., Marcilla, M., Vázquez, M. & López de Castro, J. A. HLA-B\*2704, an Allotype Associated with Ankylosing Spondylitis, Is Critically Dependent on Transporter Associated with Antigen Processing and Relatively Independent of Tapasin and Immunoproteasome for Maturation, Surface Expression, and T Cell Recognition: Relationship to B\*2705 and B\*2706. *J. Immunol.* **177**, 7015–7023 (2006).
160. Peh, C. A. *et al.* HLA-B27–Restricted Antigen Presentation in the Absence of Tapasin Reveals Polymorphism in Mechanisms of HLA Class I Peptide Loading. *Immunity* **8**, 531–542 (1998).
161. Hermann, C., Trowsdale, J. & Boyle, L. H. TAPBPR: a new player in the MHC class I presentation pathway: TAPBPR in MHC class I antigen presentation. *Tissue Antigens* **85**, 155–166 (2015).
162. Townsend A, Ohlén C, Bastin J, Ljunggren HG, Foster L, Kärre K. Association of class I major histocompatibility heavy and light chains induced by viral peptides. *Nature*. 1989;340(6233):443-8.
163. Knight, S. C. & Stagg, A. J. Antigen-presenting cell types. *Curr. Opin. Immunol.* **5**, 374–382 (1993).
164. Choi, N. M., Majumder, P. & Boss, J. M. Regulation of major histocompatibility complex class II genes. *Curr. Opin. Immunol.* **23**, 81–87 (2011).
165. Cresswell, P. & Roche, P. A. Invariant chain–MHC class II complexes: always odd and never invariant. *Immunol. Cell Biol.* **92**, 471–472 (2014).
166. Rock, K. L., Reits, E. & Neefjes, J. Present Yourself! By MHC Class I and MHC Class II Molecules. *Trends Immunol.* **37**, 724–737 (2016).

167. Roche PA, Marks MS, Cresswell P. Formation of a nine-subunit complex by HLA class II glycoproteins and the invariant chain. *Nature*. 1991;354(6352):392-4.
168. Anderson, K. S. & Cresswell', P. A role for calnexin (IP90) in the assembly of class II MHC molecules. *J. Immunol.* **168**, 2618–2625 (2002).
169. Neefjes, J. CIIV, MIIC and other compartments for MHC class II loading. *Eur. J. Immunol.* **29**, 1421–1425 (1999).
170. Hsieh, C.-S., deRoos, P., Honey, K., Beers, C. & Rudensky, A. Y. A Role for Cathepsin L and Cathepsin S in Peptide Generation for MHC Class II Presentation. *J. Immunol.* **168**, 2618–2625 (2002).
171. Mosyak, L., Zaller, D. M. & Wiley, D. C. The Structure of HLA-DM, the Peptide Exchange Catalyst that Loads Antigen onto Class II MHC Molecules during Antigen Presentation. *Immunity* **9**, 377–383 (1998).
172. Anders, A.-K. *et al.* HLA-DM captures partially empty HLA-DR molecules for catalyzed removal of peptide. *Nat. Immunol.* **12**, 54–61 (2011).
173. Denzin, L. K., Fallas, J. L., Prendes, M. & Yi, W. Right place, right time, right peptide: DO keeps DM focused. *Immunol. Rev.* **207**, 279–292 (2005).
174. Guce, A. I. *et al.* HLA-DO acts as a substrate mimic to inhibit HLA-DM by a competitive mechanism. *Nat. Struct. Mol. Biol.* **20**, 90–98 (2013).
175. Dongre, A. R. *et al.* In vivo MHC class II presentation of cytosolic proteins revealed by rapid automated tandem mass spectrometry and functional analyses. *Eur. J. Immunol.* **31**, 1485–1494 (2001).
176. Norbury, C. C. *et al.* CD8<sup>+</sup> T Cell Cross-Priming via Transfer of Proteasome Substrates. *Science* **304**, 1318–1321 (2004).



177. Grotzke, J. E., Sengupta, D., Lu, Q. & Cresswell, P. The ongoing saga of the mechanism(s) of MHC class I-restricted cross-presentation. *Curr. Opin. Immunol.* **46**, 89–96 (2017).
178. York, I. A. & Rock, K. L. ANTIGEN PROCESSING AND PRESENTATION BY THE CLASS I MAJOR HISTOCOMPATIBILITY COMPLEX. *Annu. Rev. Immunol.* **14**, 369–396 (1996).
179. Hershko, A. & Ciechanover, A. THE UBIQUITIN SYSTEM. **57** (1998).
180. Rock, K. L. & Goldberg, A. L. DEGRADATION OF CELL PROTEINS AND THE GENERATION OF MHC CLASS I-PRESENTED PEPTIDES. *Annu. Rev. Immunol.* **17**, 739–779 (1999).
181. Swatek, K. N. & Komander, D. Ubiquitin modifications. *Cell Res.* **26**, 399–422 (2016).
182. Song, L. & Luo, Z.-Q. Post-translational regulation of ubiquitin signaling. *J. Cell Biol.* **218**, 1776–1786 (2019).
183. Zheng, N. & Shabek, N. Ubiquitin Ligases: Structure, Function, and Regulation. *Annu. Rev. Biochem.* **86**, 129–157 (2017).
184. van Wijk, S. J. L. *et al.* A comprehensive framework of E2–RING E3 interactions of the human ubiquitin–proteasome system. *Mol. Syst. Biol.* **5**, 295 (2009).
185. Deshaies, R. J. & Joazeiro, C. A. P. RING Domain E3 Ubiquitin Ligases. *Annu. Rev. Biochem.* **78**, 399–434 (2009).
186. Metzger, M. B., Pruneda, J. N., Klevit, R. E. & Weissman, A. M. RING-type E3 ligases: Master manipulators of E2 ubiquitin-conjugating enzymes and ubiquitination. *Biochim. Biophys. Acta BBA - Mol. Cell Res.* **1843**, 47–60 (2014).
187. Lorenz, S. Structural mechanisms of HECT-type ubiquitin ligases. *Biol. Chem.* **399**, 127–145 (2018).

188. Nandi, D., Tahiliani, P., Kumar, A. & Chandu, D. The ubiquitin-proteasome system. *J Biosci* 19 (2006).
189. Voges, D., Zwickl, P. & Baumeister, W. The 26S Proteasome: A Molecular Machine Designed for Controlled Proteolysis. *Annu. Rev. Biochem.* **68**, 1015–1068 (1999).
190. Budenholzer, L., Cheng, C. L., Li, Y. & Hochstrasser, M. Proteasome Structure and Assembly. *J. Mol. Biol.* **429**, 3500–3524 (2017).
191. Davis, C., Spaller, B. L. & Matouschek, A. Mechanisms of substrate recognition by the 26S proteasome. *Curr. Opin. Struct. Biol.* **67**, 161–169 (2021).
192. Grziwa, A., Baumeister, W., Dahlmann, B. & Kopp, F. Localization of subunits in proteasomes from *Thermoplasma acidophilum* by immunoelectron microscopy. *FEBS Lett.* **290**, 186–190 (1991).
193. Vigneron, N. & Van den Eynde, B. Proteasome Subtypes and Regulators in the Processing of Antigenic Peptides Presented by Class I Molecules of the Major Histocompatibility Complex. *Biomolecules* **4**, 994–1025 (2014).
194. Heinemeyer, W., Fischer, M., Krimmer, T., Stachon, U. & Wolf, D. H. The Active Sites of the Eukaryotic 20 S Proteasome and Their Involvement in Subunit Precursor Processing. *J. Biol. Chem.* **272**, 25200–25209 (1997).
195. Aki M, Shimbara N, Takashina M, Akiyama K, Kagawa S, Tamura T, et al. Interferon-gamma induces different subunit organizations and functional diversity of proteasomes. *J Biochem (Tokyo)*. 1994;115(2):257-69.
196. Gaczynska M, Rock KL, Goldberg AL. Gamma-interferon and expression of MHC genes regulate peptide hydrolysis by proteasomes. *Nature*. 1993;365(6443):264-7.
197. Driscoll J, Brown MG, Finley D, Monaco JJ. MHC-linked LMP gene products specifically alter peptidase activities of the proteasome. *Nature*. 1993;365(6443):262-4.

198. Kincaid, E. Z. *et al.* Mice completely lacking immunoproteasomes show major changes in antigen presentation. *Nat. Immunol.* **13**, 129–135 (2012).
199. Murata, S. *et al.* Regulation of CD8<sup>+</sup> T Cell Development by Thymus-Specific Proteasomes. *Science* **316**, 1349–1353 (2007).
200. Tomko, R. J. & Hochstrasser, M. Molecular Architecture and Assembly of the Eukaryotic Proteasome. *Annu. Rev. Biochem.* **82**, 415–445 (2013).
201. Pepper, I. J., Van Sciver, R. E. & Tang, A. H. Phylogenetic analysis of the SINA/SIAH ubiquitin E3 ligase family in Metazoa. *BMC Evol. Biol.* **17**, 182 (2017).
202. Qi, J., Kim, H., Scortegagna, M. & Ronai, Z. A. Regulators and Effectors of Siah Ubiquitin Ligases. *Cell Biochem. Biophys.* **67**, 15–24 (2013).
203. Zhang, Q. *et al.* The substrate binding domains of human SIAH E3 ubiquitin ligases are now crystal clear. *Biochim. Biophys. Acta BBA - Gen. Subj.* **1861**, 3095–3105 (2017).
204. Gopalsamy, A., Hagen, T. & Swaminathan, K. Investigating the Molecular Basis of Siah1 and Siah2 E3 Ubiquitin Ligase Substrate Specificity. *PLoS ONE* **9**, e106547 (2014).
205. Polekhina, G. *et al.* Siah ubiquitin ligase is structurally related to TRAF and modulates TNF- $\alpha$  signaling. *Nat. Struct. Biol.* **9**, 68–75 (2002).
206. Möller, A. *et al.* Inhibition of Siah ubiquitin ligase function. *Oncogene* **28**, 289–296 (2009).
207. Santelli, E. *et al.* Structural Analysis of Siah1-Siah-interacting Protein Interactions and Insights into the Assembly of an E3 Ligase Multiprotein Complex. *J. Biol. Chem.* **280**, 34278–34287 (2005).
208. House, C. M. *et al.* A binding motif for Siah ubiquitin ligase. *Proc. Natl. Acad. Sci.* **100**, 3101–3106 (2003).

209. Matsuzawa, S. & Reed, J. C. Siah-1, SIP, and Ebi Collaborate in a Novel Pathway for  $\beta$ -Catenin Degradation Linked to p53 Responses. *Mol. Cell* **12**.
210. Velasco, K. *et al.* An N-terminal SIAH-interacting motif regulates the stability of the ubiquitin specific protease (USP)-19. *Biochem. Biophys. Res. Commun.* **433**, 390–395 (2013).
211. Bhattacharya, S. *et al.* The Modular Structure of SIP Facilitates Its Role in Stabilizing Multiprotein Assemblies. *Biochemistry* **44**, 9462–9471 (2005).
212. Skowyra, D., Craig, K. L., Tyers, M., Elledge, S. J. & Harper, J. W. F-Box Proteins Are Receptors that Recruit Phosphorylated Substrates to the SCF Ubiquitin-Ligase Complex. *Cell* **91**, 209–219 (1997).
213. Topolska-Woś, A. M., Chazin, W. J. & Filipek, A. CacyBP/SIP — Structure and variety of functions. *Biochim. Biophys. Acta BBA - Gen. Subj.* **1860**, 79–85 (2016).
214. Habelhah, H. Stress-induced decrease in TRAF2 stability is mediated by Siah2. *EMBO J.* **21**, 5756–5765 (2002).
215. Couillault, C. *et al.* TLR-independent control of innate immunity in *Caenorhabditis elegans* by the TIR domain adaptor protein TIR-1, an ortholog of human SARM. *Nat. Immunol.* **5**, 488–494 (2004).
216. House, C. M., Möller, A. & Bowtell, D. D. L. Siah Proteins: Novel Drug Targets in the Ras and Hypoxia Pathways. *Cancer Res.* **69**, 8835–8838 (2009).
217. Baba, K., Miyazaki, T., & Department of Probiotics Immunology, Institute for Genetic Medicine, Hokkaido University, North 15 West 7, Kita-ku, Sapporo City, Hokkaido, Japan. Critical function of Siah2 in tumorigenesis. *AIMS Mol. Sci.* **4**, 415–423 (2017).
218. Hara, M. R. *et al.* S-nitrosylated GAPDH initiates apoptotic cell death by nuclear translocation following Siah1 binding. *Nat. Cell Biol.* **7**, 665–674 (2005).

219. Russell, W. C., Graham, F. L., Smiley, J. & Nairn, R. Characteristics of a Human Cell Line Transformed by DNA from Human Adenovirus Type 5. *J. Gen. Virol.* **36**, 59–72 (1977).
220. Belfiore et al. - De nova HLA Class II and Enhanced HLA Class I Mole.pdf.
221. Lyapun, I. N., Andryukov, B. G. & Bynina, M. P. HeLa Cell Culture: Immortal Heritage of Henrietta Lacks. *Mol. Genet. Microbiol. Virol.* **34**, 195–200 (2019).
222. Fogh J., Trempe, G. (1975). New Human Tumor Cell Lines. In: Fogh, J. (eds) Human Tumor Cells in Vitro . Springer, Boston.
223. Ramos-Fernández, A., Paradela, A., Navajas, R. & Albar, J. P. Generalized Method for Probability-based Peptide and Protein Identification from Tandem Mass Spectrometry Data and Sequence Database Searching. *Mol. Cell. Proteomics* **7**, 1748–1754 (2008).
224. Lopez-Serra, P. et al. A DERL3-associated defect in the degradation of SLC2A1 mediates the Warburg effect. *Nat. Commun.* **5**, 3608 (2014).
225. Liiv, I. et al. DNA-PK contributes to the phosphorylation of AIRE: Importance in transcriptional activity. *Biochim. Biophys. Acta BBA - Mol. Cell Res.* **1783**, 74–83 (2008).
226. Zhu, Z.-C., Liu, J.-W., Li, K., Zheng, J. & Xiong, Z.-Q. KPNB1 inhibition disrupts proteostasis and triggers unfolded protein response-mediated apoptosis in glioblastoma cells. *Oncogene* **37**, 2936–2952 (2018).
227. Meloni, A. et al. DAXX Is a New AIRE-interacting Protein. *J. Biol. Chem.* **285**, 13012–13021 (2010).
228. Liiv, I. et al. AIRE-induced apoptosis is associated with nuclear translocation of stress sensor protein GAPDH. *Biochem. Biophys. Res. Commun.* **423**, 32–37 (2012).
229. Yang, F., Wei, Q., Adelstein, R. S. & Wang, P. J. Non-muscle myosin IIB is essential for cytokinesis during male meiotic cell divisions. *Dev. Biol.* **369**, 356–361 (2012).

230. Vicente-Manzanares, M., Ma, X., Adelstein, R. S. & Horwitz, A. R. Non-muscle myosin II takes centre stage in cell adhesion and migration. *Nat. Rev. Mol. Cell Biol.* **10**, 778–790 (2009).
231. Wang, Y. *et al.* Myosin Heavy Chain 10 (MYH10) Gene Silencing Reduces Cell Migration and Invasion in the Glioma Cell Lines U251, T98G, and SHG44 by Inhibiting the Wnt/ $\beta$ -Catenin Pathway. *Med. Sci. Monit.* **24**, 9110–9119 (2018).
232. Kumar, P. G. *et al.* The Autoimmune Regulator (AIRE) Is a DNA-binding Protein. *J. Biol. Chem.* **276**, 41357–41364 (2001).
233. Pinto, S. *et al.* An Organotypic Coculture Model Supporting Proliferation and Differentiation of Medullary Thymic Epithelial Cells and Promiscuous Gene Expression. *J. Immunol.* **190**, 1085–1093 (2013).
234. Monteleone-Cassiano, A. C. *et al.* The absence of the autoimmune regulator gene (AIRE) impairs the three-dimensional structure of medullary thymic epithelial cell spheroids. *BMC Mol. Cell Biol.* **23**, 15 (2022).
235. Zheng, Y. *et al.* An integrated overview of spatiotemporal organization and regulation in mitosis in terms of the proteins in the functional supercomplexes. *Front. Microbiol.* **5**, (2014).
236. Tykocinski, L.-O. *et al.* Epigenetic regulation of promiscuous gene expression in thymic medullary epithelial cells. *Proc. Natl. Acad. Sci.* **107**, 19426–19431 (2010).
237. Sansom, S. N. *et al.* Population and single-cell genomics reveal the *Aire* dependency, relief from Polycomb silencing, and distribution of self-antigen expression in thymic epithelia. *Genome Res.* **24**, 1918–1931 (2014).
238. Autoimmune Regulator Enhanced the Expression of Caspase-3 and Did Not Induce Massive Germ Cell Apoptosis in GC1-Spg Cells. *Cell. Physiol. Biochem.* **54**, 40–52 (2019).
239. Peng, Y.-T., Chen, P., Ouyang, R.-Y. & Song, L. Multifaceted role of prohibitin in cell survival and apoptosis. *Apoptosis* **20**, 1135–1149 (2015).

240. Ma, Y. *et al.* Plastin 3 down-regulation augments the sensitivity of MDA-MB-231 cells to paclitaxel via the p38 MAPK signalling pathway. *Artif. Cells Nanomedicine Biotechnol.* **47**, 684–694 (2019).
241. Zhong, S. *et al.* Promyelocytic Leukemia Protein (Pml) and Daxx Participate in a Novel Nuclear Pathway for Apoptosis. *J. Exp. Med.* **191**, 631–640 (2000).
242. Sparks, A. E., Chen, C., Breslin, M. B. & Lan, M. S. Functional Domains of Autoimmune Regulator (AIRE) Modulate INS-VNTR Transcription in Human Thymic Epithelial Cells. *J. Biol. Chem.* **291**, 11313–11322 (2016).
243. Palacios-Rodríguez, Y. *et al.* Polypeptide Modulators of Caspase Recruitment Domain (CARD)-CARD-mediated Protein-Protein Interactions. *J. Biol. Chem.* **286**, 44457–44466 (2011).
244. van Laar, G. G., van Hamburg, J. P. & Tas, S. W. Extrathymic AIRE-expressing cells: Friends or foes in autoimmunity and cancer? *Autoimmun. Rev.* **21**, 103141 (2022).
245. Nguyen, C. T. K. *et al.* AIRE is induced in oral squamous cell carcinoma and promotes cancer gene expression. *PLOS ONE* **15**, e0222689 (2020).
246. Uchiumi, A., Yamashita, M. & Katagata, Y. Downregulation of keratins 8, 18 and 19 influences invasiveness of human cultured squamous cell carcinoma and adenocarcinoma cells. *Exp. Ther. Med.* **3**, 443–448 (2012).
247. Crowe, D. L., Milo, G. E. & Shuler, C. E. Keratin 19 Downregulation by Oral Squamous Cell Carcinoma Lines Increases Invasive Potential. *J. Dent. Res.* **78**, 1256–1263 (1999).
248. Zhong, Z., Zhang, F. & Yin, S.-C. Effects of TESTIN Gene Expression on Proliferation and Migration of the 5-8F Nasopharyngeal Carcinoma Cell Line. *Asian Pac. J. Cancer Prev.* **16**, 2555–2559 (2015).
249. Wang, Q. *et al.* Extracellular calumenin suppresses ERK1/2 signaling and cell migration by protecting fibulin-1 from MMP-13-mediated proteolysis. *Oncogene* **34**, 1006–1018 (2015).

250. Bianchi, F. *et al.* Expression and prognostic significance of the autoimmune regulator gene in breast cancer cells. *Cell Cycle* **15**, 3220–3229 (2016).
251. Abbas, T. & Dutta, A. p21 in cancer: intricate networks and multiple activities. *Nat. Rev. Cancer* **9**, 400–414 (2009).
252. Huoh, Y.-S. *et al.* Dual functions of Aire CARD multimerization in the transcriptional regulation of T cell tolerance. *Nat. Commun.* **11**, 1625 (2020).
253. Björse, P. *et al.* Mutations in the AIRE Gene: Effects on Subcellular Location and Transactivation Function of the Autoimmune Polyendocrinopathy–Candidiasis–Ectodermal Dystrophy Protein. *Am. J. Hum. Genet.* **66**, 378–392 (2000).
254. Ramsey, C. Systematic mutagenesis of the functional domains of AIRE reveals their role in intracellular targeting. *Hum. Mol. Genet.* **11**, 3299–3308 (2002).
255. Heery, D. M., Kalkhoven, E., Hoare, S. & Parker, M. G. A signature motif in transcriptional co-activators mediates binding to nuclear receptors. *Nature* **387**, 733–736 (1997).
256. Ishii, T., Suzuki, Y., Ando, N., Matsuo, N. & Ogata, T. Novel Mutations of the Autoimmune Regulator Gene in Two Siblings with Autoimmune Polyendocrinopathy–Candidiasis–Ectodermal Dystrophy. **85**, 5 (2000).
257. Žumer, K., Plemenitaš, A., Saksela, K. & Peterlin, B. M. Patient mutation in AIRE disrupts P-TEFb binding and target gene transcription. *Nucleic Acids Res.* **39**, 7908–7919 (2011).
258. Org, T. *et al.* The autoimmune regulator PHD finger binds to non-methylated histone H3K4 to activate gene expression. *EMBO Rep.* **9**, 370–376 (2008).
259. Somma, M. P. *et al.* Moonlighting in Mitosis: Analysis of the Mitotic Functions of Transcription and Splicing Factors. *Cells* **9**, 1554 (2020).



260. Ji, L. *et al.* The SIAH E3 ubiquitin ligases promote Wnt/ $\beta$ -catenin signaling through mediating Wnt-induced Axin degradation. *Genes Dev.* **31**, 904–915 (2017).
261. Ball, L. J., Kühne, R., Schneider-Mergener, J. & Oschkinat, H. Recognition of Proline-Rich Motifs by Protein-Protein-Interaction Domains. *Angew. Chem. Int. Ed.* **44**, 2852–2869 (2005).
262. Pitkänen, J. *et al.* The Autoimmune Regulator Protein Has Transcriptional Transactivating Properties and Interacts with the Common Coactivator CREB-binding Protein. *J. Biol. Chem.* **275**, 16802–16809 (2000).
263. Buchwald, M. *et al.* SIAH ubiquitin ligases target the nonreceptor tyrosine kinase ACK1 for ubiquitinylation and proteasomal degradation. *Oncogene* **32**, 4913–4920 (2013).
264. Germani, A. *et al.* SIAH-1 interacts with  $\alpha$ -tubulin and degrades the kinesin Kid by the proteasome pathway during mitosis. **10**.
265. Filipek, A., Jastrzebska, B., Nowotny, M. & Kuznicki, J. CacyBP/SIP, a Calcyclin and Siah-1-interacting Protein, Binds EF-hand Proteins of the S100 Family. *J. Biol. Chem.* **277**, 28848–28852 (2002).
266. Uchida, D. *et al.* AIRE Functions As an E3 Ubiquitin Ligase. *J. Exp. Med.* **199**, 167–172 (2004).
267. Coscoy, L. PHD domains and E3 ubiquitin ligases: viruses make the connection. *Trends Cell Biol.* **13**, 7–12 (2003).
268. Boname, J. M. & Stevenson, P. G. MHC Class I Ubiquitination by a Viral PHD/LAP Finger Protein. *Immunity* **15**, 627–636 (2001).
269. Shao, W., Zumer, K., Fujinaga, K. & Peterlin, B. M. FBXO3 Protein Promotes Ubiquitylation and Transcriptional Activity of AIRE (Autoimmune Regulator). *J. Biol. Chem.* **291**, 17953–17963 (2016)

270. Huoh, Y.-S. *et al.* Dual functions of Aire CARD multimerization in the transcriptional regulation of T cell tolerance. *Nat. Commun.* **11**, 1625 (2020).
271. Zhao, R., Bodnar, M. S. & Spector, D. L. Nuclear neighborhoods and gene expression. *Curr. Opin. Genet. Dev.* **19**, 172–179 (2009).
272. Chang, H. R. *et al.* The functional roles of PML nuclear bodies in genome maintenance. *Mutat. Res. Mol. Mech. Mutagen.* **809**, 99–107 (2018).
273. Fiucci, G. *et al.* Siah-1b is a direct transcriptional target of p53: Identification of the functional p53 responsive element in the siah-1b promoter. *Proc. Natl. Acad. Sci.* **101**, 3510–3515 (2004).



