

TESIS DOCTORAL

Programa de Doctorado en Biomedicina (B11.56.1)

**PAPEL DEL *ENHANCER* DEL GEN DE LA  
CADENA  $\alpha$  DEL RECEPTOR DE  
LINFOCITOS T EN LA EXPRESIÓN DEL  
GEN REORDENADO**

**JESÚS ÁLVAREZ SANTIAGO**

INSTITUTO DE PARASITOLOGÍA Y BIOMEDICINA “LÓPEZ-NEYRA”  
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS (CSIC)

**UNIVERSIDAD DE GRANADA**

**IPBLN, CSIC**

**GRANADA, 2022**



**UNIVERSIDAD  
DE GRANADA**



**DIRECCIÓN: María Cristina Hernández López de Munain**

Editor: Universidad de Granada. Tesis Doctorales  
Autor: Jesús Álvarez Santiago  
ISBN: 978-84-1117-656-9  
URI: <https://hdl.handle.net/10481/79645>



Esta tesis ha sido financiada por proyectos del Plan Nacional de I+D+i del Ministerio de Economía y Competitividad (BFU2013-44660-R y BFU2016-79699-P), del Consejo Superior de Investigaciones Científicas (2019AEP202), y de la Junta de Andalucía (P20\_01271).

## **AGRADECIMIENTOS**

Al finalizar un trabajo tan duro y lleno de dificultades como es una tesis doctoral es inevitable que te asalte un sentimiento tan humano como el egocentrismo que te lleva a concentrar la mayor parte del mérito en uno mismo. Sin embargo, un análisis objetivo muestra inmediatamente que este trabajo hubiera sido imposible sin la participación, apoyo y fortaleza otorgadas por tantas personas e instituciones, por lo que para concluir esta gran etapa de mi vida quiero extender mi más sincero agradecimiento a quienes han hecho posible este logro.

En primer lugar, debo agradecer de manera especial a mi directora de tesis, la Doctora M<sup>a</sup> Cristina Hernández López de Munain. Su apoyo y confianza en mi trabajo y su capacidad para guiar y tener nuevas ideas ha sido un aporte esencial para que este trabajo llegase a terminar bien y para ofrecerme una buena formación como investigador. Gracias por todo y deseo que siga teniendo éxito en su carrera laboral.

Quiero dar las gracias también al investigador jefe de nuestro laboratorio vecino, al Doctor Carles María Suñé Negre, por sus ideas y aportaciones en tantos seminarios conjuntos. No cabe duda de que su participación ha enriquecido y mejorado el presente trabajo.

A mis compañeros de laboratorio, que han sido muchos en este tiempo, pero en especial a Jenifer y al ya doctor Alonso, gracias por enseñarme tantas cosas durante este tiempo, por vuestra gran ayuda en tantos experimentos, por vuestra paciencia, por vuestro apoyo en las fatigas más largas, pero, sobre todo, por vuestras risas que hacían más livianos los días más duros. También a mis compañeros del López-Neyra con los que he compartido tiempo y buenos momentos durante este periodo, en especial al laboratorio 109, la salud mental la conservo, en parte, gracias a vosotros.

Agradecer también a Marga, Max y Pablo que de una forma u otra habéis sido mis principales compañeros de vida en los últimos años, el apoyo y la chispa para que todos y cada uno de los días terminaran siendo buenos por muy mal que se pusieran las cosas. Quiero agradecer también a mis compañeros de

generación con los que empecé esta etapa de mi vida en el Máster, por todos los buenos momentos que compartí con ellos y por brindarme su amistad.

Por último, aunque no menos importante, a mi familia, principalmente a mis padres. Por ser los principales promotores de mis sueños, por confiar y creer en mis expectativas, por los consejos, por los valores y principios que me habéis inculcado y por mucho más, os quiero, aunque no siempre lo demuestre como debiera, gracias de todo corazón.

## ÍNDICE

RESUMEN .....	15
ABREVIATURAS .....	21
INTRODUCCIÓN .....	27
1. Linfocitos.....	27
1.1 Desarrollo de los linfocitos T .....	28
1.2. El receptor del linfocito T (TCR).....	30
1.2.1. Estructura y función del TCR.....	30
1.2.2. Estructura de los loci <i>Tcrb</i> y <i>Tcra/Tcrd</i> .....	33
2. La recombinación V(D)J .....	34
2.1. Regulación de la recombinación V(D)J .....	37
2.2. Transcripción germinal y recombinación V(D)J.....	38
2.3. Recombinación V(D)J de los genes del TCR durante el desarrollo de los linfocitos T .....	40
3. Organización espacial del genoma .....	44
3.1. Estructura de la cromatina.....	44
3.2. Dominios asociados topológicamente (TAD).....	47
3.3. Organización tridimensional del locus <i>Tcra/Tcrd</i> .....	49
4. Elementos reguladores en <i>cis</i> : <i>enhancers</i> y promotores .....	54
4.1 <i>Enhancers</i> .....	54
4.2 Promotores.....	56
4.3. Holocomplejos promotor- <i>enhancer</i> .....	57
5. El <i>enhancer</i> del gen <i>Tcra</i> (Ea) .....	58
5.1. Papel de Ea en la regulación transcripcional del gen <i>Tcra</i> .....	59
5.2. Papel de Ea en la regulación transcripcional de los genes del 3' sub- TAD .....	60
5.3. <i>Enhanceosoma</i> de Ea durante el desarrollo de los linfocitos T $\alpha\beta$ .....	61
6. Antecedentes.....	63
La actividad de Ea se inhibe en el contexto del gen no reordenado en linfocitos T $\alpha\beta$ .....	63
HIPÓTESIS .....	69
OBJETIVOS .....	69
MATERIALES Y MÉTODOS .....	73
1. Líneas celulares .....	73
2. Cepas de ratón .....	73
3. Obtención de timocitos y linfocitos T $\alpha\beta$ de ratón.....	74
4. Purificación celular mediante citometría de flujo.....	75

5. Extracción de ARN.....	76
6. Retrotranscripción (RT) .....	77
7. PCR cuantitativa (qPCR) .....	77
8. Eliminación de E $\alpha$ mediante edición génica por CRISPR-Cas9.....	82
8.1 Generación de clones de Jurkat E $\alpha$ <sup>+/-</sup> .....	82
8.1.1) Preparación de las guías de ARN (gARN) y plásmidos utilizados.....	82
8.1.2) Transfección de células Jurkat para la delección de E $\alpha$ .....	86
8.1.3) Genotipado de clones Jurkat .....	87
8.2 Generación de clones Jurkat E $\alpha$ <sup>-/-</sup> .....	89
9. Transformación bacteriana y purificación de plásmidos.....	90
9.1 <i>Midipreps</i> .....	90
9.2 <i>Minipreps</i> .....	91
10. Preparación de muestras para secuenciación .....	92
11. 4C-seq .....	92
12. Análisis estadísticos .....	97
RESULTADOS .....	100
1. La transcripción del gen reordenado de la cadena TCR $\alpha$ depende de E $\alpha$ en linfocitos T $\alpha\beta$ .....	100
2. Secuencias distales colaboran con E $\alpha$ para el mantenimiento de la transcripción del gen <i>Tcra</i> reordenado en linfocitos T $\alpha\beta$ .....	112
3. Análisis de las interacciones globales en el contexto del gen <i>Tcra</i> no reordenado en timocitos DP y del gen <i>Tcra</i> reordenado en linfocitos T $\alpha\beta$ .....	117
CONCLUSIONES.....	126
DISCUSIÓN .....	130
BIBLIOGRAFÍA .....	143
PUBLICACIONES.....	167

## ÍNDICE DE FIGURAS

Figura 1. Estadios de diferenciación de los timocitos.....	29
Figura 2. Representación de la estructura del complejo del TCR y de la interacción TCR-célula presentadora de antígeno (APC) .....	31
Figura 3. Esquema de la estructura genómica del locus <i>Tcra/Tcrd</i> .....	33
Figura 4. Esquema del proceso de recombinación V(D)J .....	37
Figura 5. Representación de los loci del TCR sin reordenar y reordenados durante el desarrollo de los timocitos .....	41
Figura 6. Esquema representativo de los diferentes grados de compactación del ADN en el núcleo celular .....	46
Figura 7. Esquema representativo de la formación de un lazo de cromatina mediado por interacciones CTCF y cohesina .....	48
Figura 8. Modelos de configuración de los lazos de cromatina en el locus <i>Tcra/Tcrd</i> .....	52
Figura 9. Posición genómica del 5' sub-TAD y los genes del 3' sub-TAD respecto a la de Ea .....	54
Figura 10. Esquema representativo de las regiones de unión de factores de transcripción a Ea .....	62
Figura 11. La transcripción germinal no codificante del gen <i>Tcra</i> dependiente de Ea se encuentra fuertemente inhibida en timocitos SP y linfocitos T $\alpha\beta$ .....	64
Figura 12. Inhibición de Ea introducido en el locus <i>Tcrb</i> en timocitos SP y linfocitos T en ratones <i>Knock-In</i> (E $\alpha$ KI).....	66
Figura 13. Esquema del proceso de formación de las construcciones que contienen dos gARN .....	83
Figura 14. Lugar de anillamiento y orientación de los oligonucleótidos utilizados sobre la secuencia genómica humana .....	88
Figura 15. Transcripción no codificante dependiente de Ea del gen <i>Tcra</i> durante el desarrollo de los timocitos .....	101
Figura 16. Transcripción del gen <i>Tcra</i> en timocitos y linfocitos T de ratones <i>wild-type</i> (WP) .....	102
Figura 17. Transcripción del gen <i>Tcra</i> en linfocitos T $\alpha\beta$ de ratones WT y E $\alpha$ <sup>-/-</sup> y expresión de TCR $\alpha\beta$ y Va2 en timocitos y linfocitos T .....	104
Figura 18. Comparación de las secuencias de Ea humano y de ratón y estrategia de edición CRISPR-Cas9 para la eliminación de Ea en células Jurkat .....	107

Figura 19. Expresión en superficie del TCR $\alpha\beta$ y transcripción del gen <i>TCRA</i> reordenado en clones Jurkat E $\alpha^{+/-}$ y E $\alpha^{-/-}$ .....	108
Figura 20. Secuencias de los sitios de delección en los clones Jurkat E $\alpha^{+/-}$ .....	109-111
Figura 21. Secuencias de los sitios de delección en los clones Jurkat E $\alpha^{-/-}$ .....	112
Figura 22. Estructura de los transgenes <i>Tcra</i> DO11.10 y <i>Tcra</i> OT-I. Transcripción de los transgenes <i>Tcra</i> en timocitos DP y linfocitos T $\alpha\beta$ de ratones <i>Rag2</i> <sup>-/-</sup> x DO11.10 y <i>Rag2</i> <sup>-/-</sup> x OT-I .....	115
Figura 23. Análisis de la cromatina mediante ATAC-seq durante la diferenciación de los timocitos a células T $\alpha\beta$ .....	119
Figura 24. Contactos a larga distancia en el ADN en el contexto del gen <i>Tcra</i> no reordenado en timocitos DP y del gen <i>Tcra</i> reordenado en linfocitos T $\alpha\beta$ .....	121
Figura 25. Cuantificación de las interacciones de las regiones C $\alpha$ peak y E $\alpha$ con las regiones del 5' sub-TAD y del 3' sub-TAD .....	124

## ÍNDICE DE TABLAS

Tabla 1. Secuencias de los oligonucleótidos usados .....	78-82
Tabla 2. Condiciones de la PCR para obtener la construcción del bloque U6-gARN-5' -H1-gARN-3' .....	84
Tabla 3. Pasos de PCR convencional .....	89
Tabla 4. Condiciones 1 <sup>a</sup> tanda PCR 4C-seq .....	95
Tabla 5. Condiciones 2 <sup>a</sup> tanda PCR 4C-seq .....	96



**RESUMEN**



Los linfocitos T  $\alpha\beta$  expresan en su superficie un receptor para antígeno denominado receptor de linfocitos T (TCR) formado por dos cadenas variables, TCR $\alpha$  y TCR $\beta$ , y una serie de cadenas invariables pertenecientes al complejo CD3. Los genes que codifican por las cadenas TCR $\alpha$  y TCR $\beta$  están formados por segmentos génicos dispersos en el genoma que necesitan reordenarse para dar lugar a una configuración génica capaz de generar un ARN mensajero que pueda traducirse en una proteína funcional. Por lo tanto, estos genes existen en dos configuraciones distintas: no reordenada y reordenada. Para que se den estos reordenamientos es necesario que se active la transcripción no codificante del gen no reordenado o transcripción germinal a través de la acción de potenciadores transcripcionales (*enhancers*) y promotores asociados a los segmentos génicos.

El *enhancer* del gen de la cadena TCR $\alpha$ , E $\alpha$ , se activa durante el paso de timocitos dobles negativos (DN), denominados así por no expresar los marcadores de superficie CD4 y CD8, a timocitos dobles positivos (DP), denominados así por expresar ambos marcadores en superficie. E $\alpha$  es esencial para activar la transcripción germinal y la recombinación V $\alpha$ J $\alpha$  de este gen. Experimentos previos de nuestro laboratorio compararon la actividad de E $\alpha$  para activar la transcripción germinal del gen *Tcra* en timocitos DP y en linfocitos T  $\alpha\beta$ , y mostraron que su actividad está fuertemente reducida en estos últimos, siendo esta inhibición irreversible durante la proliferación y diferenciación de los linfocitos maduros. A pesar de que E $\alpha$  se encuentra fuertemente inhibido, no se observó ninguna caída en la transcripción del gen *Tcra* reordenado en linfocitos T  $\alpha\beta$  frente a su transcripción en el contexto del gen no reordenado en timocitos DP, lo cual sugirió que E $\alpha$  podría no estar implicado en la transcripción del gen o que requiere de la colaboración de otras secuencias distales para mantener la transcripción génica en el contexto del gen reordenado.

Nuestros datos indican que E $\alpha$  es necesario para la transcripción del gen reordenado de la cadena TCR $\alpha$  y la expresión del TCR $\alpha\beta$  en membrana de los linfocitos T  $\alpha\beta$ , tanto en ratones como en humanos. Para la transcripción

eficiente del gen *Tcra* reordenado y la expresión del TCR $\alpha\beta$  en linfocitos T, este *enhancer* necesita colaborar con otras secuencias no presentes dentro del transgén *Tcra* DO11.10, el cual consiste en el mayor transgén *Tcra* que existe y contiene una región de 33,5 kilobases que va desde *Traaj13* hasta el tercer exón del gen vecino *Dad1*. La posición génica de E $\alpha$  se encuentra en el límite entre dos sub-dominios asociados topológicamente (sub-TAD): 5' sub-TAD de 525 kilobases, constituido por secuencias del gen *Tcra*, y 3' sub-TAD de 400 kilobases, constituido por los genes que van desde *Dad1* a *Cdh24*. Con el fin de identificar las secuencias que colaboran con E $\alpha$  en el contexto del gen *Tcra* reordenado en linfocitos T  $\alpha\beta$ , hemos realizado estudios comparativos de las interacciones genómicas que ocurren entre determinadas secuencias del gen *Tcra* con el 5' sub-TAD y el 3' sub-TAD, tanto en el contexto del gen *Tcra* no reordenado en timocitos DP como en el contexto del gen *Tcra* reordenado en linfocitos T  $\alpha\beta$  mediante 4C-seq. Las secuencias utilizadas como anzuelo fueron seleccionadas en función de su localización genómica y de la accesibilidad de su cromatina en el paso de timocitos DP a linfocitos T  $\alpha\beta$  según análisis de ATAC-seq. Como era esperable, se observó que E $\alpha$  establece fuerte interacciones con el 5' sub-TAD en el contexto del gen *Tcra* no reordenado en timocitos DP, las cuales se pierden tras la recombinación del gen en linfocitos T  $\alpha\beta$ . Además, se observó que E $\alpha$  establece interacciones con regiones del 3' sub-TAD en ambos contextos génicos y tipos celulares. Estos análisis también identificaron que la secuencia denominada Ca peak, situada dentro de una región intrónica del dominio constante del gen *Tcra*, actúa junto con E $\alpha$  en las interacciones observadas con el 5' sub-TAD en timocitos DP en el contexto del gen *Tcra* no reordenado, y también con el 3' sub-TAD tanto en el contexto del gen *Tcra* no reordenado en timocitos DP como en el contexto del gen *Tcra* reordenado en linfocitos T  $\alpha\beta$ . El análisis cuantitativo de los datos por 4C-ker indicó que las interacciones genómicas de E $\alpha$  y Ca peak, tanto de forma conjunta como por separado, son más fuertes en el contexto del gen *Tcra* reordenado en linfocitos T  $\alpha\beta$  que en el contexto del gen *Tcra* no reordenado en timocitos DP. Estos datos indican que tanto E $\alpha$  como Ca peak participan en la formación de un nudo de cromatina en el 5' sub-TAD en el contexto del gen *Tcra* no reordenado en timocitos DP, como en el fortalecimiento de un nudo de cromatina en el 3' sub-TAD en el contexto del gen *Tcra* reordenado en

linfocitos T  $\alpha\beta$  respecto al existente en el contexto del gen *Tcra* no reordenado en timocitos DP. Puesto que los reordenamientos V $\alpha$ J $\alpha$  en linfocitos T  $\alpha\beta$  no inducen nuevas interacciones de largo alcance entre E $\alpha$  y C $\alpha$  peak y otras regiones genómicas que permanecen en *cis* tras la recombinación, las interacciones fuertes que se observan entre E $\alpha$  y C $\alpha$  peak con el 3' sub-TAD podrían prevenir la reducción de la transcripción del gen *Tcra* reordenado que sería esperable tras la inhibición de la función de E $\alpha$  que se observa en linfocitos T  $\alpha\beta$ .



## ABREVIATURAS



## ABREVIATURAS

- 3C, captura de conformación cromosómica
- 3D-FISH, *3-Dimensional-Fluorescence In Situ Hybridization*, hibridación *in situ* con sondas fluorescentes
- 4C, captura de conformación cromosómica circular
- $\beta$ tg, transgén de la cadena  $\beta$  del TCR
- ACK, *Ammonium-Chloride-Potassium*, Amonio-Cloruro-Potasio
- Actb*, gen de  $\beta$ -actina de ratón (humano, *ACTB*)
- ADN, ácido desoxirribonucleico
- ADNc, ADN complementario
- ADP, adenosina-difosfato
- APC, células presentadoras de antígeno
- ATAC, mediciones de accesibilidad a la cromatina mediadas por transposasa Tn5
- BCR, *B Cell Receptor*, receptor de linfocitos B
- BSA, *Bovine Serum Albumin*, albúmina de suero bovino
- CBE, *CTCF Binding Element*, secuencia de unión de CTCF
- CDR, *Complementary Determinant Region*, región determinante de la complementariedad
- ChIP, inmunoprecipitación de cromatina
- CRISPR, *Clustered Regularly Interspaced Short Palindromic Repeats*, repeticiones palindrómicas cortas agrupadas y regularmente interespaciadas.
- Ct, ciclo umbral de detección (referido a las qPCR)
- CTCF, *CCCTC binding factor*, factor de unión a la secuencia CCCTC
- D, *Diversity*, diversidad (referido a los segmentos génicos de los genes TCR y BCR)
- DEPC, dietilpirocarbonato
- DL, DELTA-LIKE
- DN, dobles negativos (referido a los timocitos CD4<sup>-</sup>CD8<sup>-</sup>)
- DNA-PK, proteína quinasa dependiente de ADN
- DP, dobles positivos (referido a los timocitos CD4<sup>+</sup>CD8<sup>+</sup>)

eDP, *early DP*, dobles positivos tempranos (referido a los timocitos CD4<sup>+</sup>CD8<sup>+</sup>)

EDTA sódico, etiléndiaminotetraacetato de sodio

eRNA, transcritos de *enhancer*

ETP, *Early T-cell Precursors*, Precusores tempranos de linfocitos T

E $\alpha$ , *enhancer* del gen de la cadena TCR $\alpha$

E $\alpha$ KI, E $\alpha$  *Knock-In*, referido a ratones o células con E $\alpha$  integrado en el locus *Tcrb*

E $\alpha$ KO, E $\alpha$  *Knock-Out*, referido a ratones o células con E $\alpha$  delecionado del genoma (E $\alpha$ <sup>-/-</sup>)

E $\delta$ , *enhancer* del gen *Tcrd*

FAIRE, detección de elementos reguladores mediado por formaldehído

FITC, isotiocianato de fluoresceína

GFP, *Green Fluorescent Protein*, proteína verde fluorescente

gARN, guía de ARN para experimentos de edición por CRISPR/Cas9

H3K27ac, acetilación de la lisina 27 de la histona H3

H3K27me, metilación de la lisina 27 de la histona H3

H3K4ac, acetilación de la lisina 4 de la histona H3

H3K4me (1, 2 o 3), mono-, di- y tri-metilación de la lisina 4 de la histona H3

H3K6me, metilación de la lisina 6 de la histona H3

H3K9me, metilación de la lisina 9 de la histona H3

HAT, reclutamiento de enzimas acetiltransferasas de histonas

*HBB*, gen de la subunidad  $\beta$  de la hemoglobina

HBSS, *Hank's Balanced Salt Solution*, solución salina equilibrada de Hank

HMG, *High Mobility Group*, grupo de alta movilidad (familia de proteínas)

HS, *Hypersensitivity Site*, sitio de hipersensibilidad (a la enzima DNasaI)

iFBS, *inactivated Fetal Bovine Serum* (FCS, *Fetal Calf Serum*), suero fetal bovino inactivado

IL-7, interleuquina 7

IL-7R, receptor de IL-7

ISP, *Immature Single Positives*, simples positivos inmaduros (referido a timocitos)

ITAM, *Immunoreceptor Tyrosine-based Activation Motif*, motivos de activación de inmunorreceptores basados en tirosina

J, *Joining*, unión (referido a los segmentos génicos de los genes TCR y BCR)

Kb, kilobase

LB, *Lysogeny Broth*, medio para el crecimiento bacteriano

LCR, *Locus Control Region*, región de control del locus

IDP, *late DP*, dobles positivos tardíos (referido a los timocitos CD4<sup>+</sup>CD8<sup>+</sup>)

Mb, megabase

MHC, *Major Histocompatibility Complex*, complejo principal de histocompatibilidad

NHEJ, *Non Homologous End Joining*, reparación de extremos no homólogos

NK, *Natural Killer*, células asesinas naturales

PAM, *Protospacer Adjacent Motif*, secuencia de reconocimiento de CAS9

Pb, pares de bases

PBS, *phosphate buffered solution*, solución buffer fosfato

PCR, reacción en cadena por polimerasa

PE, *Phycoerythrin*, ficoeritrina

PHD, *Plant HomeoDomain*

PMA, *Phorbol 12 Myristate 13 Acetate*

qPCR, PCR cuantitativa

R, cepa de ratón *Rag2*<sup>-/-</sup>

RAG, *Recombination Activating Gene*, gen de activación de la recombinación

Rpm, revoluciones por minuto

RPMI, *Roswell Park Memorial Institute medium*

RSS, *Recombination Signal Sequence*, secuencia señal de recombinación V(D)J

RT, retrotranscripción

Rx5C.7C, cepa de ratón *Rag2*<sup>-/-</sup> x 5C.7C

RxDO11.10, cepa de ratón *Rag2*<sup>-/-</sup> x DO11.10

RxOT-1, cepa de ratón *Rag2*<sup>-/-</sup> x OT-1

Rxβ, cepa de ratón *Rag2*<sup>-/-</sup> x βtg

SCID, *Severe Combined ImmunoDeficiency*, inmunodeficiencia combinada severa

SDS, *Sodium Dodecyl Sulfate*

SOC, *Super Optimal Broth with Catabolic repressor*

SP, simples positivos (referido a los timocitos)

Sv40, *simian virus 40*, virus del simio 40

TA, temperatura ambiente

TAD, *topologically associated domains*, dominios asociados topológicamente

T-ALL, *T-cell Acute Lymphoblastic Leukemia*, leucemia linfoblástica aguda de células T

Tc, linfocito T citotóxico

TCR, *T Cell Receptor*, receptor de linfocitos T

TCR $\alpha$ , cadena  $\alpha$  del TCR

TCR $\beta$ , cadena  $\beta$  del TCR

TCR $\delta$ , cadena  $\delta$  del TCR

TCR $\gamma$ , cadena  $\gamma$  del TCR

TdT, *Terminal deoxynucleotidyl-Transferase*, transferasa de desoxinucléotidos terminales

TEA, *T Early  $\alpha$* , T  $\alpha$  temprano (exón localizado 3' de los segmentos J $\alpha$ )

Th, linfocito T *helper* o ayudador

V, variable (referido a los segmentos génicos de los genes TCR y BCR)

WT, *Wild-Type*, tipo salvaje (referido a una cepa de ratón)

# INTRODUCCIÓN



## INTRODUCCIÓN

### 1. Linfocitos

El sistema inmunitario es la principal barrera que poseen los animales superiores para protegerse de las infecciones. Se pueden diferenciar dentro del sistema inmunitario dos respuestas principales. Por un lado, ocurre una respuesta innata que se activa ante cualquier agente extraño, actuando siempre igual frente a una misma infección y en la que participan barreras físicas como la piel y las membranas mucosas, barreras químicas como la lisozima o el ácido clorhídrico del jugo gástrico, y una importante variedad de células como los macrófagos, los neutrófilos o las células asesinas naturales (*Natural Killer*, NK), además de mediadores humorales, tales como el complemento y diversas linfoquinas (1). Por otro lado, ocurre una respuesta adaptativa o específica que puede ser a su vez humoral y celular. La respuesta humoral específica está mediada por los anticuerpos producidos por las células plasmáticas (linfocitos B activados y diferenciados), mientras que la respuesta celular está mediada por los linfocitos T, originados en el timo y en los que se centra el trabajo de la presente tesis. Los linfocitos tienen diversas funciones, las cuales incluyen la secreción de linfoquinas para ayudar tanto a la respuesta innata como adaptativa, la eliminación de células infectadas o tumorales, y la inactivación y destrucción de agentes patógenos y sus derivados tóxicos (1,2).

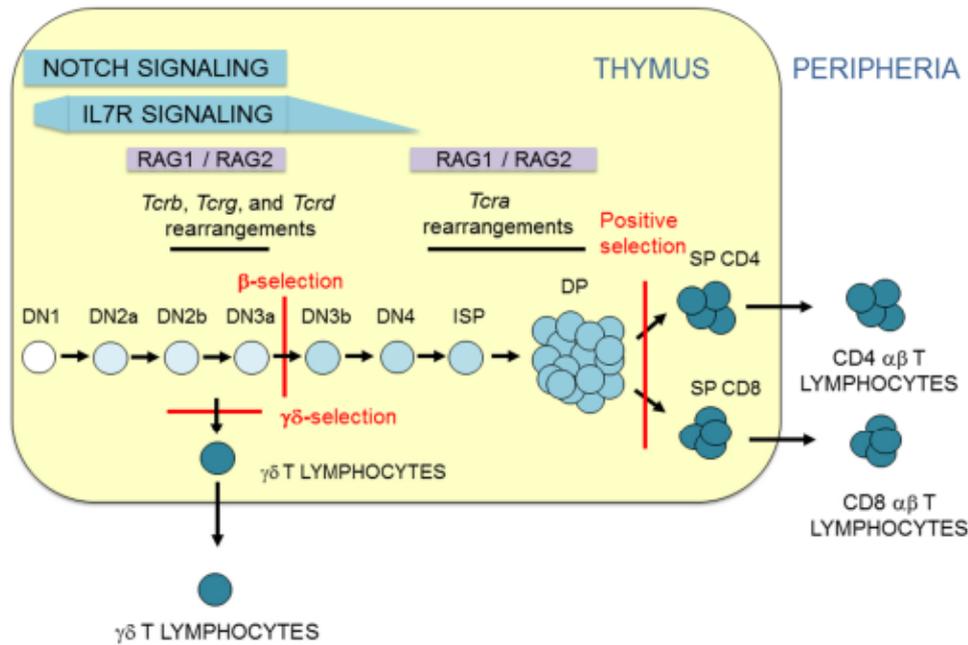
Los linfocitos presentan en la superficie de su membrana unos receptores que les permiten identificar específicamente a los antígenos; el linfocito T posee el denominado receptor de linfocitos T (*T Cell Receptor*, TCR) y el linfocito B posee el denominado receptor de linfocitos B (*B Cell Receptor*, BCR), siendo ambos receptores clonotípicos, es decir, únicos en cada célula virgen que no ha reconocido previamente a ningún antígeno o molécula extraña (2,3). La presencia de estos receptores distintos en millones de linfocitos permite la protección frente a un número ilimitado de antígenos. De hecho, las posibles variantes solo para la cadena  $\beta$  del TCR (TCR $\beta$ ) es de  $10^{21}$ , siendo mucho mayores que el número de linfocitos del cuerpo humano que es del orden de  $10^{11}$  (3).

El TCR está codificado por tres loci, tanto en ratones como en humanos, los cuales codifican para cuatro cadenas (TCR $\alpha$ , TCR $\beta$ , TCR $\gamma$  y TCR $\delta$ ) que se unen de manera heterodimérica ( $\alpha\beta$  o  $\gamma\delta$ ) para dar lugar a los dos posibles TCR: TCR $\alpha\beta$  o TCR $\gamma\delta$ . Estos tres loci son: el locus *Tcrb* en ratón (o *TCRB* en humanos) que da lugar a la cadena TCR $\beta$ , el locus *Tcrd* en ratón (o *TCRD* en humanos) que da lugar a la cadena TCR $\delta$  y el locus *Tcrd* en ratón (o *TCRA/TCRD* en humanos) que da lugar a las cadenas TCR $\alpha$  y TCR $\delta$ . De esta forma los linfocitos T se dividen en dos grupos atendiendo a las cadenas que forman su TCR: linfocitos T  $\alpha\beta$  o linfocitos T  $\gamma\delta$  (4).

### 1.1 Desarrollo de los linfocitos T

Los linfocitos T se originan a partir de progenitores de médula ósea o hígado fetal que migran al timo para su maduración, selección y posterior exportación a la periferia (5). Estos progenitores son células multipotentes que en un principio tienen potencial mieloide y linfoide, dando lugar además de a linfocitos T, a distintos linajes celulares como células NK, células dendríticas, monocitos y granulocitos (6). La migración de estos progenitores al timo está mediada por la expresión de los receptores de quimioquinas CCR9 y CCR7 en estas células y por un ligando de la P-selectina que se expresa en el epitelio del timo. Una vez en el timo, el epitelio tímico favorece la diferenciación, la proliferación y la supervivencia de estos progenitores (7,8). El factor más importante en este proceso de diferenciación al linaje T es la expresión de los ligandos DELTA-LIKE (DL), DL1 y DL4, en el microambiente tímico que se unen a los receptores NOTCH expresados en los progenitores (6,7).

Durante su desarrollo en el timo (Figura 1), los timocitos pasan por una serie de estadios de desarrollo que se pueden distinguir por la expresión en superficie de los marcadores CD4 y CD8. De más temprano a más tardío se distinguen los siguientes estadios: timocitos CD4<sup>-</sup> CD8<sup>-</sup> doble negativos (DN), timocitos inmaduros simples positivos (*Immature Single Positive*, ISP) que son CD4<sup>-</sup> CD8<sup>+</sup> en ratones o CD4<sup>+</sup> CD8<sup>-</sup> en humanos, timocitos CD4<sup>+</sup> CD8<sup>+</sup> doble positivos (DP) y timocitos CD4<sup>+</sup> o CD8<sup>+</sup> simples positivos (SP) (4).



**Figura 1. Estadios de diferenciación de los timocitos.** Representación de los diferentes estadios de diferenciación en la maduración de los timocitos y los reordenamientos de los genes del TCR. El timo se representa como un rectángulo amarillo. Se indican las poblaciones de timocitos y linfocitos T. La selección  $\beta$ , la selección  $\gamma\delta$  y la selección positiva se representan con líneas rojas. La duración de las señalizaciones de NOTCH1 y del receptor de la interleucina 7 (IL-7R) se indica con una barra azul. Se muestran las dos olas de expresión de las proteínas de activación de la recombinación génica (RAG), RAG1 y RAG2, con un rectángulo morado. Imagen obtenida de Rodríguez-Caparrós, et al. (4).

Estos estadios se pueden subdividir en base a otros marcadores de membrana (Figura 1) (4). De esta forma, el estadio DN se subdivide en otros cuatro atendiendo a la presencia o no de los marcadores CD25 y CD44 en la membrana celular: DN1 (CD25<sup>-</sup> CD44<sup>+</sup>) que corresponden a los progenitores recién llegados al timo (*Early T-cell Progenitors*, ETP), DN2 (CD25<sup>+</sup> CD44<sup>+</sup>), DN3 (CD25<sup>+</sup> CD44<sup>-</sup>) y DN4 (CD25<sup>-</sup> CD44<sup>-</sup>). La población de timocitos DN2 se subdivide en DN2a y DN2b en función de la expresión de CD117 (KIT), pasando de alta a baja expresión de este marcador. Es en la transición de timocitos DN2a a DN2b cuando ocurre el compromiso definitivo al linaje de los linfocitos T mediado por la señalización de NOTCH1. La población de timocitos DN3 se subdivide en dos subpoblaciones atendiendo a la alta o baja expresión de CD27: CD27<sup>baja</sup> para la población DN3a y CD27<sup>alta</sup> para la población DN3b. Por otra parte, los timocitos DP también se dividen en dos

subpoblaciones en base a la expresión de CD71: timocitos DP tempranos (*early DP*, eDP) son CD71<sup>+</sup> y timocitos DP tardíos (*late DP*, lDP) son CD71<sup>-</sup> (9–11).

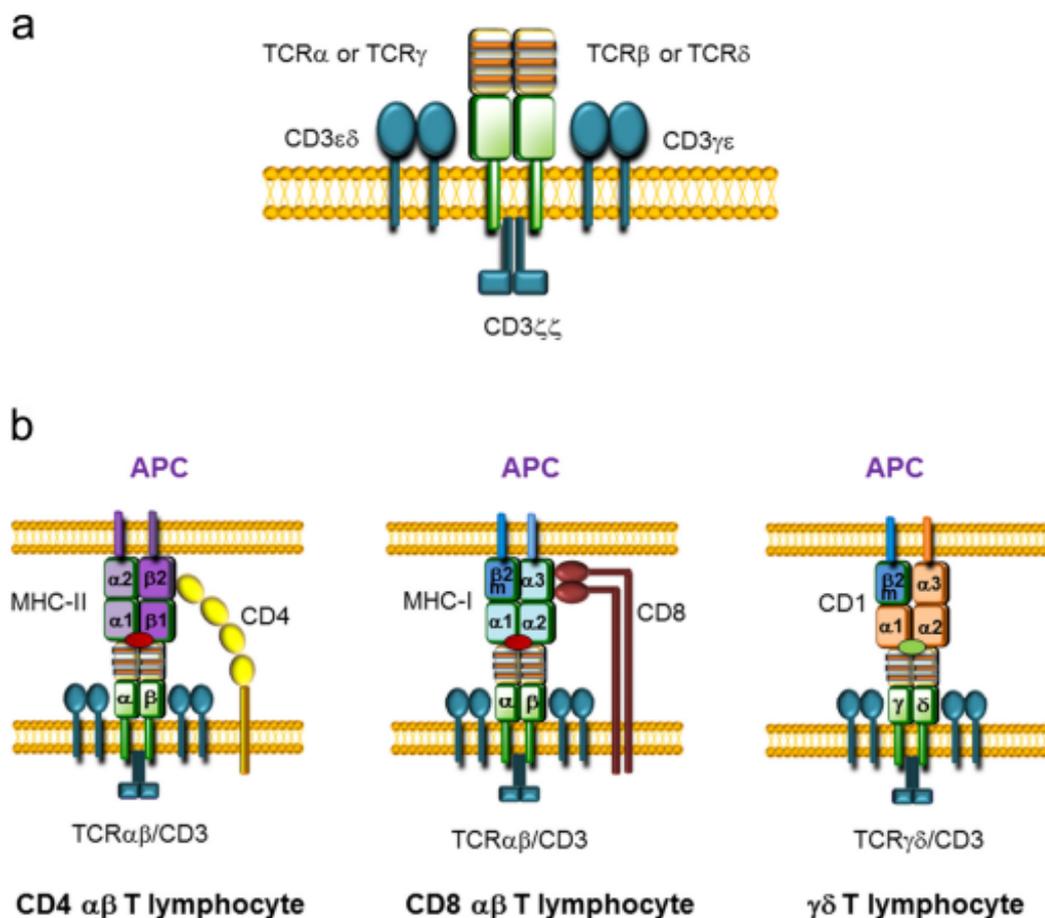
El tránsito a través de los diferentes estadios descritos ocurre durante el desarrollo de los linfocitos T  $\alpha\beta$ , mientras que para la generación de los linfocitos T  $\gamma\delta$ , los timocitos pasan solo por los estadios DN y no progresan al estadio DP (Figura 1) (4,12).

## 1.2. El receptor del linfocito T (TCR)

### 1.2.1. Estructura y función del TCR

El TCR está compuesto por un heterodímero polimórfico y glicosilado en el que las dos cadenas, TCR $\alpha$  y TCR $\beta$  o TCR $\gamma$  y TCR $\delta$ , se unen entre sí por un puente de cisteína cercano a la superficie celular (4). Este heterodímero está asociado a un complejo proteico denominado CD3 que se encuentra anclado a la membrana y que consiste en dos heterodímeros CD3 $\gamma\epsilon$  y CD3 $\delta\epsilon$ , junto con un homodímero de cadenas CD3 $\zeta\zeta$  (Figura 2). La función del complejo CD3 es esencial, puesto que se encarga de transmitir la señal derivada de la interacción del TCR con un antígeno al interior celular, ya que el TCR es incapaz de hacerlo por sí solo al tener una porción intracelular muy corta de tan solo de 2 a 7 aminoácidos (4,13).

El complejo proteico CD3 presenta un total de diez motivos de activación de inmunorreceptores basados en tirosina (*Immunoreceptor Tyrosine-based Activation Motif*, ITAM), uno en cada cadena CD3 $\epsilon$ , CD3 $\gamma$  y CD3 $\delta$ , y tres en cada cadena CD3 $\zeta$ , que al ser fosforilados por las quinasas LCK y FYN de la familia de las Src transmiten la señal desencadenada por el TCR, cuando éste reconoce un antígeno, al interior de la célula a través de la quinasa ZAP70, lo que da lugar a la activación de RAS y la proteína-quinasa C, y a la movilización de Ca<sup>2+</sup>, todo lo cual finalmente conduce a la activación final de determinados factores de transcripción inducibles (14).



**Figura 2. Representación de la estructura del complejo del TCR y de la interacción TCR-célula presentadora de antígeno (ACP).** (A) En color verde se identifica el heterodímero formado por las cadenas TCR $\alpha$  y TCR $\beta$  o TCR $\gamma$  y TCR $\delta$ . Las franjas naranjas dentro de cada cadena del TCR indican las regiones más hipervariables (regiones determinantes de la complementariedad, CDR). Junto al heterodímero TCR, en color azul se muestran las cadenas que forman el complejo CD3: dos heterodímeros, CD3 $\epsilon\delta$  y CD3 $\gamma\epsilon$ , y un homodímero, CD3 $\zeta\zeta$ , en la cara interna de la membrana celular. (B) Representación del TCR $\alpha\beta$ /CD8 o el TCR $\alpha\beta$ /CD4 presente en los linfocitos T  $\alpha\beta$  y el complejo mayor de histocompatibilidad (MHC) cargado con un péptido (óvalo rojo) presente en una APC. El correesceptor CD4 se representa en amarillo y el correesceptor CD8 se representa en marrón. El MHC de clase I (MHC-I) está formado por una cadena  $\alpha$  y una cadena microglobulina  $\beta_2$ , mientras que el MHC de clase II (MHC-II) está formado por una cadena  $\alpha$  y una cadena  $\beta$ . Los linfocitos T CD4<sup>+</sup> interactúan con el dominio  $\beta_2$  de la cadena  $\beta$  del MHC-II, mientras que los linfocitos T CD8<sup>+</sup> interactúan con el dominio  $\alpha_3$  de la cadena  $\alpha$  del MHC-I. También se representa la interacción de los linfocitos T  $\gamma\delta$  con la molécula CD1, formada por una cadena  $\alpha$  y una cadena  $\beta_{2m}$  cargada con un antígeno lipídico (óvalo verde), que también se expresa en una APC. Ilustración obtenida de Rodríguez-Caparros, et al. 2020 (4).

Las funciones del TCR $\alpha\beta$  y del TCR $\gamma\delta$  son diferentes, así como la forma en la que reconocen a los antígenos (Figura 2B). El TCR $\alpha\beta$  necesita que el complejo principal de histocompatibilidad (*Major Histocompatibility Complex*, MHC) le presente un antígeno procesado en forma de péptidos,

mientras que el TCR $\gamma\delta$  es capaz de reconocer otras estructuras no asociadas al MHC, tales como antígenos lipídicos presentados por la molécula CD1, aunque también se han identificado casos en los que pueden reconocer determinados péptidos asociados a moléculas de MHC (4,15–17). Entre los linfocitos T  $\alpha\beta$  se distinguen dos tipos, los pertenecientes al linaje CD8<sup>+</sup> (Tc, T citotóxicos) y los pertenecientes al linaje CD4<sup>+</sup> (Th, T *helper* o ayudadores). Los linfocitos Tc se encargan de la eliminación de células infectadas o tumorales, mientras que los linfocitos Th funcionan como coordinadores de la respuesta inmunitaria controlando la actividad de otras células inmunitarias, principalmente linfocitos y macrófagos. Los linfocitos Tc reconocen péptidos presentados por moléculas de MHC de clase I (MHC-I) presentes en prácticamente todas las células nucleadas, mientras que los linfocitos Th reconocen péptidos presentados a través de moléculas de MHC de clase II (MHC-II) presentes en las llamadas células presentadoras de antígeno (APC) (4,18).

En cada una de las cadenas que forman parte del heterodímero del TCR se pueden distinguir una región constante en el extremo carboxilo-terminal y una región variable en el extremo amino-terminal. La región variable, que se sitúa en el exterior celular, es la parte del receptor que reconoce a los antígenos (4) (Figura 2A). En la estructura de la región variable se pueden distinguir tres zonas hipervariables denominadas regiones determinantes de complementariedad (*Complementary Determinant Regions*, CDR) (19,20). La disposición en el espacio de estas regiones hipervariables conforma el sitio de reconocimiento del antígeno (21).

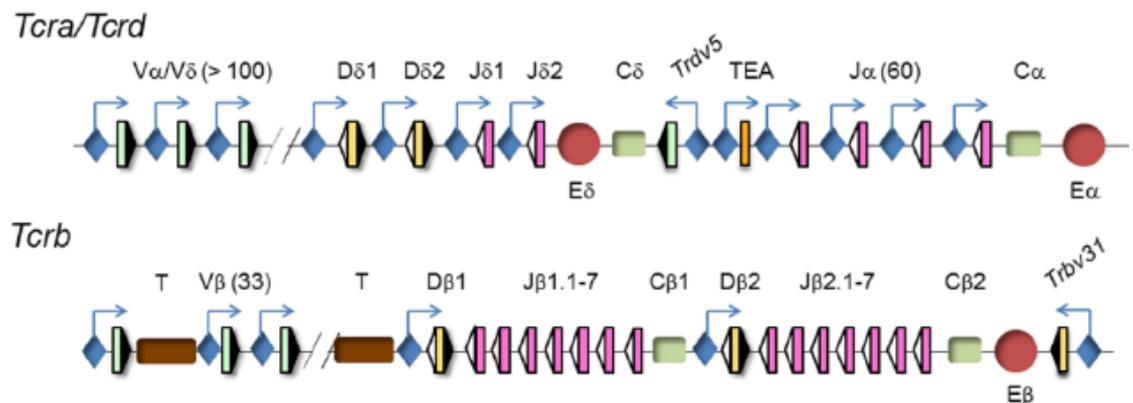
A nivel génico, las regiones hipervariables están formadas por la combinación de segmentos génicos que se denominan V (Variable), D (Diversidad) y J (*Joining* o unión) (4). Las regiones variables de las cadenas TCR $\alpha$  y TCR $\gamma$  están codificadas por la unión de segmentos génicos V y J, mientras que las de las cadenas TCR $\beta$  y TCR $\delta$  están codificadas por la unión de segmentos génicos V, D y J, tanto en ratones como en humanos (Figura 3). La recombinación de estos segmentos génicos durante la diferenciación de los timocitos es la que genera, mediante un proceso llamado “recombinación V(D)J”, una configuración génica capaz de dar lugar a un TCR $\alpha\beta$  o TCR $\gamma\delta$

funcional que es distinto en cada linfocito T (4,11,22). Este proceso se detalla más adelante el apartado 2 de la Introducción.

### 1.2.2. Estructura de los loci *Tcrb* y *Tcrα/Tcrδ*

Los segmentos génicos V(D)J están localizados en los loci que codifican para las distintas cadenas que componen el TCR, teniendo una distribución compleja y conservada en ratón y humano. Esta organización de segmentos génicos está también presente en los genes de las inmunoglobulinas y es responsable de la generación de la diversidad de los BCR y anticuerpos (4,21).

**Figura 3. Esquema de la estructura genómica de los loci *Tcrα/Tcrδ* y del *Tcrb*.** Los rectángulos verdes, amarillos y rosas representan a los segmentos génicos V, D y J respectivamente, y los



triángulos blancos o negros, las regiones señal de recombinación (RSS), diferenciando las secuencias espaciadoras de 12 o 23 pares de bases en color blanco o negro, respectivamente. Con rectángulos horizontales verdes se representan a las regiones constantes. Los promotores se representan con rombos azules y los potenciadores transcripcionales (*enhancers*) se muestran con círculos rojos. Ilustración obtenida de Rodríguez-Caparrós et al. 2020 (4).

El locus que codifica para la cadena TCRβ se encuentra en el cromosoma 6 de ratón y el cromosoma 7 en humanos, y tiene una longitud aproximada de 1 megabase (Mb) (4) (Figura 3). El locus *Tcrb* de ratón está compuesto por 35 segmentos Vβ, de los cuales 25 son funcionales, situados en la región 5' seguidos por dos subgrupos de segmentos Dβ y Jβ formados por 1 segmento Dβ y 6 segmentos Jβ funcionales. Cada subgrupo DβJβ está asociado a una región constante, Cβ1 o Cβ2, en la zona 3' del locus (4,21).

La estructura del locus que codifica para las cadenas TCRα y TCRδ es más compleja debido a que alberga dos genes en el mismo locus genético: el

gen de la cadena TCR $\alpha$  y el gen de la cadena TCR $\delta$  (Figura 3) (4). Este locus se encuentra en el cromosoma 14 de ratón y humano teniendo una longitud de 1,7 Mb en ratón y 1,1 Mb en humano; el mayor tamaño del locus de ratón se debe a que ha ocurrido un mayor número de duplicaciones de los segmentos V situados en la región 5', los cuales ocupan la mayor parte del locus. Entre estos segmentos V hay algunos que sólo se unirán con segmentos génicos D $\delta$ , otros que se unirán únicamente con segmentos J $\alpha$  y otros que pueden reordenar con ambos. Seguidamente en la región 3' del locus, ocupando las últimas 100 kilobases (Kb) aproximadamente, se encuentra el resto del gen de la cadena TCR $\delta$  con sus segmentos D $\delta$ , J $\delta$  y la región constante C $\delta$ , así como una región con 61 segmentos génicos J $\alpha$  y la región constante C $\alpha$  (23–25). El hecho que el gen de la cadena TCR $\delta$  se sitúe entre los segmentos V $\alpha$  y J $\alpha$  implica que cualquier reordenamiento del gen de la cadena TCR $\alpha$  conlleva la eliminación del gen de la cadena TCR $\delta$  (Figuras 3 y 4), tanto en humanos como en ratones (4). Debido a esta organización, la regulación temporal de los reordenamientos en este locus es fundamental para permitir la generación de linfocitos T  $\gamma\delta$ , de forma que los reordenamientos V $\delta$ D $\delta$ J $\delta$  han de completarse antes de que ocurran los primeros reordenamientos V $\alpha$ J $\alpha$  durante el desarrollo de los timocitos, debido a que los reordenamientos del gen TCR $\alpha$  provocan la delección del gen TCR $\delta$  y su subsiguiente silenciamiento génico (23,26).

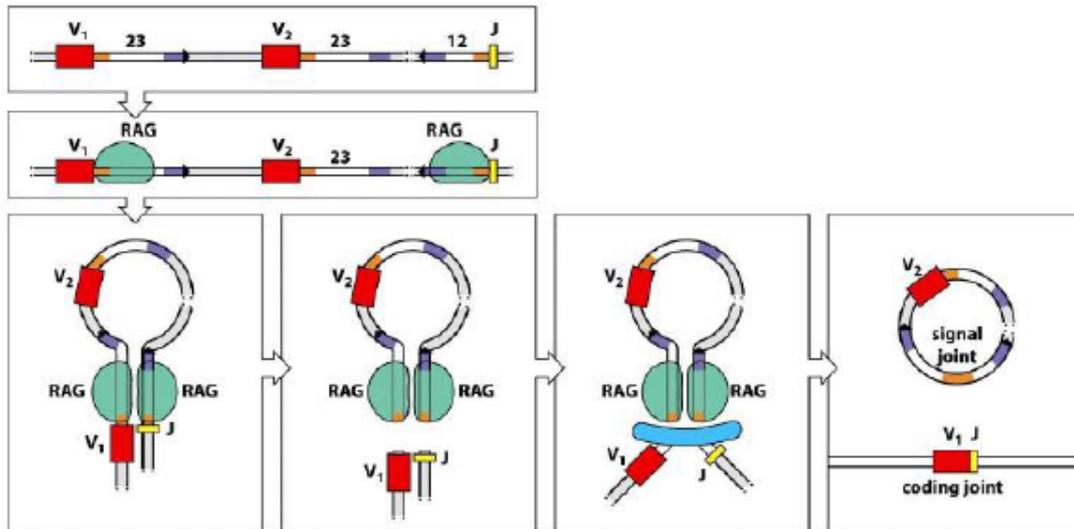
## 2. La recombinación V(D)J

La recombinación V(D)J es el mecanismo por el cual se ensamblan los diferentes segmentos génicos VJ o VDJ de los genes codificantes de las cadenas que forman el BCR y el TCR (2). Una vez finaliza el proceso de recombinación, la estructura del gen queda modificada, permitiendo la expresión de una cadena funcional. Se trata de un proceso que permite que con una cantidad relativamente pequeña de ADN se consiga un gran número de receptores distintos, lo cual permite el reconocimiento específico de múltiples antígenos. Por lo tanto, la recombinación V(D)J es el mecanismo en el que se basa la diversidad del reconocimiento antigénico que caracteriza al sistema inmunitario adaptativo.

Unos elementos esenciales para el inicio de la recombinación V(D)J son las secuencias que están presentes al inicio y/o final de cada segmento génico V, D y J, denominadas secuencias señal de recombinación (*Recombination Signal Sequences*, RSS) (27) (Figuras 3-5). Las RSS son necesarias para que se unan unas proteínas específicas de linfocitos inmaduros denominadas RAG1 y RAG2 (*Recombination Activation Gene*, RAG) que son responsables del proceso. Cada RSS está formada por un heptámero y un nonámero separados por una secuencia de 12 o 23 pares de bases (pb) (Figura 5). Para que se pueda dar la recombinación entre dos segmentos génicos es necesario que ocurra entre un segmento que tenga una RSS con un espaciador de 12 pb y un segmento que tenga una RSS con un espaciador de 23 pb, condición que da nombre a la denominada “regla 12/23” (28,29). Sin embargo, esta restricción no es suficiente para explicar la especificidad de los reordenamientos de los segmentos V $\beta$  con los segmentos D $\beta$  y no con los segmentos J $\beta$ . Esta excepción de la regla 12/23 viene dada por la presencia de determinadas secuencias dentro de las regiones espaciadoras de las RSS que son capaces de restringir las recombinaciones V $\beta$ -J $\beta$  (30,31). Además de la regla 12/23, existen otros niveles de regulación para la recombinación V(D)J. Uno de ellos deriva de la transcripción no codificante de los genes no reordenados (transcripción germinal), ya que la trimetilación de la lisina K4 de la histona H3 (H3K4me3) derivada de la transcripción activa favorece la unión de RAG2 y la actividad catalítica de RAG1 (32). Otro nivel de regulación depende de la localización intracelular de los loci respecto a la de RAG2 y la ARN polimerasa 2 (33).

Una vez que RAG1 y RAG2 se han unido a las RSS, estas proteínas se unen entre sí formando una endonucleasa que corta la doble cadena de ADN entre la RSS y la región codificante de dos segmentos génicos compatibles, dando como resultado la formación de cuatro extremos, dos pertenecientes a los segmentos V, D, o J y otros dos pertenecientes a las RSS (4) (Figura 5). A continuación, se unen los dos segmentos codificantes y las dos RSS entre sí. Aunque RAG1 y RAG2 son suficientes para la escisión *in vitro* del sustrato, las proteínas del grupo de alta movilidad (*High Mobility Group B*, HMGB), HMGB1 o HMGB2 se unen de forma inespecífica al ADN e induciendo su flexión, lo cual facilita el ensamblaje de RAG1 y RAG2 a las RSS (34,35).

Los cortes en la doble cadena del ADN inducidos por RAG1 y RAG2 son reparados por la maquinaria responsable de la reparación del ADN de segmentos no homólogos (*Non Homologous End Joining*, NHEJ) (39). Para las uniones codificantes entre dos segmentos génicos, en primer lugar, se origina una ligación entre las dos hebras del ADN en cada extremo, formándose una horquilla. Esta horquilla se abre de forma asimétrica por acción de la enzima ARTEMISA, rellenándose los extremos protuberantes que se originen con nucleótidos complementarios o “nucleótidos P” para dar lugar a extremos romos. En este punto actúa la enzima deoxinucleotidiltransferasa terminal (*Terminal deoxynucleotidyl-Transferase*, TdT) que añade nucleótidos al azar o “nucleótidos N” antes de que se produzca la unión definitiva de los dos segmentos génicos, lo cual aumenta aún más la variabilidad en las uniones resultantes en este proceso. A continuación ocurre la ligación de los extremos, tanto codificantes entre los segmentos génicos como no codificantes entre las RSS, mediante la unión de un heterodímero Ku70/Ku80 a los extremos del ADN, lo cual recluta a la subunidad catalítica de la proteína quinasa dependiente de ADN (DNA-PK) para formar una holoenzima que fosforila múltiples sustratos y promueve el empalme entre los extremos de ADN (38,39). Para que se produzca la unión final de los extremos cortados actúa la proteína XRCC4 que se une y activa a la ligasa IV. Como resultado se forma una unión codificante VJ, VD o DJ, que en caso de ser productiva permitirá la formación de una cadena del TCR o BCR, y otra unión no codificante entre las dos RSS que dará lugar a un anillo extracromosómico que desaparecerá por dilución durante la expansión clonal (2,40).



**Figura 4. Esquema del proceso de recombinación V(D)J.** En rojo se representan los segmentos génicos V y en amarillo los segmentos génicos J. Las proteínas RAG se muestran en verde. Las RSS aparecen como un espacio blanco de distinta longitud atendiendo a si son de 12 o 23 pb flanqueado por un rectángulo naranja y otro morado, que representan el heptámero y el nonámero. Ilustración obtenida de “*The Immune System*, 2Ed. (Garland Science 2005).

## 2.1. Regulación de la recombinación V(D)J

El conocimiento actual sobre los mecanismos moleculares y celulares que regulan el proceso de recombinación V(D)J procede principalmente de los estudios realizados en modelos de ratón, habiéndose comprobado que este proceso está conservado en humanos.

Durante este proceso existen 4 puntos de control (2): 1º) control del linaje celular, según el cual la recombinación de ciertos genes ocurre exclusivamente en determinados precursores de linfocitos, por ejemplo, el gen *Igh* solo reordena en los precursores de linfocitos B mientras que el gen *Tcra* solo reordena en precursores de linfocitos T; 2º) control temporal, según el cual se establece el orden en que los diferentes genes reordenan, por ejemplo, los genes *Tcrb*, *Tcrg* y *Tcrd* reordenan en los estadios DN2 y DN3a, antes que lo haga el gen *Tcra* en los estadios DN4 y DP; 3º) control en el orden en el que se reordenan los diferentes segmentos génicos, según el cual determinados genes que contienen segmentos V, D y J, como es el caso de los genes *Tcrb* e *Igh*, reordenan primero los segmentos D con los segmentos J y posteriormente se une el segmento V al segmento D previamente reordenados; y 4º) control

de la exclusión alélica, según el cual solo uno de los dos alelos se reordena en un momento dado, como ocurre en el caso de los genes *Tcrb*, *Igh*, *Igk* e *Igl*; aunque no hay exclusión alélica en la recombinación V(D)J de los demás genes, existe una restricción de expresión alélica determinada por la afinidad para el emparejamiento de las cadenas expresadas (4,41).

Estos niveles de regulación fueron postulados en 1985 por Yancopoulos y Alt mediante el denominado “modelo de accesibilidad” (27,42), según el cual la estructura de la cromatina, la cual es específica para cada gen de los receptores de antígeno y que varía a lo largo del desarrollo de los precursores de los linfocitos, determina la unión de RAG1 y RAG2 a las RSS y, por tanto, el comienzo del proceso de recombinación V(D)J. La validez de este modelo ha sido demostrada experimentalmente y está ya totalmente aceptada (41). La apertura de la cromatina que permite la unión de RAG1 y RAG2 depende de una serie de marcas epigenéticas como son la hipometilación del ADN y determinadas modificaciones covalentes en las histonas, que en última instancia dependen de la activación de elementos *cis* reguladores de la transcripción, concretamente promotores y potenciadores transcripcionales (*enhancers*), además de la posición de los loci a nivel subnuclear (4,27).

## 2.2. Transcripción germinal y recombinación V(D)J

Dentro de los genes que codifican para el TCR y el BCR, hay dos tipos de elementos reguladores de la transcripción que son determinantes para la estructura de la cromatina y, por tanto, para el proceso de recombinación V(D)J durante el desarrollo de las células: 1) los promotores asociados a los segmentos génicos V, D y J, y 2) los *enhancers*, que se sitúan en las zonas próximas a las regiones constantes (4,43).

En este punto, cabe aclarar que en estos genes se pueden distinguir dos tipos de transcripción (4). 1) Una transcripción inicial no codificante, denominada también transcripción germinal, que es previa a la recombinación V(D)J y se debe a la activación de los promotores de los segmentos génicos V, D y J por los *enhancers*; esta transcripción es responsable de la apertura de la cromatina y permite el reconocimiento de las RSS por RAG1 y RAG2. 2)

Después del reordenamiento V(D)J se produce otra transcripción que se inicia en el promotor del segmento V reordenado; esta transcripción da lugar a un ARN mensajero que será codificante o no dependiendo de si el reordenamiento ha sido productivo o no, es decir, si se mantiene el correcto marco de lectura impuesto por la secuencia del segmento V reordenado.

En los experimentos pioneros de Yancopoulos y Alt (42) se observó que la activación de la recombinación en los segmentos  $V_H$  del locus *Igh* coincidía con la detección de transcritos germinales iniciados en los segmentos  $V_H$ . Posteriormente la asociación entre la transcripción germinal y el inicio de la recombinación V(D)J se ha demostrado en todos los genes que codifican a los receptores de antígeno (44–46). Está demostrado que los promotores y *enhancers* activan la transcripción germinal y el reclutamiento de RAG1 y RAG2 a las RSS, siendo especialmente importante este proceso en el reordenamiento de los segmentos génicos presentes en loci largos donde la acción de los promotores necesita extenderse a largas distancias (47).

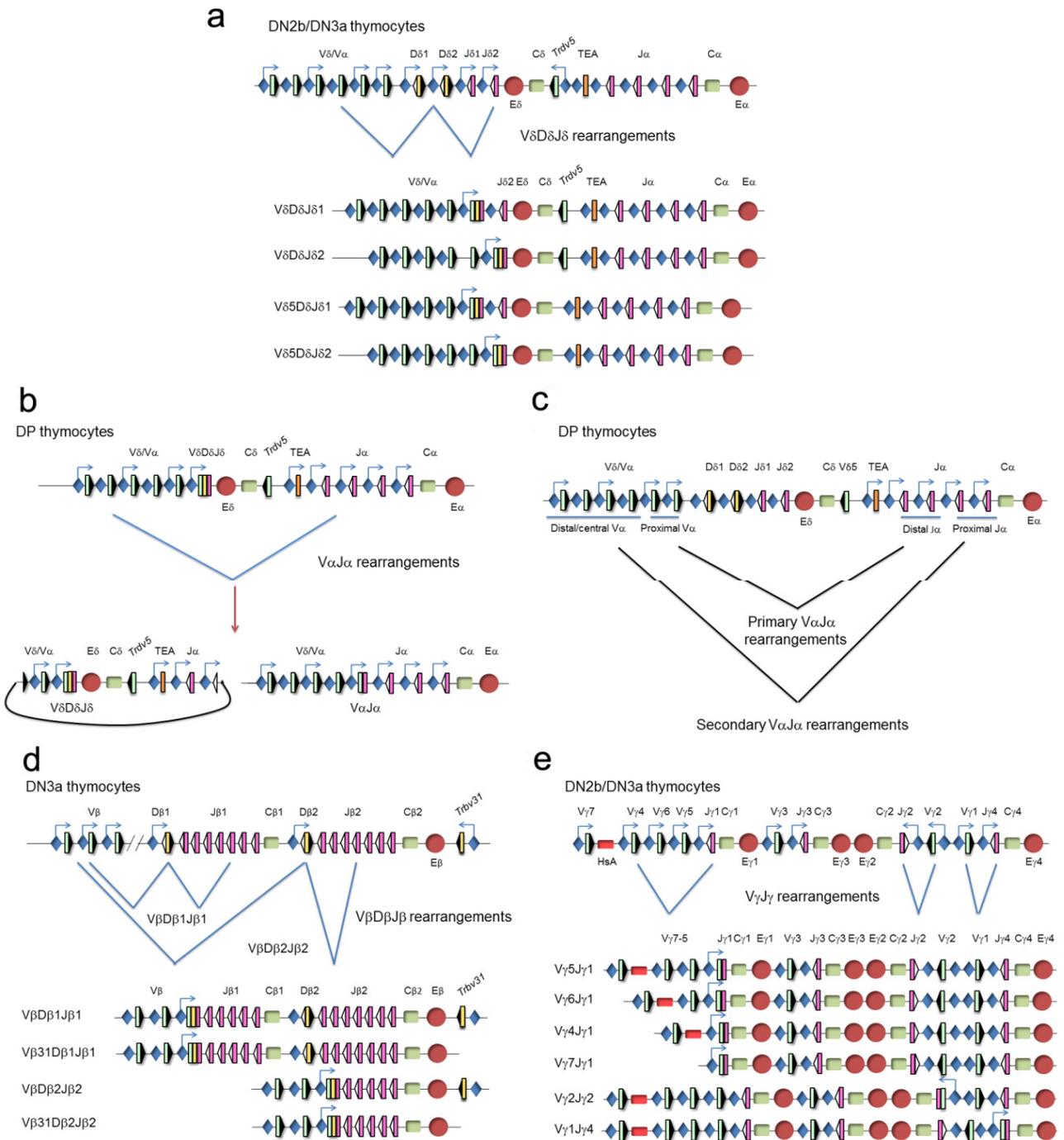
Aunque la relación entre la transcripción no codificante germinal y la recombinación V(D)J es general, se ha encontrado algún caso muy concreto en el que la recombinación V(D)J y la transcripción germinal no están totalmente asociados, de forma que la transcripción germinal no es suficiente para activar la recombinación V(D)J. Por ejemplo, los segmentos génicos  $V_H$  distales tienen altos niveles de transcripción en las células pro-B de ratones *Pax5*<sup>-/-</sup> y, sin embargo no recombinan (48). En este caso concreto, aunque hay transcripción germinal y la cromatina tiene una configuración adecuada para la recombinación, la ausencia del factor de transcripción PAX5 impide su recombinación. PAX5 es esencial para la recombinación del locus *Igh*, haciendo que la cohesina permanezca más tiempo unida a la cromatina y facilitando su extrusión y el acercamiento de los elementos  $V_H$  distales (48). Sin embargo, no se ha demostrado de forma clara la situación contraria, esto es la activación de la recombinación V(D)J en ausencia total de transcripción germinal. Por lo tanto, se puede concluir que la transcripción germinal y la accesibilidad a la cromatina son sucesos necesarios, pero no suficientes para que ocurra la recombinación V(D)J. La identificación de determinados casos

en los que estos procesos no están acoplados indica la existencia de un tipo de regulación de la recombinación V(D)J que va “más allá” de la accesibilidad.

### 2.3. Recombinación V(D)J de los genes del TCR durante el desarrollo de los linfocitos T

La regulación de la recombinación V(D)J durante el desarrollo se ha estudiado de forma detallada en ratones. El compromiso de los progenitores tímicos al linaje de los linfocitos T no ocurre hasta la transición de timocitos DN2a a DN2b, dependiendo de la señalización mediada por NOTCH1 (4). A lo largo de su maduración, los timocitos reordenan sus genes del TCR de una forma programada para finalmente dar lugar a los linfocitos T  $\alpha\beta$  o  $\gamma\delta$  (Figuras 1 y 5). En timocitos DN2b se detectan los primeros reordenamientos en los genes *Tcr $\gamma$* , *Tcr $\delta$*  y *Tcr $\beta$*  derivados de una primera oleada de expresión de RAG1 y RAG2 que se observa en timocitos DN2b y DN3a (49). Para el caso de los genes *Tcr $\gamma$*  y *Tcr $\delta$* , la recombinación  $V\gamma J\gamma$  y  $V\delta D\delta J\delta$  está completa en los estadios DN2b y DN3a, mientras que para el gen *Tcr $\beta$*  la recombinación  $V\beta D\beta J\beta$  se completa en timocitos DN3a. El gen *Tcr $\alpha$*  no comienza a reordenar hasta el estadio DN4, siendo estos reordenamientos  $V\alpha J\alpha$  muy abundantes en timocitos DP, coincidiendo con un segundo pico de expresión de RAG1 y RAG2 (49). La bajada de la expresión de RAG1 y RAG2 en la transición de timocitos DN3a a DP se debe a la inhibición de la transcripción de los genes *Rag1* y *Rag2* durante la proliferación celular que acompaña a este proceso (49,50).

La recombinación productiva de los genes *Tcr $\gamma$*  ( $V\gamma J\gamma$ ) y *Tcr $\delta$*  ( $V\delta D\delta J\delta$ ) en los estadios DN2b y DN3a resulta en la expresión de un TCR $\gamma\delta$  funcional, lo cual permite que estos timocitos pasen por un proceso denominado selección  $\gamma\delta$  (Figura 1), basado en la expresión de este receptor, antes de salir a la periferia como linfocitos T  $\gamma\delta$  maduros (4).



**Figura 5. Representación de los loci del TCR sin reordenar y reordenados durante el desarrollo de los timocitos.** Representación de los reordenamientos de los genes **(a) *Tcrd***, **(b, c) *Tera***, **(d) *Tcrb***, y **(e) *Tcrg***. Los segmentos génicos variable (V), diversidad (D) y unión (J) se representan con rectángulos verdes, amarillos y rosas respectivamente. Las secuencias señal de recombinación (RSS) se representan con triángulos adyacentes a los segmentos V, D y J, siendo negro si la RSS es de 12 pares de bases y blanco si es de 23 pares de bases. Las regiones constantes (C) se representan con rectángulos horizontales verdes. El exón *early*  $\alpha$  (TEA) se representa con un rectángulo naranja. Los genes del tripsinógeno (T) se representan con rectángulos marrones. Los *enhancers* se representan con círculos rojos y los promotores con rombos azules. La transcripción germinal está representada con flechas azules. En el gen *Tcrd* **(a)** los reordenamientos V $\delta$ D $\delta$ J $\delta$  ocurren en timocitos DN2b y DN3a mientras que

los reordenamientos  $V\alpha J\alpha$  del gen *Tcra* ocurren en timocitos DP (**b**, **c**). Como consecuencia de los reordenamientos primarios  $V\alpha J\alpha$  (**c**), el gen *Tcrd* se deletiona del genoma en un círculo extracromosómico (**b**). En el locus *Tcrb* (**d**), los reordenamientos  $V\beta D\beta J\beta$  suceden en timocitos DN3a. En el locus *Tcrg* (**e**), los reordenamientos  $V\gamma J\gamma$  ocurren en timocitos DN2b y DN3a. Ilustración obtenida de Rodríguez-Caparrós et. Al 2020 (4).

En el proceso de maduración de timocitos al linaje de linfocitos T  $\alpha\beta$ , uno de los puntos más importantes es la recombinación funcional  $V\beta D\beta J\beta$  del locus *Tcrb* en el estadio DN3a (4) (Figura 5). Al expresarse una cadena TCR $\beta$  en estas células, ésta se ensambla con una cadena temporal invariable denominada pre-T $\alpha$ , formando un receptor de membrana denominado pre-TCR, el cual no reconoce ningún ligando específico (51). La expresión de un pre-TCR permite al timocito DN3a continuar con su diferenciación hasta dar lugar a timocitos DP mediante un proceso denominado selección  $\beta$ , llamado así porque depende de la expresión de la cadena TCR $\beta$  (Figura 1). Posteriormente, la recombinación funcional  $V\alpha J\alpha$  del gen *Tcra* en timocitos DN4 y sobre todo en timocitos DP conducirá a la expresión de una cadena TCR $\alpha$  productiva y a su ensamblaje con la cadena TCR $\beta$  previamente expresada, formando de esta manera el TCR $\alpha\beta$ . La selección  $\beta$ , además de permitir el desarrollo de timocitos hacia linfocitos T  $\alpha\beta$ , inhibe de forma estable la expresión de las cadenas TCR $\gamma$  y TCR $\delta$ , evitando que dicha expresión pueda interferir en el ensamblaje de las cadenas TCR $\alpha$  y TCR $\beta$  para dar lugar a un TCR $\alpha\beta$  (12,52,53).

Es interesante que la recombinación de los genes *Tcrg*, *Tcrd* y *Tcrb* ocurra en precursores potenciales tanto para linfocitos T  $\gamma\delta$  como  $\alpha\beta$ . Está aceptado que la fuerza de señalización del TCR $\gamma\delta$  y pre-TCR determina el destino de estos linajes de una forma instructiva (54–56). La señal fuerte derivada de la expresión del TCR $\gamma\delta$  determina la diferenciación de los precursores al linaje T  $\gamma\delta$ , mientras que la señal débil derivada de la expresión del pre-TCR determina la diferenciación de los precursores al linaje  $\alpha\beta$ .

Debido a su estructura, y a diferencia de los otros genes del TCR (Figuras 3 y 5) (4), el gen *Tcra* puede sufrir sucesivos reordenamientos  $V\alpha J\alpha$ , desde el estadio DN4 hasta el estadio IDP para garantizar la consecución de una cadena TCR $\alpha$  funcional (57). De esta forma, los reordenamientos que no

sean productivos o den lugar a una cadena TCR $\alpha$  que no pueda ensamblarse con la cadena TCR $\beta$  previamente expresada serán eliminados y reemplazados por otro reordenamiento (58). Esto es posible debido a la estructura de este gen y a la forma en la que ocurren los reordenamientos en el mismo, distinguiéndose dos tipos de reordenamientos: primarios y secundarios (4) (Figura 5). Los reordenamientos primarios son los primeros y ocurren entre los segmentos V $\alpha$  más proximales y los J $\alpha$  más distales, tomando como referencia la región constante C $\alpha$ , en timocitos DN4 y eDP (58). Para que se den lugar estos reordenamientos se necesita la activación de dos promotores localizados en la región más 5' que contiene a segmentos J $\alpha$ : el promotor asociado a un exón no codificante denominado TEA (*T Early Alpha*, TEAp) y el promotor asociado al segmento génico *Traj49* (J $\alpha$ 49p); la activación de estos promotores depende de la función del *enhancer* de *Tera* (E $\alpha$ ) (59). En el caso de que el reordenamiento primario no haya sido productivo, se irán sucediendo reordenamientos secundarios de segmentos V $\alpha$  situados cada vez más distalmente con segmentos J $\alpha$  situados más proximalmente respecto a la posición de C $\alpha$  en timocitos IDP (55). La activación de los promotores TEAp y J $\alpha$ 49p en timocitos DN4 y eDP depende de la activación previa de E $\alpha$ , que a su vez es consecuencia de la señalización desencadenada por el pre-TCR (60). Las recombinaciones V $\delta$ D $\delta$ J $\delta$  que han ocurrido previamente en cada alelo aumentan aún más la diversidad de las cadenas TCR $\alpha$ , ya que determinan el punto de inicio del uso de los segmentos V $\alpha$  en las recombinaciones V $\alpha$ J $\alpha$  (61).

Antes de salir como linfocitos T  $\alpha\beta$ , los timocitos IDP han de pasar por otros procesos de selección. Aquellos timocitos que no reconocen un MHC propio son eliminados por apoptosis en un proceso denominado negligencia, aquellos timocitos que reconocen péptidos propios con alta afinidad en el contexto de un MHC propios son eliminados por apoptosis en un proceso denominado selección negativa para evitar autoinmunidad celular, y solo aquellos timocitos que reconocen péptidos propios presentados por el MHC propio con una afinidad baja o media sobreviven en un proceso denominado selección positiva (4) (Figura 1). Los timocitos IDP que superan estos procesos de selección constituyen el 3-5% del total de timocitos DP y diferencian a

timocitos SP CD4<sup>+</sup> o CD8<sup>+</sup>, los cuales finalmente abandonan el timo y salen a la periferia como linfocitos T  $\alpha\beta$  (4,5,44).

Durante todo el proceso de recombinación V(D)J de los genes del TCR son muy importantes las señalizaciones recibidas por los timocitos durante su diferenciación, tales como las mediadas por NOTCH1, el receptor de interleuquina 7 (IL-7R) y el pre-TCR (4) (Figura 1). Las señalizaciones mediadas por los receptores NOTCH1 e IL-7R son predominantes hasta el estadio DN3a, momento en el que se expresa el pre-TCR; después, en la transición de timocitos DN3a a DP (selección  $\beta$ ), predomina la señalización mediada por el pre-TCR (4,10). La señalización mediada por NOTCH1 es esencial para la expresión del pre-TCR, tanto de la cadena pre-T $\alpha$  como de la cadena TCR $\beta$ , a través de la activación de la transcripción del gen *Ptcra* que codifica para la cadena pre-T $\alpha$  y de la recombinación del locus *Tcrb* (63,64). Además esta señalización está implicada en la transcripción germinal de los genes *Tcrd* y *Tcrg* a través de sus *enhancers* (12,65).

La señalización mediada por IL-7R es clave para la maduración de los timocitos. para la transcripción y recombinación del gen *Tcrg* y para la transcripción germinal del gen *Tcrd*, por lo tanto, para la generación de los linfocitos T  $\gamma\delta$  (62,65,66). La señalización mediada por el pre-TCR activa la transcripción y recombinación del gen *Tcra* e inhibe la señalización de NOTCH1, y por tanto, la transcripción de los genes *Tcrd* y *Tcrg* (12,41,60,67). Además, la señalización de NOTCH1 activa la transcripción del gen de la cadena de IL-7R (68,69). Por tanto, al inhibirse la señalización de NOTCH1 por la expresión del pre-TCR (67), también se inhibe la señalización de IL-7R. De esta manera, la señalización de NOTCH1 baja bruscamente en el paso de DN3a a DN3b (10), y como consecuencia desaparece la señalización de IL-7R en el paso de DN4 a DP (62).

### 3. Organización espacial del genoma

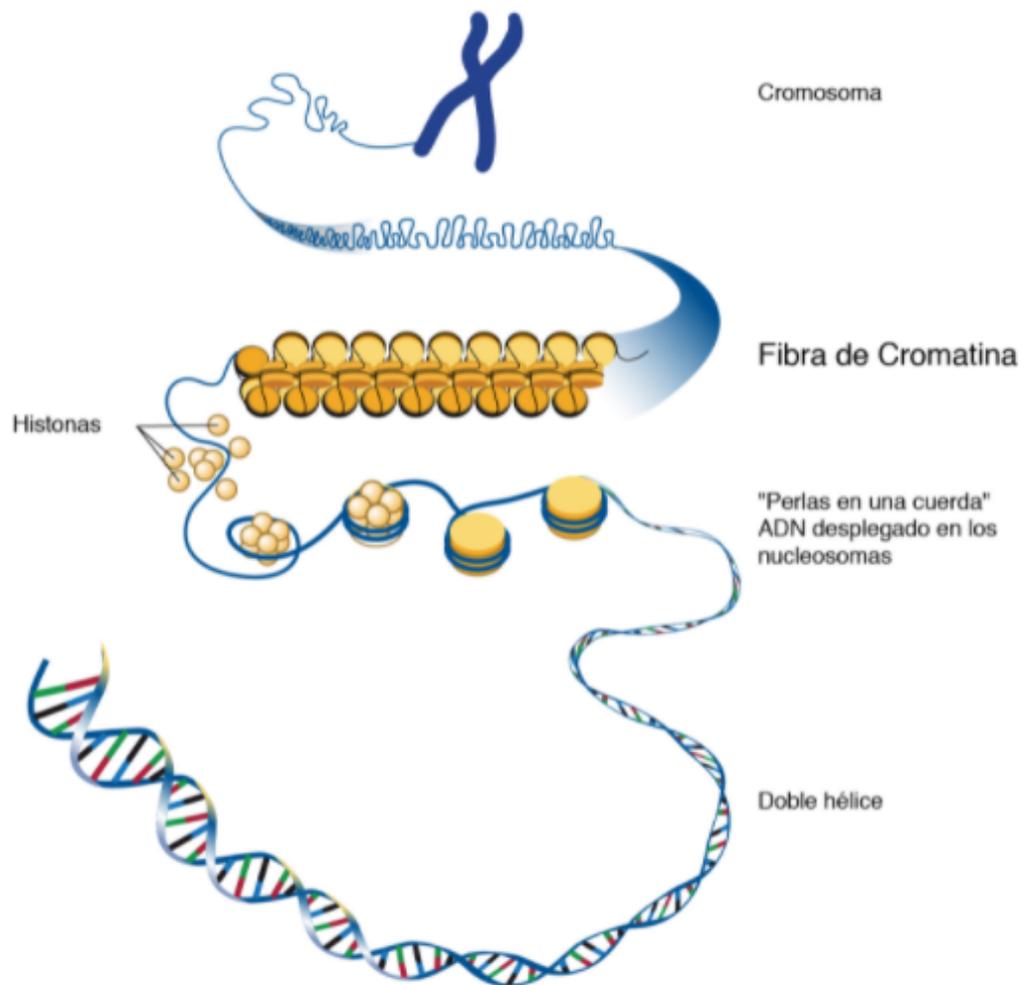
#### 3.1. Estructura de la cromatina

El término cromatina fue acuñado por Walther Fleming en 1879 para denominar a estructuras fibrosas teñidas en el núcleo (70). La cromatina es

una estructura compacta formada por ADN asociado a proteínas, la cual permite alojar 2 metros de ADN dentro del núcleo celular. Estas proteínas son de dos tipos, proteínas básicas de bajo molecular denominadas histonas y proteínas no básicas. Dentro de las histonas se diferencian cinco distintas, denominadas H1, H2A, H2B, H3 y H4. Dentro de las proteínas no básicas se encuentran enzimas implicadas en la replicación, transcripción y regulación del ADN (71). El empaquetamiento del ADN se realiza alrededor de una estructura formada por ocho unidades de histonas (dos moléculas de cada una de las histonas H2A, H2B, H3 y H4), constituyendo la unidad básica de empaquetamiento, denominada nucleosoma (Figura 6). Los nucleosomas se unen entre sí mediante ADN unido a la histona H1.

Los sucesivos empaquetamientos de esta estructura son los responsables de los diferentes aspectos del ADN vistos al microscopio, así como de la diferente accesibilidad de la maquinaria de transcripción (72,73). Se pueden diferenciar bajo el microscopio dos tipos de cromatina, la eucromatina, que es la más abundante en la interfase y se corresponde con la cromatina menos compactada y, por lo tanto, más accesible a la transcripción, y la heterocromatina que tiene un mayor grado de compactación. La heterocromatina puede aparecer en dos formas: la heterocromatina constitutiva, la cual corresponde a un ADN muy condensado que no se transcribe y está presente en zonas concretas en todas las células, y la heterocromatina facultativa, la cual se encuentra en zonas diferentes en distintas células e incluye a un conjunto de genes que se inactivan o activan de forma específica en cada estirpe celular durante la diferenciación.

La estructura de la cromatina es dinámica, cambiando de un estado de empaquetamiento máximo durante la mitosis a un estado más desempaquetado durante la interfase, permitiendo el acceso al ADN de diferentes complejos de proteínas para realizar los diferentes procesos moleculares (72,73). Por ejemplo, los nucleosomas pueden “aflojarse” permitiendo la disposición de nuevos octámeros tras la replicación del ADN o para “abrir” elementos reguladores, como los promotores y *enhancers*, facilitando el acceso a factores o complejos de transcripción que se unen al ADN.



**Figura 6. Esquema representativo de los diferentes grados de compactación del ADN en el núcleo celular.** De abajo hacia arriba se representa el paso de un estado desempaquetado del ADN visto como una doble hélice hasta el nivel más compacto que corresponde con el cromosoma. En la parte intermedia de la figura se muestra como el ADN interacciona con las histonas para conseguir estados con distinto nivel de compactación. Ilustración obtenida de *National Human Genome Research Institute* (<https://www.genome.gov/es/genetics-glossary/Cromatina>).

Los extremos amino-terminales de las histonas que conforman los nucleosomas interaccionan con los de otras histonas adyacentes y con el ADN. Existen enzimas que modifican estos extremos amino-terminales e interfieren en su interacción con el ADN, permitiendo que el ADN esté más accesible para su interacción con otras proteínas (74). La primera modificación de histonas descubierta fue la acetilación en 1.964 (75). Desde entonces se han identificado de más de 100 modificaciones postraduccionales distintas de las histonas. Las

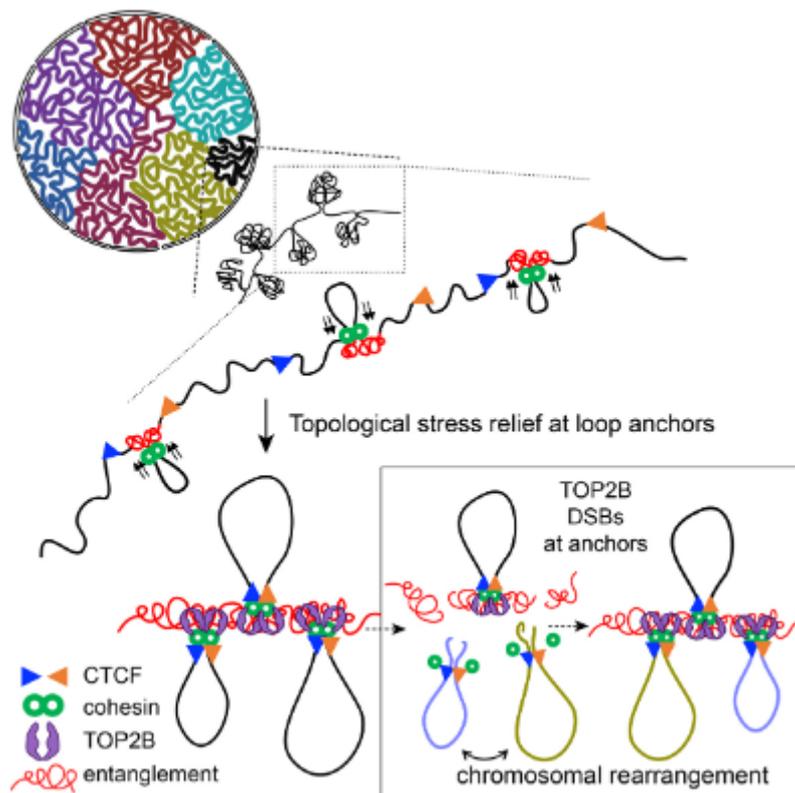
principales modificaciones postraduccionales de los extremos amino-terminales de las histonas se clasifican en 7 tipos: acetilación, metilación, fosforilación, adenosina-difosfato (ADP)-ribosilación, glicosilación, sumoilación y ubiquitilación. Cada tipo tiene una función distinta y un mecanismo regulador (74). Por ejemplo, la acetilación, proceso que se da en los residuos de lisina, neutraliza la carga positiva de las histonas promoviendo su desempaquetamiento y favoreciendo los procesos que necesitan tener una cromatina accesible (76). La metilación de histonas, por su parte, ocurre también en residuos de lisina (desde monometilación hasta trimetilación) o de arginina (solo hasta dimetilación), en este caso interfiriendo con factores de unión a la cromatina pudiendo favorecer la transcripción, como en los casos de la H3K4me y la metilación de la lisina 6 de la histona H3 (H3K6me), o reducir la transcripción, como en los casos de la metilación de la lisina 9 y la lisina 27 de la histona H3 (H3K9me y H3K27me) (77,78).

En relación con lo establecido previamente, se ha demostrado que las marcas H3K4me<sub>3</sub> y la hiperacetilación de las histonas H3 y H4 se asocian con loci que se encuentran en recombinación V(D)J activa (4,79,80). De hecho, RAG2 posee un dominio de dedo PHD (*Plant HomeoDomain*) que reconoce de forma específica la marca H3K4me<sub>3</sub> y estimula la actividad catalítica de la proteína, siendo este reconocimiento esencial para la recombinación V(D)J (81,82). La ausencia del dedo PHD de RAG2 causa un tipo de inmunodeficiencia combinada severa (*Severe Combined ImmunoDeficiency*, SCID) o el síndrome de Ommen (una forma leve de SCID). Desde el punto de vista clínico, alteraciones de la marca H3K4me<sub>3</sub> en determinados puntos críticos pueden llevar a una actividad anómala del complejo RAG1/RAG2 y dar lugar a tumores linfoides malignos (79).

### 3.2. Dominios asociados topológicamente (TAD)

El genoma está compartimentalizado en dominios de alrededor de 1 Mb denominados dominios asociados topológicamente (*Topologically Associated Domains*, TAD) (Figura 7), que se caracterizan porque las regiones dentro del mismo dominio tienen interacciones físicas entre sí más frecuentes que con

otras regiones fuera del mismo (83,84). Estos TAD están delimitados por regiones enriquecidas en genes de expresión constitutiva (*housekeeping*) o por sitios de unión para el factor de unión a la secuencia CCCTC (*CCCTC binding factor*, CTCF) (*CTCF Binding Element*, CBE). CTCF es un factor ampliamente distribuido con funciones tanto en activación como en represión transcripcional, con capacidad para aislar un promotor de la influencia de un *enhancer* y para aislar dos TAD entre sí.



**Figura 7. Esquema representativo de la formación de un lazo de cromatina mediado por interacciones entre CTCF y cohesina.** Se forman lazos de cromatina mediados por la unión de CTCF a sitios convergentes de unión a CTCF y el subsiguiente reclutamiento de cohesina y topoisomerasa (TOP2B). Ilustración obtenida de Canela, A., Maman, Y., et al. 2017 (84).

Los datos sugieren que los TAD actúan regulando la probabilidad de contacto entre los promotores y los *enhancers*, puesto que, si se deleta el límite entre dos TAD se producen contactos entre regiones que estaban anteriormente aisladas, dando lugar a una expresión génica aberrante que deriva en problemas de desarrollo y en enfermedades como el cáncer (85).

También se ha visto que en la formación de estas interacciones de cromatina se crean unas tensiones en el ADN que son eliminados por cortes en la doble hebra y que crean cierta diversidad genética al ser reparados (84).

### 3.3. Organización tridimensional del locus *Tcra/Tcrd*

El locus *Tcra/Tcrd* está organizado de tal forma que permite que haya una compleja regulación temporal para pasar de los reordenamientos V $\delta$ D $\delta$ J $\delta$  del gen *Tcrd* en los estadios DN2 y DN3a al reordenamiento V $\alpha$ Ja del gen *Tcra* en el estadio DP durante el desarrollo de los timocitos (4) (Figura 5). Esta estructura evita la expresión conjunta de ambos genes, ya que la expresión de un TCR $\gamma\delta$  en timocitos DN2 y DN3a dirige su diferenciación hacia linfocitos T  $\gamma\delta$  (4) (Figura 1). Los estadios DN2/DN3a y DP del desarrollo de los timocitos coinciden con dos oleadas de expresión de RAG1 y RAG2, responsables de que ocurran estos reordenamientos diferenciados durante estas dos fases durante el desarrollo de los timocitos (49) (Figura 1). Por otro lado, los *enhancers* de los genes *Tcra* y *Tcrd*, E $\alpha$  y E $\delta$  respectivamente, son responsables de la transcripción germinal y la recombinación V(D)J de su gen correspondiente durante el desarrollo de los timocitos, mediante la inducción de modificaciones en la estructura de la cromatina dentro del área de influencia de cada uno de dichos *enhancers* (4,22,23,25,26,43,86). Además, los promotores de los diferentes segmentos génicos V $\alpha$ , V $\delta$ , D $\delta$ , J $\delta$  y Ja regulan de forma local la accesibilidad de RAG1 y RAG2 (59), siendo la elongación transcripcional necesaria para que ocurra una apertura de la cromatina en los segmentos génicos que están lejos de los promotores (43,47,59,87).

Para que ocurran los reordenamientos V $\delta$ D $\delta$ J $\delta$  y se exprese la cadena TCR $\delta$  en timocitos DN2 y DN3a, se activa primero la transcripción germinal del gen *Tcrd* a partir de los promotores asociados a los segmentos D $\delta$  y J $\delta$  en un proceso que depende de E $\delta$  (23,26) (Figura 5). Por otro lado, para los reordenamientos V $\alpha$ Ja y la consecuente expresión de la cadena TCR $\alpha$  en timocitos DP, la transcripción germinal del gen *Tcra* se inicia en los promotores asociados a los segmentos génicos V $\alpha$  y Ja implicados en los reordenamientos primarios en un proceso dependiente de E $\alpha$  (Figura 5). Una

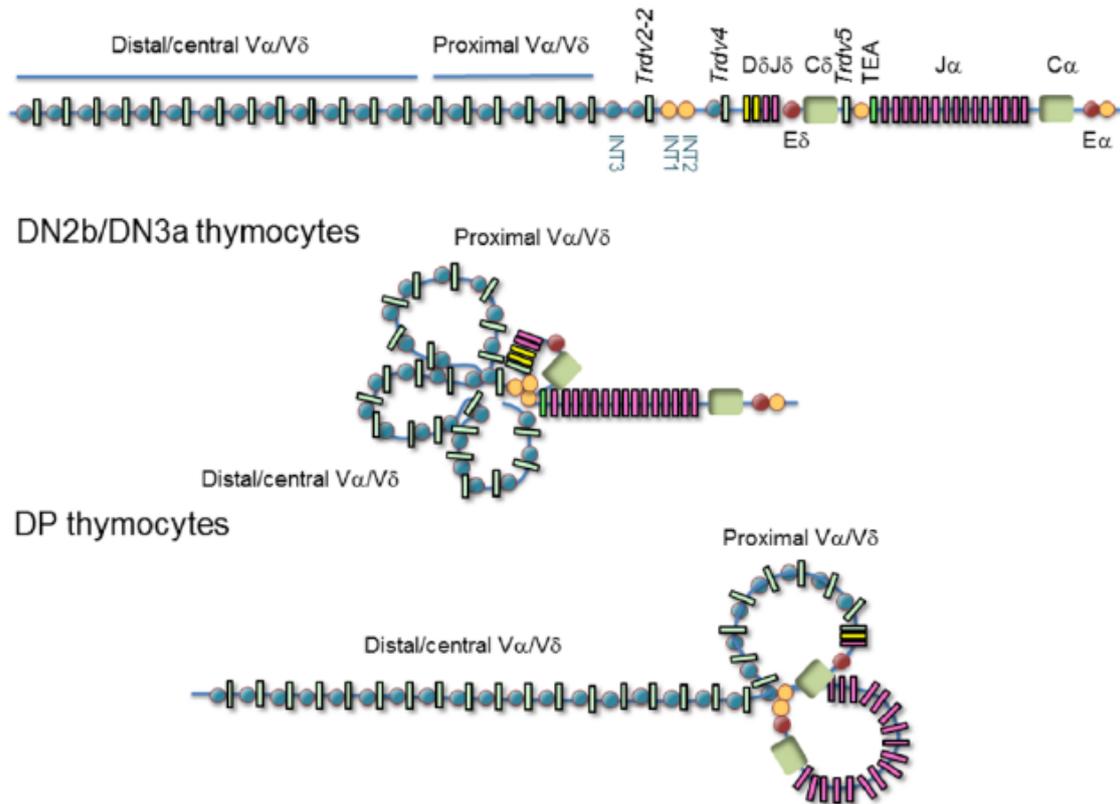
vez que los genes han reordenado, su transcripción depende de la activación de los promotores asociados a los segmentos  $V\alpha$  o  $V\delta$  implicados en el reordenamiento (4).

Como se indica en la Figura 5, en el caso de los reordenamientos  $V\delta D\delta J\delta$  del gen *Tcrd* solo hay un intento de recombinación en cada alelo, mientras que los intentos de reordenamiento del gen *Tcra* pueden ser múltiples en cada alelo debido al gran número de segmentos  $V\alpha$  y  $J\alpha$  y a la disposición de los mismos (4,59,88). Los reordenamientos  $V\alpha J\alpha$  iniciales o primarios en timocitos eDP ocurren debido a que las RSS de los segmentos  $V\alpha$  proximales y  $J\alpha$  distales respecto a  $C\alpha$  quedan accesibles a RAG1 y RAG2 en un proceso regulado por los promotores asociados a estos segmentos génicos tras su activación por  $E\alpha$  (Figura 5) (89,90). Dentro de los promotores implicados, los más importantes para que sucedan los reordenamientos primarios son TEAp y  $J\alpha_{49p}$  (59). Como consecuencia de los reordenamientos primarios se produce la delección del gen *Tcrd* del genoma en el estadio DP, que queda incluido en un círculo extracromosómico que se diluye durante la expansión clonal (Figura 5). Si ninguno de los reordenamientos primarios del gen *Tcra* en ambos alelos es productivo, se producen los reordenamientos secundarios en timocitos lDP entre segmentos génicos  $V\alpha$  situados progresivamente hacia 5' (o distales respecto a la posición de  $C\alpha$ ) con segmentos génicos  $J\alpha$  situados progresivamente más hacia 3' (o proximales respecto a la posición de  $C\alpha$ ) (Figura 5). Los reordenamientos secundarios son dependientes de la transcripción iniciada en los promotores de los segmentos génicos  $V\alpha$  involucrados en la recombinación primaria previa (91). Este mecanismo por el que puede haber, sucesivamente, diferentes reordenamientos  $V\alpha J\alpha$  en cada alelo *Tcra*, ha sido demostrado experimentalmente (58) y asegura la expresión de una cadena TCR $\alpha$  en todos los timocitos DP, para que todos ellos puedan expresar un TCR $\alpha\beta$  que les permita pasar por los procesos de selección tímica.

La conformación espacial del locus *Tcra/Tcrd* ha sido determinada mediante estudios de microscopía confocal con sondas fluorescentes (3-Dimensional-Fluorescence In Situ Hybridation, 3D-FISH) (92). Se observó que el locus está totalmente compactado en timocitos DN3a, cuando ocurren

los reordenamientos V $\delta$ D $\delta$ J $\delta$  del gen *Tcrd* (Figura 8). No obstante, al observar los timocitos DP, cuando ocurren los reordenamientos VaJa del gen *Tcra*, el locus tiene una configuración estirada en la región 5', donde se encuentran los segmentos Va, y una configuración contraída en la región 3', donde se localizan los segmentos Ja y Ea (Figura 8). La conformación contraída en timocitos DN3a permite que el gen *Tcrd*, en el que sólo puede haber un reordenamiento por alelo, pueda utilizar cualquier segmento V $\delta$  situado a lo largo de todo el *cluster* de segmentos Va/V $\delta$ . Por el contrario, en timocitos DP, el gen *Tcra* puede tener sucesivos ciclos de reordenamientos secuenciales, primarios y secundarios, de forma ordenada que se ven favorecidos por la conformación estirada de la región 5' del locus donde se encuentran los segmentos Va distales y centrales, junto con la conformación contraída de la región 3' donde están los segmentos Va proximales, Ja, Ca y Ea. La región contraída en la parte 3' del locus, de alrededor de 525kb, corresponde con el área de influencia de Ea (89). Los experimentos 3D-FISH fueron realizados en el contexto de ratones *Rag2*<sup>-/-</sup> por lo que se realizaron con el locus en una configuración no reordenada, por tanto los cambios de conformación del locus observados ocurren de forma independiente al proceso de recombinación V(D)J.

Se han identificado interacciones físicas mediante lazos de cromatina que se forman en el locus *Tcra/Tcrd* durante el desarrollo de los timocitos mediante análisis de 3C y 4C (86,89,93,94). Estos lazos de cromatina junto con los cambios en su estructura mediados por las interacciones promotor-*enhancer* controlan la accesibilidad de RAG1 y RAG2 a las RSS asociadas a los segmentos V, (D) y J, siendo ésta la base molecular del control del proceso de la recombinación V(D)J (4,95,96).



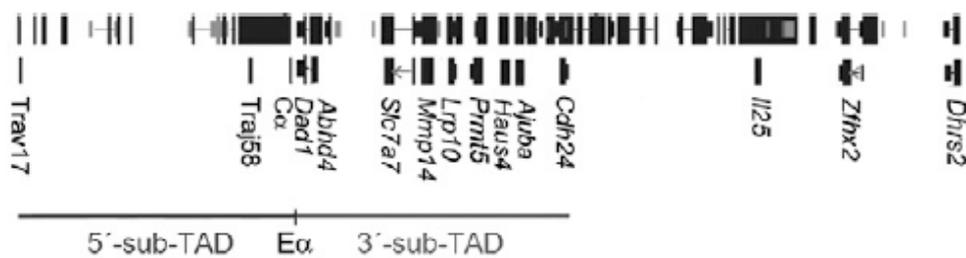
**Figura 8. Modelos de configuración de los lazos de cromatina en el locus *Tcrα/Tcrδ*.** Los segmentos génicos variable (V), diversidad (D) y unión (J) se representan con rectángulos verdes, amarillos y rosas respectivamente. El exón TEA se representa con un rectángulo naranja. Los genes del tripsinógeno (T) se representan con rectángulos marrones. Los *enhancers* se representan con círculos rojos. Los sitios de unión a CTCF (CBE) se representan con círculos naranjas o azules (los naranjas representan los CBE más relevantes involucrados en la estructura génica tridimensional). En el centro de recombinación (segmentos D $\delta$  y J $\delta$ ) del gen *Tcrδ* se establece con un nudo de cromatina formado entre los CBE situados en INT1/2 y el CBE del promotor TEA (TEAp) en timocitos DN2b y DN3a, mientras que los segmentos génicos V $\alpha$ /V $\delta$  de la región 5' del locus forman nudos adicionales haciendo que el locus se contraiga. En timocitos DP, en el centro de recombinación del gen *Tcrα* (segmentos V $\alpha$ /V $\delta$  proximales y J $\alpha$ ) se establece un nudo de cromatina entre el CBE de TEAp y los CBE adyacentes a E $\alpha$ , mientras se forman nudos adicionales en los segmentos V $\alpha$ /V $\delta$  proximales para facilitar sucesivos reordenamientos V $\alpha$ J $\alpha$ . Ilustración obtenida de Rodríguez-Caparrós, A., et al. 2020 (4).

E $\delta$  se encuentra dentro de un lazo de cromatina en timocitos DN2-DN3a que incluye a los segmentos D $\delta$  y J $\delta$ , la región C $\delta$  y el segmento *Trdv5*, lo que explica la restricción de la actividad de E $\delta$  a los segmentos D $\delta$  y J $\delta$  (93) (Figura 8). La formación de este lazo se debe a interacciones físicas mediadas por el factor CTCF que se une a dos CBE situados 5' de los segmentos D $\delta$ , denominados INT1 e INT2, orientados inversamente, con otro CBE asociado a TEAp (4,93,97). Si se destruye este lazo, mediante la delección de INT1 e INT2, los segmentos D $\delta$  solo reordenan con un único segmento V $\delta$  situado

inmediatamente en dirección 5', *Trdv4*, por lo que se reduce de forma considerable el repertorio de las cadenas TCR $\delta$  (93). El repertorio de cadenas TCR $\alpha$  también se ve afectado por la delección de INT1 e INT2, puesto que los reordenamientos del gen *Tcrd* cambian la estructura del locus debido a que los segmentos V $\delta$  se encuentran dispersos a lo largo del *cluster* de segmentos V $\alpha$ /V $\delta$  (58,61,93). Por tanto, el hecho de que la actividad de E $\delta$  quede aislada dentro del lazo INT1/INT2-TEAp, es importante para generar la variedad de repertorios tanto de la cadena TCR $\delta$  como de la cadena TCR $\alpha$  (58,93,98).

Por otro lado, en timocitos DP, mediante análisis 3C (*Chromosomal Conformation Capture*), se ha demostrado la existencia de un lazo de cromatina de unos 525kb que corresponde con la región del locus *Tcr $\alpha$ /Tcr $\delta$*  contraída observada en los experimentos 3D-FISH (89,92) (Figura 8). Este lazo se forma como consecuencia de la unión de CTCF/cohesina a los CBE asociados a los promotores de los segmentos génicos V $\alpha$  proximales, TEAp y la región 3' inmediatamente adyacente a E $\alpha$  (4,89,96–99). Para que puedan formarse estos lazos de cromatina, los CBE asociados a los promotores V $\alpha$  y TEAp están orientados en dirección opuesta a los asociados a E $\alpha$  (97). Estos lazos de cromatina favorecen la activación transcripcional dependiente de E $\alpha$  de los promotores V $\alpha$  proximales y los J $\alpha$  distales respecto a C $\alpha$ , promoviendo los reordenamientos primarios, y si no se produce un reordenamiento productivo ocurran reordenamientos secundarios, mediante la recombinación de otros segmentos V $\alpha$  en orden 3' - 5' con otros segmentos J $\alpha$  en dirección 5' - 3' (4,58,59,89,91,92,98,99). Es decir, la formación de este lazo dirige los reordenamientos sucesivos del gen *Tcr $\alpha$*  (Figura 5). El orden en los reordenamientos V $\alpha$ J $\alpha$  determina que tras los reordenamientos primarios quede un *pool* de segmentos V $\alpha$  centrales y distales y de J $\alpha$  centrales y proximales necesarios para los reordenamientos V $\alpha$ J $\alpha$  secundarios (58,98). Aunque E $\alpha$  es esencial para la formación de este lazo en la cromatina en timocitos DP, tanto E $\alpha$ , E $\delta$  y CTCF son prescindibles para las contracciones del locus que se observan mediante 3D-FISH (92), sin que se conozca por el momento el mecanismo de control de las contracciones del locus detectadas mediante microscopía confocal.

Es de interés que los dos CBE adyacentes a  $E\alpha$  ( $E\alpha$ CBE) están localizados en el límite entre dos sub-TAD, siendo estos dos CBE fundamentales para el aislamiento entre ambos sub-TAD (99). La formación de estos dos sub-TAD en timocitos DP dependen de la formación de lazos de cromatina hacia ambos lados: El 5' sub-TAD depende de la formación de lazos entre el 5' - $E\alpha$ CBE y los CBE asociados a los promotores  $V\alpha/V\delta$  proximales en una región de 525 kb, con límite en *Trav17*, mientras que el 3' sub-TAD depende de la formación de lazos entre el 3' - $E\alpha$ CBE y los CBE asociados a los genes que van desde *Dad1* hasta *Chd24* en una región de 400 kb (Figura 9). Los datos de Zhao et al. (99) indican que la delección de los dos  $E\alpha$ CBE afecta negativamente a la transcripción de los segmentos génicos *Trav19* y *Trav21*, situados en el extremo 5' del 5' sub-TAD, y de *Ajuba* y *Cdh24*, situados en el extremo 3' del 3' sub-TAD, apoyando un mecanismo de acción de estos CBE en la formación de ambos sub-TAD mediado por la extrusión de la cromatina de ambos lados de cromatina.



**Figura 9. Posición genómica del 5' sub-TAD y los genes del 3' sub-TAD respecto a la de  $E\alpha$ .** Se indica la posición de  $E\alpha$  entre ambos sub-TAD. Ilustración obtenida de Rodríguez-Caparrós, A., et al 2022 (100).

## 4. Elementos reguladores en *cis*: *enhancers* y promotores

### 4.1 *Enhancers*

Los *enhancers* se describieron por primera vez en la década de 1.980, definiéndose como secuencias de ADN no codificante que aumentan la transcripción de sus genes diana de manera independiente de su distancia y orientación (101). El primer *enhancer* descubierto fue una secuencia de 72 pb del virus de simio 40 (*Simian Virus 40*, SV40), capaz de aumentar la expresión

del gen de la  $\beta$ -globina humana (*ACTB*) en plásmidos transfectados (102), mientras que el primer *enhancer* en mamíferos descrito en 1983 fue el del locus *Igh* de ratón, E $\mu$  (103). Desde entonces se han descubierto multitud de *enhancers* (su número excede en mucho al número de genes, estimándose en cientos de miles) y se ha demostrado que actúan mediante el reclutamiento de factores de transcripción específicos para regular la expresión génica espacio-temporal durante el desarrollo celular, tanto en situaciones de fisiología normal como durante la enfermedad (104).

La transcripción de los genes diana de un *enhancer* está mediada por la unión cooperativa de factores y cofactores de transcripción, enzimas de remodelación de la cromatina y, en última instancia, por el complejo de la ARN polimerasa II (105). No obstante, el ensamblaje de estas proteínas y enzimas necesita que el ADN se encuentre accesible. Por lo tanto, la accesibilidad de la cromatina es una característica de los *enhancers* activos que puede medirse con distintas técnicas, tales como la hipersensibilidad a la acción de la enzima DNaseI, la detección de elementos reguladores por tratamiento con formaldehído (FAIRE) o mediciones de accesibilidad a la cromatina por la función de la transposasa Tn5 (ATAC) (106).

El estudio del estado epigenético de los *enhancers* se usa para conocer su estado funcional. Por ejemplo, los *enhancers* inactivos típicamente presentan monometilación en la lisina 4 de la histona H3 (H3K4me1) y carecen de la marca H3K4me3, mientras que los *enhancers* activos presentan acetilación de la lisina 27 de la histona H3 (H3K27ac) y H3K4me3 (107–109). Otras características de los *enhancers* activos incluyen el reclutamiento de enzimas acetil-transferasas de histonas (HAT) CBP/p300 y la detección de transcritos de *enhancer* (eRNA) (110,111).

Todos los loci de los genes de las cadenas de los receptores de antígeno, tanto TCR como BCR, presentan en su estructura génica *enhancers*, localizados generalmente en la región 3' del locus. Diversos estudios funcionales hechos para establecer el papel de dichos *enhancers* realizados mediante su delección en ratones deficientes (*Knock-Out*) o mediante ratones transgénicos que contienen construcciones reporteras, muestran el papel esencial que estos elementos tienen en la activación de la recombinación V(D)J

(4,27,43,112). Cuando se deleta el *enhancer* de un gen del TCR no ocurre la transcripción germinal, ni los reordenamientos del gen en cuestión. Al mirar las marcas epigenéticas se observa una disminución de la acetilación de las histonas H3 y H4, por lo que la cromatina no se encuentra accesible para las maquinarias de transcripción y recombinación (80).

Estos datos indican que los *enhancers* son elementos que controlan la accesibilidad del ADN a largas distancias iniciando la apertura de la cromatina a nivel general, modulando la transcripción iniciada en sus promotores específicos y, en el caso de los genes de los receptores de antígeno, la transcripción germinal y los reordenamientos de sus segmentos génicos (4).

## 4.2 Promotores

Los promotores son regiones localizadas en el inicio de un gen, que sirven como sitio de reclutamiento de la maquinaria transcripcional, definiendo el sitio donde comienza la transcripción, así como su regulación (113). El primer lugar de iniciación de la transcripción descrito fue la caja TATA (101). Los promotores activos presentan un enriquecimiento de la marca H3K4me3 y una reducción de la marca H3K4me1 (114).

Tanto en los loci de los TCR como en los de los BCR, los promotores se encuentran asociados a los segmentos génicos (4). Antes de que ocurra la recombinación V(D)J, los promotores de estos segmentos génicos controlan la transcripción germinal no codificante, pero una vez ocurridos los reordenamientos productivos, el promotor del segmento génico V implicado en la recombinación es el responsable de la iniciación de la transcripción del gen reordenado que dará lugar a la cadena del receptor de antígeno. Al contrario que la acción de los *enhancers*, la acción de los promotores tiene un efecto localizado en la estructura de la cromatina que se restringe a un segmento génico asociado o a un grupo de segmentos génicos cercanos (4). Además, la mayoría de los promotores presentes en los loci TCR o BCR dependen de la actividad de su *enhancer* correspondiente, teniendo muchos de ellos una leve o nula actividad independiente de *enhancer* (4,28).

Para demostrar el efecto local de los promotores en la activación de la recombinación V(D)J se realizaron diversas deleciones de estos para comprobar sus efectos. Por ejemplo, la deleción del promotor asociado al segmento génico D $\beta$ 1 resulta en la ausencia de transcripción del segmento D $\beta$ 1 y de la recombinación D $\beta$ 1J $\beta$ , pero no afecta a la transcripción del segmento D $\beta$ 2 y a la recombinación D $\beta$ 2J $\beta$ , las cuales dependen del promotor D $\beta$ 2 (115). Respecto a las marcas epigenéticas, se observa que la acetilación de histonas en la zona del promotor delecionado es menor (116). Efectos similares se han visto al delecionar el promotor del segmento génico *Trbv14* en el locus *Tcrb* (117), TEAp o Ja49p en el gen *Tcra* (59,118), entre otros, comprobándose que la inhibición de la transcripción y recombinación solo afecta a los segmentos génicos específicamente controlados por el promotor eliminado.

Hay diferentes mecanismos moleculares que abren la cromatina en la región del promotor y que, por lo tanto, facilitan el comienzo de la transcripción. Entre estos mecanismos moleculares se encuentran la presencia de regiones empobrecidas en nucleosomas (detectados por hipersensibilidad a DNasaI) o el enriquecimiento de modificaciones específicas en las histonas, como la dimetilación de la lisina 4 de la histona H3 (H3K4me2) o H3K4me3 y la acetilación de la histona H3 (H3K4ac y H3K27ac) (119). La proteína RAG2 con su dominio de dedo PHD es capaz de reconocer la marca H3K4me3 y, por lo tanto, reconocer promotores activos (73,74). Debido a que la unión de RAG2 es general en todos promotores activos del genoma, la especificidad recombinasa en los loci de TCR y BCR se debe a la unión específica de RAG1 a las RSS (90). Además la marca H3K4me3 estabiliza la asociación del complejo RAG1/RAG2 y aumenta su actividad enzimática (81).

#### 4.3. Holocomplejos promotor-*enhancer*

Como ya se ha mencionado, en la interacción entre los *enhancers* y los promotores se produce un lazo en el ADN que acerca ambos elementos reguladores (120). Los genes que codifican las diferentes cadenas de los TCR y BCR están dispuestos de tal manera que promotores y *enhancers* se encuentran situados a largas distancias y es mediante la formación de lazos o

nudos en el ADN que se forma un holocomplejo transcripcional. De esta forma los *enhancers* reclutan diversos factores de transcripción y cofactores directamente sobre los promotores mediante este hipotético mecanismo descrito en estudios realizados en *ACTB* (121,122). Por ejemplo, ensayos de inmunoprecipitación de cromatina (ChIP) en *Tcrb* han mostrado que factores de transcripción asociados al segmento génico D $\beta$ 1, como son SP1 o TBP, están asociados también a E $\beta$ , hecho que se debe a la proximidad espacial entre el promotor y el *enhancer* (116). Sin embargo, el modelo actual sostiene que en la comunicación entre los *enhancers* y sus promotores específicos se forma un lazo de ADN que permite una interacción entre ambos elementos dentro de unos condensados que se forman por separación de fases donde se acumulan los factores y cofactores de transcripción (123,124).

Además, se ha demostrado que existe especificidad en la interacción de los promotores y los *enhancers*, especificidad que aún no está bien entendida, pero en la que influye la compatibilidad bioquímica, restricciones impuestas por la conformación tridimensional de los cromosomas, elementos aislantes y los efectos del entorno local de la cromatina (119).

## 5. El *enhancer* del gen *Tcr $\alpha$* (E $\alpha$ )

E $\alpha$  se localiza entre la región C $\alpha$  del gen de la cadena TCR $\alpha$  y el siguiente gen en dirección 3' que codifica por DAD1, de una forma conservada entre humanos y ratones (Figura 9). En concreto, E $\alpha$  se encuentra a unas 4 kb en dirección 3' de la región C $\alpha$  dentro de una región de 7,4 kb descrita como reguladora del locus (*Locus Control Region*, LCR) del gen *Tcr $\alpha$*  (125). Este LCR consta de varias regiones identificadas en ensayos de hipersensibilidad a la digestión por DNasaI (*Hypersensitivity Site*, HS), de las cuales la región *core* de E $\alpha$  de 275 pb (E $\alpha$ core) constituye la región HS1 (125). A 500 pb 3' de E $\alpha$ core se encuentra la región HS1' que corresponde a los dos E $\alpha$ CBE mencionados anteriormente involucrados en la separación del 5' sub-TAD y el 3' sub-TAD (99,126,127). Aunque se propuso que los demás HS del LCR podrían contribuir a aislar al gen *Tcr $\alpha$*  de la influencia de las regiones reguladoras presentes en el

3' sub-TAD y viceversa (128), la delección de la región HS2-HS6 no tiene consecuencias funcionales (99).

En esta tesis la definición de E $\alpha$  incluye a la región de 1 kb que incluye a E $\alpha$ core y los dos E $\alpha$ CBE adyacentes que fue delecionada para la generación de ratones E $\alpha^{-/-}$  (22). La actividad de E $\alpha$  se ha demostrado que es imprescindible para el correcto desarrollo de los linfocitos T, ya que este elemento es el encargado de activar la transcripción germinal no codificante necesaria para que ocurran los reordenamientos V $\alpha$ J $\alpha$  del gen *Tcra* en timocitos DN4 y DP para dar lugar a linfocitos T  $\alpha\beta$  (22). Además, su actividad es esencial para la transcripción y expresión normal del gen *Tcrd* reordenado en linfocitos T  $\gamma\delta$ , así como la transcripción normal de los genes situados en los extremos del 3' sub-TAD en timocitos DP (22,99,100).

### 5.1. Papel de E $\alpha$ en la regulación transcripcional del gen *Tcra*

Como se ha mencionado anteriormente, E $\alpha$  se encuentra inactivo mientras ocurren los reordenamientos del gen *Tcrd* en los estadios DN2b a DN3a, y no es hasta que se recibe la señal del pre-TCR cuando se activa en timocitos DN4 y eDP, suceso que da lugar a la detección de transcritos germinales del gen *Tcra* (4,23,60). La influencia de E $\alpha$  se extiende a 525 kb hacia 5', estando dentro de su área de influencia 1/3 de todos los segmentos V $\alpha$ /V $\delta$  más cercanos con límite en *Trav17*, que constituyen los llamados V $\alpha$  proximales que están implicados en los reordenamientos V $\alpha$ J $\alpha$  primarios, y todos los segmentos J $\alpha$ , formando los lazos de cromatina que conforman el 5' sub-TAD identificado en timocitos DP (Figuras 8 y 9) (89,99). Los cambios en la cromatina provocados por la activación de E $\alpha$ , aumentan la accesibilidad de toda esta región y se deben a modificaciones epigenéticas, principalmente a acetilación de las histonas H3 y H4, y metilaciones de H3K4 (H3K4me1, H3K4me2 y H3K4me3) y H3K36 (H3K36me3) (4,59,80). De esta forma, la activación de E $\alpha$  da lugar a la transcripción germinal del gen *Tcra* iniciada en los promotores de los segmentos V $\alpha$  proximales, TEAp y J $\alpha$ 49p, además de otros promotores J $\alpha$  distales (59,88), siendo TEAp y J $\alpha$ 49p imprescindibles para que ocurran los reordenamientos primarios V $\alpha$ J $\alpha$  (59,118).

A pesar de que  $E\alpha$  es esencial para la recombinación germinal no codificante necesaria para la iniciación de los reordenamientos primarios del gen *Tcra*, su ausencia no impide la presencia en ratones  $E\alpha^{-/-}$  de una pequeña población de linfocitos T  $\alpha\beta$  que presentan unos reordenamientos que implican a los segmentos génicos *Trav14* ( $V\alpha 2$ , según la nomenclatura antigua) con segmentos diversos  $J\alpha$  (22,129). Estos reordenamientos son debidos a la acción de  $E\delta$  en los estadios DN2 y DN3a (22,130).

Cuando los timocitos ya han completado su desarrollo, es decir, son timocitos SP o linfocitos T  $\alpha\beta$ ,  $E\alpha$  se encuentra fuertemente inhibido, con un 15% de la actividad que tenía en timocitos DP, medida en análisis de la transcripción germinal no codificante en ratones *Rag2*<sup>-/-</sup> (94). Aunque se intente recuperar la actividad del *enhancer* en linfocitos T  $\alpha\beta$  a través de su estimulación con tratamientos con ésteres del forbol (*Phorbol 12 Miristate 13 Acetate*, PMA) e ionomicina, a través de la activación del TCR $\alpha\beta$  o de la inducción de la diferenciación a linfocitos  $Th_0$  a timocitos  $Th_1$ ,  $Th_2$  o  $T_{reg}$ , el estado inhibido del *enhancer* se mantiene inalterado, por lo que se concluyó que dicho estado inhibido es estable durante la homeostasis de los linfocitos T  $\alpha\beta$  maduros (94).

## 5.2. Papel de $E\alpha$ en la regulación transcripcional de los genes del 3' sub-TAD

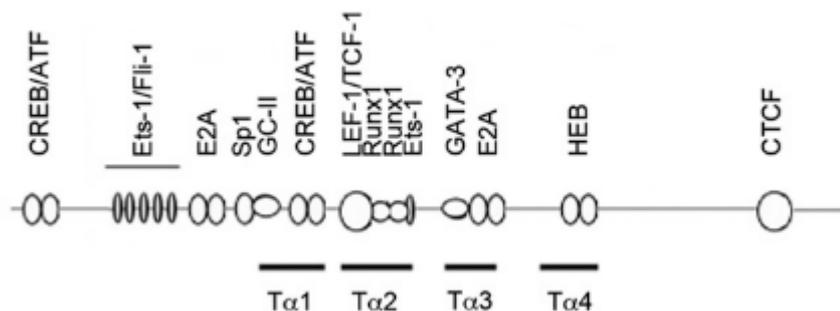
Como se ha mencionado, los dos  $E\alpha$ CBE adyacentes a  $E\alpha$ core se encuentran en el límite entre dos sub-TAD. Estos  $E\alpha$ CBE juegan un papel relevante en la transcripción de los segmentos  $V\alpha$  situados en el extremo del 5' sub-TAD y de los genes situados en el extremo del 3' sub-TAD en timocitos DP procedentes de ratones *Rag2*<sup>-/-</sup> (sus timocitos DN3a diferencian a DP al ser inyectados con un anticuerpo para CD3 $\epsilon$ ) (99). Los datos de Zhao et al. (99) mostraron que la delección de  $E\alpha$  (incluyendo  $E\alpha$ core junto con los dos  $E\alpha$ CBE) inhibe la transcripción de todos los promotores que fueron analizados situados en ambos sub-TAD en una región de alrededor de 1Mb en timocitos DP. Estos datos sugirieron que  $E\alpha$ core tiene una influencia positiva en la transcripción de los genes de ambos sub-TAD, que en el caso del 3' sub-TAD puede ser

bloqueada por los E $\alpha$ CBE para los promotores de los genes más proximales (*Dad1* y *Abdh4*) y facilitada para los promotores de los genes más distales (*Prmt5*, *Ajuba* y *Cdh24*) (99).

Nuestro laboratorio ha analizado recientemente la transcripción de los genes del 3' sub-TAD en timocitos DP (100). Los resultados indican que la delección de E $\alpha$  (E $\alpha$ core + E $\alpha$ CBE) afecta solo a los genes del 3' sub-TAD situados en los extremos 5' (*Dad1* y *Abdh4*) y 3' (*Haus4* y *Cdh24*) en timocitos DP (en este caso procedentes de ratones *Rag2*<sup>-/-</sup> que expresan una cadena transgénica TCR $\beta$  (*Rag2*<sup>-/-</sup> x  $\beta$ tg)). Estos datos sugieren que la formación de lazos entre el 3' -E $\alpha$ CBE y los CBE presentes en los promotores localizados de los genes situados en el extremo del 3' sub-TAD permiten una mayor interacción entre E $\alpha$ core y los promotores implicados.

### 5.3. *Enhanceosoma* de E $\alpha$ durante el desarrollo de los linfocitos T $\alpha\beta$

La región E $\alpha$ core de 275 pb está formada por cuatro elementos para la unión de factores de transcripción denominados Ta1, Ta2, Ta3 y Ta4. Esta región Ta1-Ta4 constituye la mínima región imprescindible para la correcta regulación de la recombinación V(D)J durante el desarrollo de los timocitos (131). El fragmento Ta1-Ta2 se considera el *enhancer* mínimo, debido a que es la secuencia más corta que ha mostrado tener actividad transcripcional por sí sola en experimentos de transfecciones celulares para la expresión de genes reporteros (132). No obstante, la región Ta1-Ta2 presenta una regulación temporal inadecuada, ya que se activa prematuramente durante el desarrollo de los timocitos en experimentos de construcciones reporteras transgénicas de recombinación V(D)J en ratón (133). Por tanto, los factores de transcripción que se unen a la región Ta3-Ta4 son necesarios para una regulación temporal adecuada del *enhancer*, retrasando su activación hasta los estadios DN4 y DP (131). Además, la región Ta1-Ta2 no es capaz por sí sola de activar la recombinación VaJa del gen *Tcra* endógeno, cuando E $\alpha$  fue sustituido por Ta1-Ta2 (134), por lo que la región Ta3-Ta4 junto con los E $\alpha$ CBE adyacentes parecen necesarios para la función de E $\alpha$ .



**Figura 10. Esquema representativo de las regiones de unión de factores de transcripción a Eα.** Ilustración obtenida de Del Blanco, B., et al. 2015.

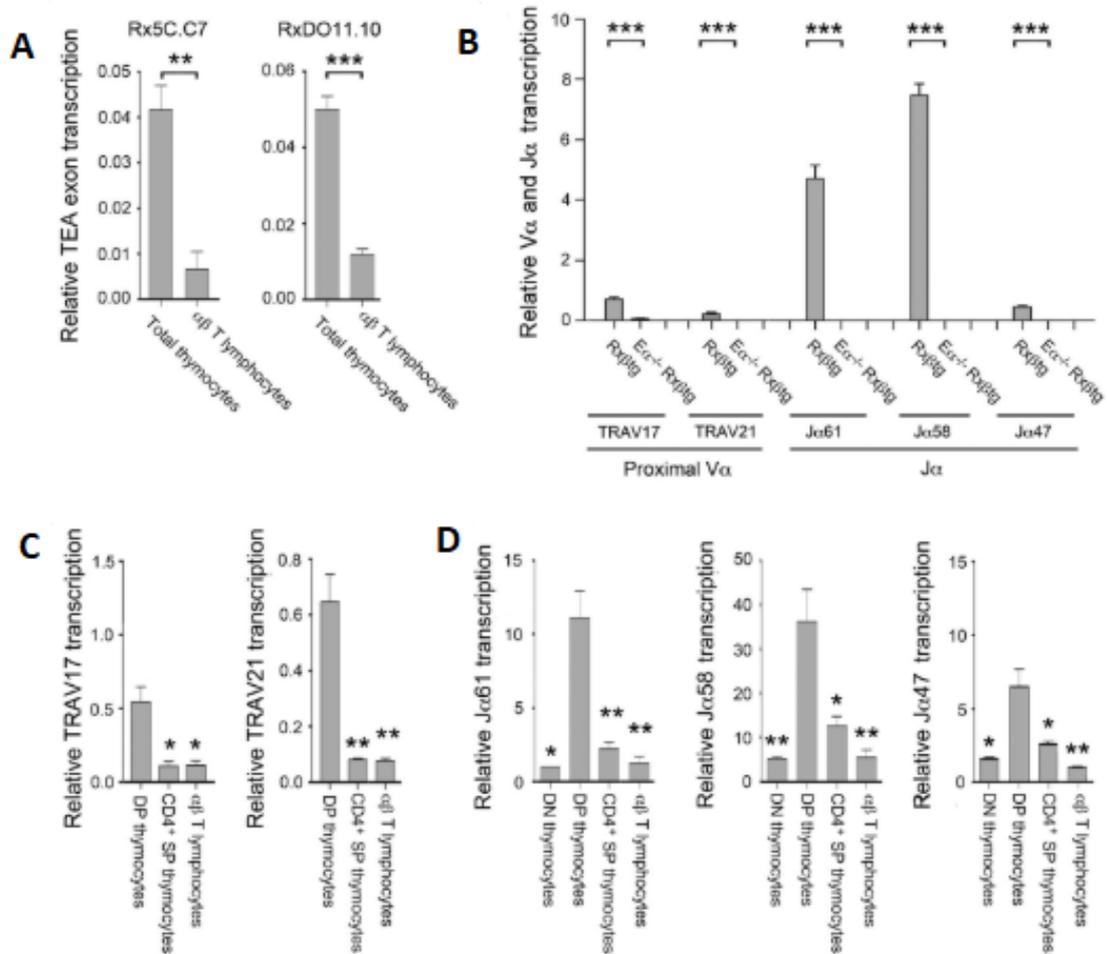
A la región Tα1-Tα4 se unen numerosos factores de transcripción constitutivos durante el desarrollo de los timocitos, tales como CREB, LEF-1/TCF-1, RUNX1 y ETS1, que acceden de manera sinérgica al ADN formando un *enhanceosoma* estable en Tα1-Tα2 (135), mientras que a la región Tα3-Tα4 se unen GATA3 y las proteínas E, E2A y HEB (136). Además de estos factores, en la región 5' de Tα1 se unen ATF2, FLI1, ÍKAROS y SP1 (136,137). En la figura 10 se muestra un esquema en el que se representan los diversos factores de transcripción que se unen a Eα en timocitos. Además, se ha visto recientemente que el reclutamiento de factores supresores de la familia Homeobox A (HOXA) al *enhancer* en timocitos DN3a a través de ETS1 está implicado en la inhibición de la función de Eα (138).

La señalización desencadenada por el pre-TCR permite que factores de transcripción inducibles, tales como NFAT, AP1 y EGR1/EGR3, se unan a Eα, mientras que se disocian los factores HOXA, haciendo que el *enhancer* se active en timocitos DN4 y eDP (60,138). En el estadio IDP, los factores de transcripción inducibles ya no están presentes pero el *enhancer* se mantiene activo, debido posiblemente por la fuerte unión de las proteínas E, E2A y HEB, junto con la ausencia de los factores HOXA (94,138). Por el contrario, el estado fuertemente inhibido del Eα en linfocitos T αβ se relaciona con una fuerte disminución de la unión de HEB y E2A, aunque la expresión forzada de E2A no es suficiente para recuperar su actividad en linfocitos Tαβ maduros (94). Actualmente se desconoce el mecanismo molecular relacionado con la pérdida de actividad de Eα en linfocitos T maduros.

## 6. Antecedentes

### La actividad de E $\alpha$ se inhibe en el contexto del gen no reordenado en linfocitos T $\alpha\beta$

Para conocer y comparar la actividad de E $\alpha$  en timocitos IDP y linfocitos T  $\alpha\beta$ , nuestro laboratorio realizó previamente un estudio (94) comparando la abundancia relativa de los transcritos *Tcra* germinales dependientes del *enhancer* en el gen no reordenado en ambos tipos celulares utilizando diferentes cepas de ratón *Rag2*<sup>-/-</sup> (Figura 11A). Los transcritos fueron cuantificados utilizando la técnica de la reacción en cadena por polimerasa cuantitativa (*quantitative Polymerase Chain Reaction*, qPCR) mediada por retro-transcripción (RT-qPCR). Se analizaron dos cepas de ratones *Rag2*<sup>-/-</sup> que expresan un TCR $\alpha\beta$  transgénico: *Rag2*<sup>-/-</sup> x 5C.C7 (Rx5C.7C) y *Rag2*<sup>-/-</sup> x DO11.10 (RxDO11.10). La ausencia de RAG2 impide la recombinación V(D)J de los genes endógenos y permite el análisis de los transcritos no codificantes germinales de los genes de los receptores de antígeno. La expresión de un TCR $\alpha\beta$  transgénico permite que los timocitos DN3a puedan completar su desarrollo a linfocitos T $\alpha\beta$ , ya que los ratones *Rag2*<sup>-/-</sup> tienen un bloqueo total en el estadio DN3a (139,140). Así pues, estos modelos de ratón permitieron el análisis de los transcritos no codificantes germinales endógenos del gen *Tcra* en timocitos y linfocitos T. Como se puede observar en la figura 11A-D, los transcritos del exón TEA, de los segmentos J $\alpha$  distales (J $\alpha$ 61, J $\alpha$ 58 y J $\alpha$ 47) y los segmentos V $\alpha$  proximales (V $\alpha$ 17/*Trav17* y V $\alpha$ 21/*Trav21*), cuya transcripción depende de E $\alpha$  puesto que en ratones E $\alpha$ <sup>-/-</sup> no se detectaron (Figura 11B), mostraron una transcripción fuertemente disminuida en aproximadamente un 80-85% en timocitos SP y linfocitos T  $\alpha\beta$  respecto a la transcripción observada en timocitos DP (Figuras 11A, 11C y 11D).



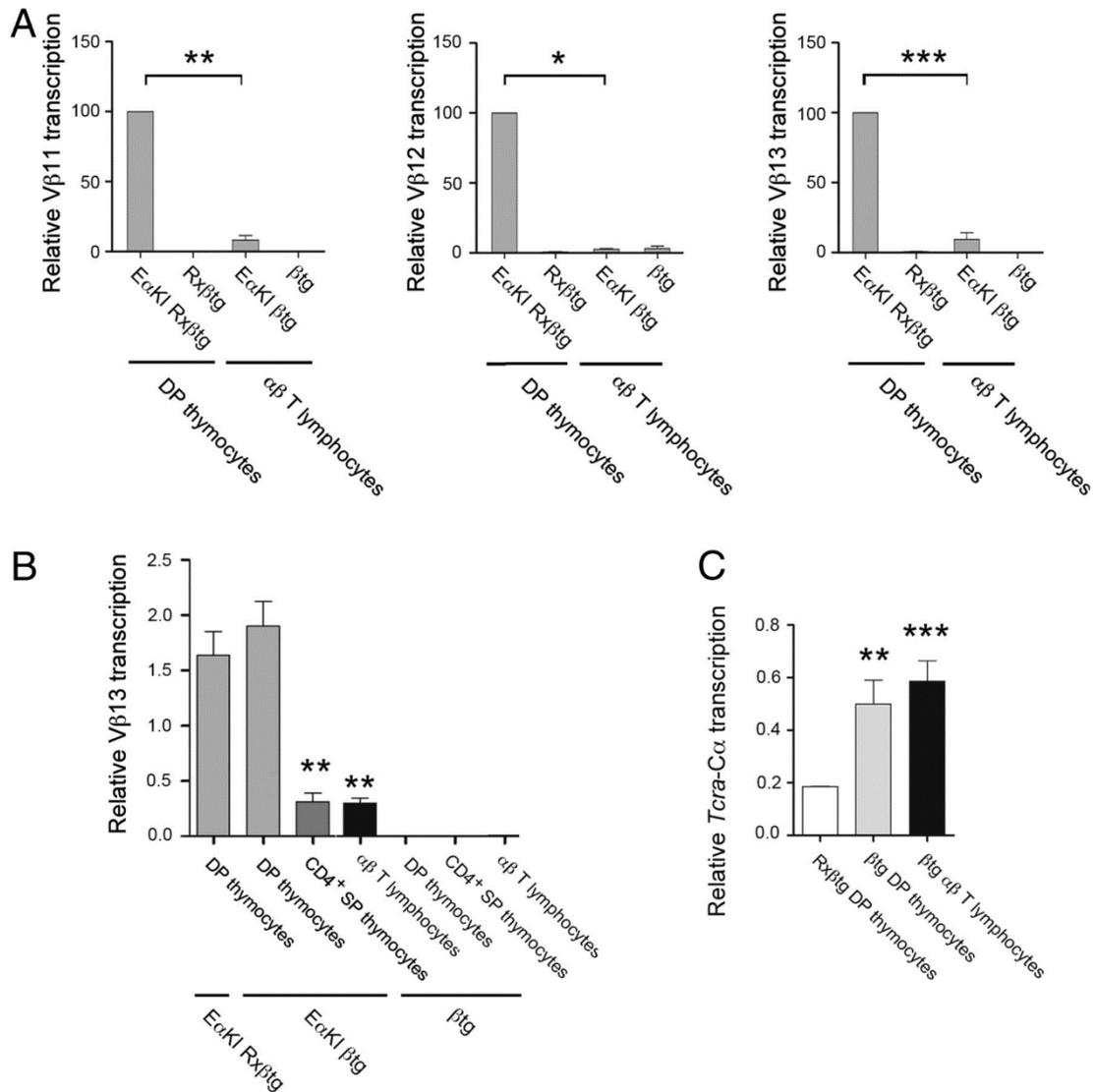
**Figura 11. La transcripción germinal no codificante dependiente de Ea del gen *Tcrα* se encuentra fuertemente inhibida en timocitos SP y linfocitos T  $\alpha\beta$ .** (A) Transcripción del exón TEA en timocitos totales y linfocitos  $\alpha\beta$  de ratones Rx5C.C7 y RxDO11.10. (B) Dependencia de Ea de la transcripción de los segmentos Va proximales *Trav17* y *Trav21*, y de los segmentos Ja distales *Traj61*, *Traj58* y *Traj47* en timocitos DP de ratones *Rag2<sup>-/-</sup>* x transgén *Tcrb* (Rx $\beta$ tg) y ratones *Rag2<sup>-/-</sup>* x transgén *Terb* x EaKO (Ea<sup>-/-</sup> Rx $\beta$ tg). (C) Transcripción de los segmentos Va proximales *Trav17* y *Trav21* en timocitos DP y SP y linfocitos T  $\alpha\beta$  de ratones RxDO11.10. (D) Transcripción de los transcritos de Ja distales en timocitos DP y SP y linfocitos T  $\alpha\beta$  de ratones RxDO11.10. Los transcritos indicados se analizaron por RT-qPCR. Los experimentos se realizaron por duplicado de tres experimentos independientes. Los datos fueron normalizados con la transcripción de *Actb*. Se representa la media de los resultados +/- media del error estándar (*Standard Error Mean*, SEM). El valor estadístico *p* se representa por \* (\*<0,05, \*\*<0,005 y \*\*\*<0,0005). Figura tomada de (del Blanco et al., 2015).

Para confirmar que Ea es el elemento regulador que se encuentra inhibido en los timocitos SP y linfocitos T  $\alpha\beta$ , y no los promotores de los segmentos génicos analizados, se utilizaron ratones con una copia de Ea introducida (ratones Ea *Knock-In*, EaKI) en el clúster de segmentos génicos V $\beta$  del locus *Tcrb* (141), de forma que los promotores de los segmentos V $\beta$ 11, V $\beta$ 12 y V $\beta$ 13 transcribieran de una forma dependiente de Ea en los timocitos

DP de estos ratones (Figura 12.A) (94). Comparando la transcripción de estos segmentos V $\beta$  en timocitos DP de ratones E $\alpha$ KI con fondo *Rag2*<sup>-/-</sup> que además expresan una cadena TCR $\beta$  transgénica (E $\alpha$ KI Rx $\beta$ tg) con la transcripción de los segmentos V $\beta$  en linfocitos T  $\alpha\beta$  de ratones E $\alpha$ KI se observó una profunda inhibición de la transcripción germinal de los segmentos V $\beta$  de estos últimos (Figura 12B), que era comparable con la que se encontró en los segmentos V $\alpha$  proximales y J $\alpha$  distales en el gen *Tcra* no reordenado en linfocitos T  $\alpha\beta$  respecto a timocitos DP (Figura 11A, 11C y 11D). Estas observaciones demostraron que E $\alpha$  es el elemento responsable de la inhibición de la transcripción germinal de los promotores dependientes de este *enhancer* en el contexto del gen *Tcra* no reordenado en linfocitos T  $\alpha\beta$ .

Así pues, E $\alpha$ , una vez se activa por las señalizaciones desencadenadas por el pre-TCR, es esencial para que se transcriba el gen *Tcra* no reordenado y ocurran los reordenamientos V $\alpha$ J $\alpha$  primarios en timocitos DN4 y DP (22). No obstante, una vez que el gen ha reordenado y se expresa en timocitos DP, la señalización mediada por el TCR $\alpha\beta$  responsable de la selección positiva inhibe la actividad de E $\alpha$  en timocitos SP y linfocitos T  $\alpha\beta$  (94).

Es importante destacar que la transcripción del gen *Tcra* reordenado en linfocitos T  $\alpha\beta$  es igual o incluso mayor que la del gen *Tcra* no reordenado en timocitos IDP (Figura 12C). Varias posibilidades podrían explicar este hecho (94). Una posibilidad es que, aunque inhibido, E $\alpha$  sea suficiente para la activación eficiente de los promotores de los segmentos V $\alpha$  del gen *Tcra* reordenado, debido a que en la nueva estructura génica derivada de la recombinación V $\alpha$ J $\alpha$ , los promotores se encuentran mucho más cerca del *enhancer*. Otra posibilidad es que E $\alpha$  podría necesitar de la colaboración de otro elemento para la alta transcripción normal del gen *Tcra* reordenado. Una tercera posibilidad es que E $\alpha$  no sea el elemento responsable de la transcripción del gen *Tcra* reordenado en timocitos SP y linfocitos T  $\alpha\beta$ , pudiendo existir otro *enhancer* responsable de esta función.



**Figura 12. Inhibición de Ea introducido en el locus *Tcrb* en timocitos SP y linfocitos T de ratones *Knock-In* (EaKI). (A) Transcripción dependiente de EaKI de los segmentos Vβ11, Vβ12 y Vβ13 en timocitos DP de ratones *Rag2*<sup>-/-</sup> x transgén *Tcrb* (Rxβtg) y EaKI *Rag2*<sup>-/-</sup> x transgén *Tcrb* (EaKI Rxβtg) y linfocitos T αβ de ratones transgénicos del gen *Tcrb* reordenado (βtg) y EaKI transgénicos del gen *Tcrb* reordenado (Ea βtg). (B) Transcripción dependiente de EaKI del segmento Vβ13 en timocitos DP de ratones EaKI Rxβtg, timocitos DP, SP y linfocitos T αβ de ratones EaKI βtg y βtg. (C) Transcritos totales Ca del gen *Tcra* no reordenado en timocitos DP de ratones Rxβtg, y del gen *Tcra* reordenado en timocitos DP y linfocitos T αβ de ratones βtg. Los transcritos indicados se analizaron por RT-qPCR. Los experimentos se realizaron por duplicado de tres experimentos independientes. Los datos fueron normalizados con la transcripción de *Actb*. Se representa la media de los resultados +/- SEM. El valor estadístico *p* se representa por asteriscos (\*<0,05, \*\*<0,005 y \*\*\*<0,0005). Figura tomada de (del Blanco et al., 2015).**

## HIPÓTESIS Y OBJETIVOS



## HIPÓTESIS

Ea, a pesar de estar fuertemente inhibido, es necesario para la transcripción del gen *Tcra* reordenado en linfocitos T  $\alpha\beta$ .

## OBJETIVOS

1. Análisis de la función de Ea en la transcripción del gen *Tcra* reordenado en linfocitos T  $\alpha\beta$  de ratón.
2. Análisis de la función de Ea en la transcripción del gen *TCRA* reordenado en células humanas: generación de células Jurkat deficientes de Ea.
3. Detección de posibles regiones relevantes para la transcripción del gen *Tcra* por ATAC-seq en bases de datos.
4. Análisis de las interacciones entre regiones del gen *Tcra* en timocitos DP y linfocitos T  $\alpha\beta$  por 4C-seq.



# MATERIALES Y MÉTODOS



## MATERIALES Y MÉTODOS

### 1. Líneas celulares

Las células Jurkat son una línea celular de origen humano que fue establecida por Ulrich Schneider, Hans Schewenk y George Bornkamm en el año 1.977. Estas células fueron obtenidas de un niño de 14 años que sufría una leucemia linfoblástica aguda de células T (*T-cell Acute Lymphoblastic Leukemia*, T-ALL) (142). Para el mantenimiento de esta línea celular se ha utilizado medio RPMI-1640 (*Roswell Park Memorial Institute medium*) suplementado con suero fetal bovino previamente inactivado 30 minutos a 56°C (*inactivated Fetal Bovine Serum*, iFBS) al 10%, 2 mM glutamina y 100 unidades/mL de penicilina/estreptomicina (Gibco/Life Technologies). Las células se mantuvieron en un incubador a 37°C y al 5% de CO<sub>2</sub>.

### 2. Cepas de ratón

Para la elaboración del presente trabajo se han utilizado las siguientes cepas de ratón: *Rag2*<sup>-/-</sup>, *Rag2*<sup>-/-</sup> x  $\beta$ tg, *Rag2*<sup>-/-</sup> x *Ea*<sup>-/-</sup>, *Rag2*<sup>-/-</sup> x  $\beta$ tg x *Ea*<sup>-/-</sup>, *Rag2*<sup>-/-</sup> x OT-1, *Rag2*<sup>-/-</sup> x DO11.10 y *Ea*<sup>-/-</sup>.

Los ratones *Rag2*<sup>-/-</sup> son deficientes en la proteína RAG2 que es esencial para que ocurra la recombinación V(D)J y la generación de linfocitos B y T. Estos ratones tienen un bloqueo total de los timocitos en el estadio DN3a debido a la ausencia de reordenamientos de los genes que codifican el TCR (139). Los ratones *Rag2*<sup>-/-</sup> x  $\beta$ tg son ratones *Rag2*<sup>-/-</sup> que además son transgénicos para una cadena reordenada productivamente del locus *Tcrb*, lo que les permite expresar una cadena TCR $\beta$  funcional y diferenciar a los timocitos desde el estadio DN3a hasta el estadio IDP debido a la expresión del pre-TCR (140). Los ratones *Rag2*<sup>-/-</sup> x OT-1 son ratones *Rag2*<sup>-/-</sup>, cuyos linfocitos T expresan un TCR $\alpha\beta$  transgénico que reconoce un péptido de ovoalbúmina en el contexto del MHC- I y, por lo tanto, expresan CD8 (143). Los ratones *Rag2*<sup>-/-</sup>

$\alpha\beta$  x DO11.10 son ratones *Rag2*<sup>-/-</sup>, cuyos linfocitos T expresan un TCR $\alpha\beta$  transgénico que reconoce un péptido de ovoalbúmina en el contexto del MHC-II y, por lo tanto, expresan CD4 (144). Los ratones *E $\alpha$* <sup>-/-</sup> son ratones deficientes del fragmento de 1-kb que incluye *E $\alpha$ core* y los dos *E $\alpha$ CBE* (22).

Estas cepas de ratón fueron obtenidas de Taconic Farms (*Rag2*<sup>-/-</sup>, *Rag2*<sup>-/-</sup> x  $\beta$ tg, *Rag2*<sup>-/-</sup> x OT1 y *Rag2*<sup>-/-</sup> x DO11.10) y del investigador que generó el modelo *E $\alpha$* <sup>-/-</sup> (Barry P. Sleckman, University of Alabama at Birmingham, EE.UU.), y fueron mantenidas en condiciones libres de patógenos en la zona de barrera de la Unidad de Experimentación Animal del Instituto de Parasitología y Biomedicina López-Neyra (IPBLN). El uso de estos animales se adhirió a las directrices bioéticas del CSIC y la Junta de Andalucía. También se han utilizado ratones normales (*Wild Type*, WT) de la cepa C57BL6 obtenidos comercialmente de los laboratorios Charles River.

### 3. Obtención de timocitos y linfocitos T $\alpha\beta$ de ratón

Los timos de ratón fueron aislados y sus células disgregadas en 10 mL de medio RPMI-1640 completo con 10% de iFBS mediante presión del tejido sobre una malla de nylon de 200  $\mu$ m, para después filtrar la suspensión de células a través de una malla de 40  $\mu$ m y centrifugarla durante 5 minutos a 1.500 revoluciones por minuto (rpm).

Los linfocitos T  $\alpha\beta$  fueron aislados de bazo y ganglios linfáticos de ratón. Los tejidos fueron disgregados de la misma forma que los timos y sus células filtradas y centrifugadas de la misma forma que los timocitos. Para la eliminación de los eritrocitos presentes en la preparación fue necesario hacer una incubación con tampón ACK (*Ammonium-Chloride-Potassium*) que contiene 155 mM de cloruro amónico, 10 mM de bicarbonato potásico y 0,1 mM de EDTA sódico (etiléndiaminotetraacetato de sodio) en una proporción de 1 mL de solución ACK por ratón durante 1 minuto en hielo. Después las células se diluyeron 10 veces con medio completo y se centrifugaron. En el caso de células provenientes de los ratones *Rag2*<sup>-/-</sup> x OT-1 y *Rag2*<sup>-/-</sup> x DO11.10 no se necesitaron más pasos de purificación porque estos ratones carecen de otros linfocitos distintos a los que expresan el TCR $\alpha\beta$  transgénico por carecer de

RAG2, por los que la población de linfocitos T obtenida es prácticamente pura según se verificó por citometría de flujo (FACScalibur, Becton-Dickinson) con un anticuerpo específico para la cadena TCR $\beta$  acoplado con isotiocianato de fluoresceína (FITC) (ver siguiente sección para ver el método de tinción de células para citometría de flujo). En el caso de la purificación de linfocitos T  $\alpha\beta$  procedentes de ratones WT se necesitaron pasos adicionales para enriquecer la suspensión celular en linfocitos T  $\alpha\beta$  mediante el uso de columnas de lana de nylon que permiten la depleción de las células con mayor adhesión al nylon, como es el caso de los linfocitos B y otros tipos celulares como monocitos y macrófagos, y/o mediante el uso de un separador celular por citometría de flujo. Para la preparación de las columnas de nylon se usaron jeringas de 10 mL con 1 g de lana de nylon por columna. Estas columnas fueron preincubadas durante 1 hora a 37°C con medio RPMI-1640 completo antes de la incubación de las células durante otra hora para permitir la adhesión de las células más adherentes al nylon. El grado de enriquecimiento de las células T  $\alpha\beta$  después de este procedimiento se analizó mediante citometría de flujo como se ha indicado anteriormente. La purificación celular de linfocitos T  $\alpha\beta$  o V $\alpha$ 2 mediante el uso de un separador celular basado en citometría de flujo se explica en el siguiente apartado.

#### 4. Purificación celular mediante citometría de flujo

El separador celular por citometría de flujo (Aria III, Becton Dickinson) se ha utilizado en dos tipos de experimentos:

1. En la purificación de linfocitos T  $\alpha\beta$  y V $\alpha$ 2 de ratón a través de la tinción de los linfocitos con un anticuerpo para TCR $\beta$  acoplado a FITC o un anticuerpo para V $\alpha$ 2 acoplado a ficoeritrina (*Phycoerythrin*, PE), respectivamente. Estos linfocitos se resuspendieron en un tampón de tinción: HBSS (*Hank's Balanced Salt Solution*) o PBS (*phosphate buffered solution*) con 2% BSA (*Bovine Serum Albumin*) y 0,1% de azida sódica y se incubaron con los anticuerpos correspondientes a una dilución 1/100 durante 30 minutos en hielo y en oscuridad. Las células se lavaron tres veces en tampón de tinción.

2. En la separación de las células Jurkat transfectadas, las cuales expresan la proteína verde fluorescente (*Green Fluorescent Protein*, GFP) durante unas 48-72 horas, de las no transfectadas que no expresan GFP. Estas células fueron tratadas en condiciones de esterilidad y en ausencia de azida sódica para mantenerlas vivas. En el tubo de recogida se pusieron 500  $\mu$ L de FBSi, 500  $\mu$ L de medio completo y 20 $\mu$ L de penicilina/estreptomicina 100X. Después estas células fueron diluidas en medio RPMI-1640 completo y se cultivaron a 37°C y 5% de CO<sub>2</sub>.

## 5. Extracción de ARN

Para la extracción del ARN de los timocitos (provenientes de los timos de ratones) y de los linfocitos T  $\alpha\beta$  (provenientes de los bazos y ganglios linfáticos de ratones) se utilizó el método de extracción basado en tiocianato de guanidinio, fenol y cloroformo, mediante el uso del reactivo comercial TriFast (peqGold Trifast, Peqlab). Para ello, se homogenizó el precipitado de las células a un máximo de 20 millones de células por 100  $\mu$ L de reactivo. Una vez obtenidos los lisados, las muestras pudieron procesarse directamente o mantenerse congeladas a -80°C antes de proceder con el aislamiento del ARN.

Para procesar las muestras guardadas a -80°C, se dejaron 5 minutos a temperatura ambiente (TA). A continuación, se añadió 1/5 del volumen inicial de cloroformo, que para 100  $\mu$ L de muestra correspondieron a 20  $\mu$ L. Las muestras se homogenizaron mediante agitación en vórtex durante 15 segundos, se incubaron 10 minutos a TA y después de centrifugaron 5 minutos a 12.000 g a TA en una minifuga (5424R, Eppendorf). Esto permitió la separación de las muestras en tres fases, una superior incolora donde se encuentra el ARN, una interfase blanquecina donde están las proteínas y una fase inferior de color donde está el ADN. La fase superior acuosa con el ARN se pasó a otro tubo de 1,5 mL con cuidado de que no se mezclaran las fases. Se añadió 1/2 del volumen inicial de isopropanol, que para 100  $\mu$ L del reactivo Trifast correspondieron a 50  $\mu$ L y se incubó durante 1 hora en hielo. El isopropanol permite que el ARN precipite. Después se centrifugaron las

muestras durante 10 minutos a 12.000 g a 4°C. Se retiró el sobrenadante y se añadió etanol 70% frío en relación 1/1 respecto al volumen inicial de la muestra para lavar los precipitados de ARN. Se centrifugaron las muestras a 7.500 g durante 5 minutos a 4°C, se eliminó el etanol y se secaron los precipitados al aire. Después se resuspendió cada precipitado en 20 µL de agua con dietilpirocarbonato (DEPC). Se calentaron las muestras en un termobloque durante 2 minutos a 60°C para terminar de disolver el precipitado de ARN y después se guardaron las muestras a -80°C. Para medir la cantidad de ARN se utilizó el espectrofotómetro NanoDrop 2000 (Thermo-Fisher Scientific).

## 6. Retrotranscripción (RT)

La retrotranscripción se utilizó para obtener el ADN complementario (ADNc) a partir del ARN total extraído de timocitos y linfocitos. Para ello se ha utilizado el kit comercial “PrimeScript RT reagent with gDNA eraser (Perfect Real Time)” de TAKARA. Se siguió el protocolo recomendado, el cual consta de dos pasos: el primero para la eliminación del ADN genómico mediante un tratamiento con DNaseI y el segundo es la retrotranscripción propiamente dicha. Una ventaja de este kit es que el tampón del segundo paso inhibe a la DNaseI sin necesidad de pasos adicionales para inactivar esta enzima. El volumen resultante fue de 20 µL de ADNc a los que se añadieron otros 80 µL de agua MilliQ para obtener un volumen final de 100 µL.

## 7. PCR cuantitativa (qPCR)

La reacción en cadena de la polimerasa cuantitativa (qPCR) se ha realizado en los equipos CFX96 Real Time System (Bio-Rad) con la utilización del software Bio-Rad CFX Manager. Se utilizaron placas de 96 pocillos (Bio-Rad). El volumen final de cada reacción fue de 10 µL: 4 µL de la muestra de ADNc, 5 µL de SYBR Green iTaq Universal SYBR Green Supermix (Bio-Rad), 0,5 µL del oligonucleótido correspondiente a la hebra superior del ADN (*Forward*, F) a 10mM y 0,5 µL del oligonucleótido correspondiente a la hebra inferior del ADN (*Reverse*, R) a 10mM (Tabla 1).

El protocolo de la reacción de qPCR utilizado fue el siguiente: un paso de 95°C durante 7 minutos seguido de 45 ciclos de 3 pasos:

1. 95°C durante 30 segundos
2. 59,5°C durante 45 segundos
3. 72°C durante 30 segundos

Por último, se añadió un último paso a 95°C durante 1 minuto.

Los cálculos de la expresión relativa de los transcritos se realizaron mediante el método  $\Delta\Delta Ct$ . Los valores se normalizaron con los valores obtenidos para los transcritos de cada muestra con su correspondiente valor de los transcritos del gen de  $\beta$ -ACTINA de ratón o humano (*Actb* o *ACTB*).

Los oligonucleótidos usados para las PCR se muestran en la tabla 1:

- **Oligonucleótidos para la cuantificación de transcritos**

Oligonucleótidos	Secuencia 5' → 3'	F/R	Modelo
<b><i>Actb</i></b>	ACACCCGCCACCAGTTC	F	Murino
	TACAGCCCGGGGAGCAT	R	
<b>C<math>\alpha</math></b>	CCTGGAGCAACCAGACAA	F	Murino
	GGCTACTTTCAGCAGGAGGA	R	
<b><i>Traj61-C<math>\alpha</math></i></b>	GCCATGACTGGAAGACTCAT	F	Murino
	AAGTCGGTGAACAGGCAGAG	R	
<b><i>Traj58-C<math>\alpha</math></i></b>	TGGGTCTAAGCTGTCATTTGG	F	Murino
	AAGTCGGTGAACAGGCAGAG	R	
<b><i>Trav14-C<math>\alpha</math></i></b>	TGGAGACTCAGCCACCTACT	F	Murino
	CACAGCAGGTTCTGGGTT	R	
<b>OT-I VaJa transgén</b>	CAGCAGCAGGTGAGACAAAGT	F	Murino
	GGCTTTATAATTAGCTTGGTCC	R	
	CAGGAGGGATCCAGTGCCAGC	F	Murino

<b>DO11.10</b> transgén	<b>V<math>\alpha</math>J<math>\alpha</math></b>	TGGCTCTACAGTGAGTTTGGT	R	
<b>DO11.10</b> transgén	<b>V<math>\alpha</math>C<math>\alpha</math></b>	CTTCTCCCTGCACATCACAG	F	Murino
		AGAGGGTGCTGTCCTGAGAC	R	
<b>DO11.10</b> transgén	<b>V<math>\beta</math>C<math>\beta</math></b>	GGGCTGAGGCTGATCCATTA	F	Murino
		AGGAGACCTTGGGTGGAGTC	R	
<b>C<math>\alpha</math></b>		CCTGACCCTGCCGTGTACCAGC	F	Humano
		GTCCATAGACCTCATGTCTAGCACAG	R	
<b>Jurkat</b> <b>TRAV18-</b> <b>TRAJ13</b>	<b>TRAV18-</b>	AAGGCATCAACGGTTTTGAG	F	Humano
		GAGTCTGGTCCCTGATCCAA	R	
<b>ACTB</b>		CGCGAGAAGATGACCCAGA	F	Humano
		TCACGATGCCAGTGGTACG	R	

- **Oligonucleótidos con las gARN utilizados para ser clonados en el plásmido pX458 y generar clones Jurkat E $\alpha$ <sup>+/-</sup> y E $\alpha$ <sup>-/-</sup>:** Las secuencias en color indican la secuencia genómica reconocida por la guía.

<b>gARN1</b>	AAAGGACGAAACACCG <b>CCACTTCCCTCCAGGTGTTT</b> GTTTTAGAGCTAGAAATAGCAAG
<b>gARN2</b>	TTCTAGCTCTAAAAC <b>CCACATCATACACTCGTGCC</b> GGATCCAAGGTGTCTCATAC
<b>gARN3</b>	AAAGGACGAAACACCG <b>TCAGTAGCCACAAGTAGAAC</b> GTTTTAGAGCTAGAAATAGCAAG
<b>gARN4</b>	TTCTAGCTCTAAAAC <b>CCTCAGTTTCCCGTGGGAC</b> GGATCCAAGGTGTCTCATAC
<b>U6-F</b>	GAGGGCCTATTTCCCATGATTCC

- **Oligonucleótidos para la detección de los clones Jurkat E $\alpha$ <sup>+/-</sup> y E $\alpha$ <sup>-/-</sup> por PCR**

<b>hTα3-Tα4</b>	AAAATACTGAGTTAGAGATAG
	CGGTGCAAGTCACCGA
<b>Delección Eα F2/R2</b>	AGGCAGGCACACATCTATCC
	AGAGGCAAACAGTGCCCTAA

- **Oligonucleótidos usados en los experimentos de 4C seq (*r=reading; nr=nonreading*):** Las secuencias en **negrita** indican la región genómica de hibridación en los experimentos de PCR inversa. Las secuencias en **rojo** indican las secuencias para diferenciar las distintas muestras en la secuenciación masiva (*barcodes*).

<b>Trav1 HindIII-r</b>	<b>TACACGACGCTCTTCCGATCTNNNNGGTCAACTGCCCCATGCTAAGC</b>
<b>Trav1 HindIII-nr</b>	ACTGGAGTTCAGACGTGTGCTCTTCCGATCT <b>GACGGTGGGAGGTAAGTTCCA</b>
<b>Trav1/2 HindIII-r</b>	TACACGACGCTCTTCCGATCTNNNNGATGGAGCTGTAGATGACTATATTAAGTG <b>ACAGAAG</b>
<b>Trav1/2 HindIII-nr</b>	ACTGGAGTTCAGACGTGTGCTCTTCCGATCT <b>CCACAAAGAAAACTAAACCCCAA</b> <b>ACATTATACACC</b>
<b>Cα peak HindIII-r</b>	TACACGACGCTCTTCCGATCTNNNNAAGACAGACCCTGCGAAGCTT
<b>Cα peak HindIII-nr</b>	ACTGGAGTTCAGACGTGTGCTCTTCCGATCT <b>GCCCTCTCTGTATCTCAGGGGAAT</b> <b>T</b>
<b>Eα HindIII-r</b>	TACACGACGCTCTTCCGATCTNNNNGGCAAAGCAGGGTCTGTAAGCT
<b>Eα HindIII-nr</b>	ACTGGAGTTCAGACGTGTGCTCTTCCGATCT <b>CCTTTCCCATCAGTCCTTTCTCCCTA</b> <b>G</b>
<b>Trav1 DpnII-r</b>	TACACGACGCTCTTCCGATCTNNNNGGTGAGGTGGTGTGGGGA

<b>Trav1 DpnII-nr</b>	ACTGGAGTTCAGACGTGTGCTCTCCGATCT <b>GGTAAGTTCCA</b> ACTTCATAGAGAA <b>CATGGCTA</b>
<b>Trav1/2 DpnII-r</b>	TACACGACGCTCTCCGATCTNNNN <b>TTATGACTTTTTATTTTCATGGTTGATC</b>
<b>Trav1/2 DpnII-nr</b>	ACTGGAGTTCAGACGTGTGCTCTCCGATCT <b>GACTGTGTAAGAAGTGACCAA</b>
<b>Cα peak DpnII-r</b>	TACACGACGCTCTCCGATCTNNNN <b>CATCTCTCTATGTTCTCTGTCTTCTGCTGGA</b>
<b>Cα peak DpnII-nr</b>	ACTGGAGTTCAGACGTGTGCTCTCCGATCT <b>GTGAATATGTCTTGCTGTCTTTACGGACT</b>
<b>Eα DpnII-r</b>	TACACGACGCTCTCCGATCTNNNN <b>GTC</b> ACTGGACTCCAGACTTGTGTGATC
<b>Eα DpnII-nr</b>	ACTGGAGTTCAGACGTGTGCTCTCCGATCT <b>CAGGTTTCAGGTTCTGCGACTTCCTAG</b>
<b>Universal A</b>	AATGATACGGCGACCACCGAGATCTACAC <b>TATAGCCT</b> ACACTCTTCCCTACACGACGCTCTCCGATCT
<b>Universal B</b>	AATGATACGGCGACCACCGAGATCTACAC <b>ATAGAGGC</b> ACACTCTTCCCTACACGACGCTCTCCGATCT
<b>Index 1</b>	CAAGCAGAAGACGGCATAACGAGAT <b>CGTGAT</b> GTGACTGGAGTTCAGACGTGTGCT
<b>Index 2</b>	CAAGCAGAAGACGGCATAACGAGAT <b>ACATCG</b> GTGACTGGAGTTCAGACGTGTGCT
<b>Index 3</b>	CAAGCAGAAGACGGCATAACGAGAT <b>GCCTAA</b> GTGACTGGAGTTCAGACGTGTGCT
<b>Index 4</b>	CAAGCAGAAGACGGCATAACGAGAT <b>TGGTCAG</b> GTGACTGGAGTTCAGACGTGTGCT
<b>Index 5</b>	CAAGCAGAAGACGGCATAACGAGAT <b>CACTGT</b> GTGACTGGAGTTCAGACGTGTGCT
<b>Index 6</b>	CAAGCAGAAGACGGCATAACGAGAT <b>ATTGGC</b> GTGACTGGAGTTCAGACGTGTGCT
<b>Index 7</b>	CAAGCAGAAGACGGCATAACGAGAT <b>GATCTG</b> GTGACTGGAGTTCAGACGTGTGCT
<b>Index 8</b>	CAAGCAGAAGACGGCATAACGAGAT <b>TCAAGT</b> GTGACTGGAGTTCAGACGTGTGCT
<b>Index 9</b>	CAAGCAGAAGACGGCATAACGAGAT <b>CTGATC</b> GTGACTGGAGTTCAGACGTGTGCT
<b>Index 10</b>	CAAGCAGAAGACGGCATAACGAGAT <b>AAGCTA</b> GTGACTGGAGTTCAGACGTGTGCT

<b>Index 11</b>	CAAGCAGAAGACGGCATAACGAGAT <b>GTAGCC</b> GTGACTGGAGTTCAGACGTGTGCT
<b>Index 12</b>	CAAGCAGAAGACGGCATAACGAGAT <b>TACAAG</b> GTGACTGGAGTTCAGACGTGTGCT
<b>Index 13</b>	CAAGCAGAAGACGGCATAACGAGAT <b>TGTTGAG</b> GTGACTGGAGTTCAGACGTGTGCT
<b>Index 14</b>	CAAGCAGAAGACGGCATAACGAGAT <b>ACGGAA</b> GTGACTGGAGTTCAGACGTGTGC T
<b>Index 15</b>	CAAGCAGAAGACGGCATAACGAGAT <b>TCTGAC</b> GTGACTGGAGTTCAGACGTGTGCT
<b>Index 16</b>	CAAGCAGAAGACGGCATAACGAGAT <b>CGGGAC</b> GTGACTGGAGTTCAGACGTGTGCT

**Tabla 1. Secuencias de los oligonucleótidos usados.**

## 8. Eliminación de E $\alpha$ mediante edición génica por CRISPR-Cas9

Esta técnica de edición génica ha sido utilizada para la delección del fragmento de 1 kb de E $\alpha$ , conteniendo E $\alpha$ core y E $\alpha$ CBE, en la línea celular Jurkat para conseguir clones E $\alpha$ <sup>+/-</sup> y E $\alpha$ <sup>-/-</sup>.

### 8.1 Generación de clones de Jurkat E $\alpha$ <sup>+/-</sup>

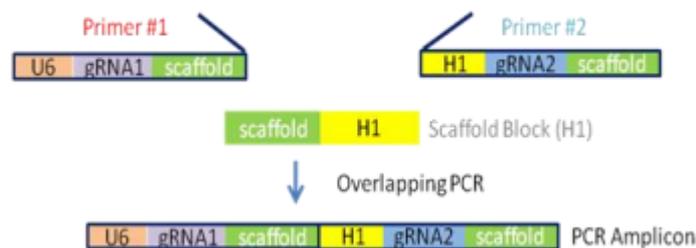
Debido a que las células Jurkat tienen células con distinta expresión de partida del receptor, previamente se realizó una selección de un clon de células Jurkat con expresión media del TCR $\alpha\beta$ . Para ello se realizó una tinción de las células con el anticuerpo FITC-TCR $\beta$  (clone T10B9) (555547, BD Bioscience) y las células positivas con expresión media se purificaron por citometría de flujo (como se detalla el apartado 2.4 de Material y Métodos), realizándose una dilución límite de las células purificadas a las 24 horas para la separación de clones. Se seleccionó el clon 4 debido a que estas células mostraron buen aspecto al microscopio y una expresión del TCR $\alpha\beta$  positiva y uniforme al ser analizadas por citometría de flujo.

#### 8.1.1) Preparación de las guías de ARN (gARN) y plásmidos utilizados

Para editar el genoma con la tecnología CRISPR/Cas9 es necesario expresar la endonucleasa CAS9 e introducir unas guías de ARN (gARN) en las células. Las guías usadas eran complementarias al ADN presente en las zonas

flanqueantes a la región de Eα a eliminar y precedidas por una secuencia denominada PAM (*Protospacer Adjacent Motif*) que es reconocida por CAS9 (Tabla 1). Las PAM de CAS9 consisten en una base nitrogenada seguida de dos guaninas (NGG). Para evitar al máximo que hubiera falsas hibridaciones de las gARN en otras regiones del genoma se utilizó el programa de diseño de gARN CHOPCHOP (<https://chopchop.rc.fas.harvard.edu/>), que genera un listado de posibles guías con sus potenciales dianas inespecíficas.

Se diseñaron los oligonucleótidos complementarios para clonar las guías seleccionadas en el plásmido de expresión para CAS9 y GFP (pSpCas9(BB)-2A-GFP), pX458, (*Addgene plasmid ID: 48138*) (145) para su posterior transfección en las células Jurkat (Tabla 1). Las guías fueron diseñadas para la eliminación del fragmento de 1-kb de Eα correspondiente a la región previamente delecionada para generar los ratones Eα<sup>-/-</sup> (22). Se diseñaron dos gARN que reconocen secuencias adyacentes al extremo 5' de Eα (gARN-5') y otras dos que reconocen secuencias adyacentes al extremo 3' (gARN-3') de Eα. Para optimizar el proceso de deleción se clonaron las dos guías, una gARN- 5' y una gARN- 3', conjuntamente (145). Para ello se usaron fragmentos de ADN conteniendo secuencias *scaffold*, U6 y/o H1, para obtener un amplicón por PCR que contiene las dos guías juntas en un único inserto para ser ligado directamente con el plásmido pX458 (Figura 13), utilizando el kit Precision Multiplex gRNA Cloning Kit (System Biosciences).



**Figura 13. Esquema del proceso de formación de las construcciones que contienen dos gARN.** Se muestra la estructura de los fragmentos de ADN utilizados para formar un amplicón conteniendo ambas guías mediante PCR.

Los oligonucleótidos (gARN-5' en morado y gARN-3' en azul) utilizados (Tabla 1) conteniendo las secuencias del promotor U6 (en naranja), *scaffold* (en verde) y promotor H1 (en negro) tuvieron la siguiente estructura:

1. *Forward*:

**AAAGGACGAAACACCG**NNNNNNNNNNNNNNNNNNNNNNNNNNNNNN**GTTT**  
**TAGAGCTAGAAATAGCAAG**

2. *Reverse*:

**TTCTAGCTCTAAAAC**NNNNNNNNNNNNNNNNNNNNNNNNNNNNNN**GGATCC**  
**AAGGTGTCTCATAC**

Para obtener la construcción completa fue necesario realizar una PCR en la que se usó el bloque *scaffold*-H1 que viene en el kit comercial. La reacción para la PCR utilizada fue la siguiente:

- |   |        |
|---|--------|
| 1. Tampón 5x Phusion                    | 10 µL  |
| 2. 10mM dNTPs                           | 1 µL   |
| 3. 10 µM oligonucleótido <i>Forward</i> | 2,5 µL |
| 4. 10 µM oligonucleótido <i>Reverse</i> | 2,5 µL |
| 5. Bloque <i>scaffold</i> -H1           | 0,5 µL |
| 6. DMSO                                 | 1,5 µL |
| 7. Phusion ADN Polimerasa               | 0,5 µL |
| 8. Agua milliQ hasta 50 µL              |        |

La PCR se realizó en las condiciones que se muestran en la tabla 2.

Paso	Temperatura	Tiempo
Desnaturalización inicial	98°C	3 minutos
30 ciclos	98°C	5 segundos
	51°C	15 segundos
	72°C	15 segundos
Extensión final	72°C	10 minutos
Mantenimiento	4°C	∞

**Tabla 2. Condiciones de la PCR para obtener la construcción del bloque U6-gARN-5'/3-H1-gARN-3'.**

Se clonaron las cuatro gARN que flanquean Eα por parejas (gARN-5' : gARN1 o gARN3 y gARN3' : gARN2 o gARN4) en las cuatro posibles combinaciones: gARN1-gARN2, gARN1-gARN4, gARN3-gARN2 y gARN3-

gARN4. Una vez se obtuvieron los cuatro productos de PCR se tomaron 2,5  $\mu\text{L}$  de cada muestra para ser analizadas en un gel de agarosa al 2%. Las bandas resultantes se purificaron usando el kit “*GeneJET Gen Extracción kit*” (KO691, Thermo-Fisher Scientific) y posteriormente se secuenciaron usando el oligonucleótido U6-F (Tabla 1).

Los amplicones resultantes fueron insertados en el plásmido px458. Este plásmido permitió el clonaje de los bloques con las gARN mediante un único sitio para el enzima de restricción BbsI para la expresión de las guías junto con CAS9 y GFP. Para ello se realizó una digestión con la enzima BbsI (BbsI, RO539, *New England Biolabs*) a 37°C toda la noche para asegurarse de que se produce una linealización completa del plásmido. Al día siguiente se inactivó la enzima aumentando la temperatura hasta los 65°C según indican las instrucciones del fabricante y se purificó el plásmido digerido después de ser sometido a electroforesis y purificación desde agarosa.

Para introducir cada amplicón con las dos gARN en el plásmido px458 se realizó una reacción de fusión siguiendo las instrucciones recomendadas por el kit (Precision X Multiplex gRNA Cloning Kit) de la siguiente manera:

1. Vector linealizado (10-100 ng/ $\mu\text{L}$ )	1 $\mu\text{L}$
2. Inserto de PCR (20-200 ng/ $\mu\text{L}$ )	1 $\mu\text{L}$
3. dH <sub>2</sub> O	x $\mu\text{L}$ hasta 50
4. 5x master mix	2 $\mu\text{L}$

Esta reacción se dejó 5 minutos a TA, seguido de 10 minutos en hielo. Se hicieron en paralelo reacciones de control positivo (con un vector y un inserto de 500 pb que vienen con el kit) y negativo (con 10-100 ng/ $\mu\text{L}$  de nuestro vector linealizado, pero sin inserto de PCR).

El siguiente paso fue realizar una transformación bacteriana para aislar los plásmidos obtenidos. Para crecer las bacterias en placas de agar con ampicilina se plaquearon 50  $\mu\text{L}$  y 100  $\mu\text{L}$  de bacterias *E. coli* DH5 $\alpha$  transformadas con cada construcción. Se crecieron 3 colonias obtenidas de cada transformación y posteriormente se purificaron los plásmidos tras crecer las bacterias en 5 mL de medio LB con ampicilina a 100  $\mu\text{g}/\text{mL}$  utilizando el kit “QUIaprep spin miniprep kit” (Qiagen). El plásmido purificado fue

secuenciado con el oligonucleótido U6-F (Tabla 1) en la Unidad de Genómica del IPBLN.

Una vez comprobadas las secuencias de los insertos, se hizo una purificación de los plásmidos a mayor escala desde un cultivo de bacterias crecidas en 200 mL de medio LB con 100 µg/mL de ampicilina y columnas de “Midiprep”. Estos plásmidos fueron utilizados para la delección de Eα.

### 8.1.2) Transfección de células Jurkat para la delección de Eα

Se transfectaron  $10^6$  células para cada plásmido px458 conteniendo cada una de las combinaciones de dos gARN flanqueantes de Eα. También se realizó un control de transfección sin plásmido. Las transfecciones se realizaron utilizando el electroporador Nucleofector y el kit de transfección Cell Line Nucleofector Kit V (VCA-1003, Lonza). Para ello las células se centrifugaron durante 5 minutos a 1.500 rpm a 4°C. Mientras tanto se preparó una placa de 6 pocillos con 4,5 mL de medio completo en 4 pocillos, uno para cada punto de transfección de las cuatro combinaciones de dos gARN flanqueantes de Eα. Las células se resuspendieron en 82 µL de tampón V y 18 µL del suplemento del kit. Tras la resuspensión, las células se introdujeron en las cubetas de electroporación del kit sumergidas en hielo. La electroporación se realizó con el programa X-005 (que tras unas pruebas fue el programa que mayor eficiencia de transfección ofrecía). Rápidamente en la cabina de flujo laminar se añadieron 500 µL de FBSi a cada cubeta y la suspensión celular se pasó a un pocillo de la placa preparada anteriormente, la cual se incubó a 37°C y 5% de CO<sub>2</sub>.

Tras 24 horas de incubación, se recogieron los 4,5 mL de cada pocillo, se centrifugaron las células y se resuspendieron en 500 µL de tampón de tinción para células vivas estéril (ver el apartado 4 de esta sección). Se prepararon también los tubos de recogida para citometría de flujo con 500 µL de iFBS, 500 µL de medio completo y 20 µL de penicilina/estreptomicina 10.000 unidades/mL. El 3% de las células GFP<sup>+</sup> con mayor intensidad de fluorescencia fueron separadas por citometría de flujo en el separador celular. Las células recogidas se centrifugaron y se incubaron en una placa de 6 pocillos con 5 mL de medio completo/pocillo a 37°C y 5% de CO<sub>2</sub>. Al día siguiente las

células se contaron y se diluyeron al límite (60 células/12 mL de medio completo aproximadamente) para plaquear <1 célula/pocillo en pocillos de placas de 96 pocillos de fondo plano con 100 µL de medio RPMI-1640 completo/pocillo. En la preparación del medio se utilizó también medio condicionado proveniente del cultivo de células Jurkat en estado exponencial, para favorecer el crecimiento de las células, en una proporción de 40% de medio condicionado y 60% de medio fresco. Las placas con los clones se dejaron expandiendo a 37°C y 5% de CO<sub>2</sub> durante 2-3 semanas, añadiendo 100 µL adicionales de medio completo fresco a los 5 días y renovándolo cada 3-5 días según fuese necesario. Una semana después del plaqueo se revisaron las placas en busca de los clones. Cuando al menos el 60% de la superficie de los pocillos estuvo cubierta de células se realizaron réplicas de las placas pasando 40 µL de cada pocillo a una placa nueva con 100 µL de medio fresco por pocillo. Los 160 µL restantes de las placas originales se usaron para aislar el ADN de esos clones y proceder a su genotipado.

### 8.1.3) Genotipado de clones Jurkat

El sobrenadante de los pocillos fue aspirado con vacío y la punta de una micropipeta de 200 µl. Las células quedaron adheridas al fondo del pocillo y fueron lisadas con 20 µL de un tampón de lisis con 0,4 µg de proteinasa K (32181, Intron Biotechnology) en solución TNES: 100 mM, Tris pH:8, 5 mM EDTA, 200 mM de cloruro de sodio y 0,2% de SDS (*sodium dodecyl sulphate*). Los lisados se transfirieron a tubos de 0,2 mL para incubarlos en un termociclador durante 30 minutos a 65°C y 10 minutos a 98°C. Después de una breve centrifugación, se añadieron 180 µL de agua Milli-Q para tener un volumen final de 200 µL. Estos lisados fueron usados para realizar genotipados por qPCR para la detección de la presencia de Eα con los oligonucleótidos Eαdel-F2/R2 y hTα3/Tα4 (100,131).

Las secuencias de los oligonucleótidos usados para los genotipados se muestran en la tabla 1, así como una representación de su lugar de anillamiento y dirección de amplificación (Figura 14), representando en color rojo la zona que contiene Eα.



**Figura 14. Lugar de anillamiento y orientación de los oligonucleótidos utilizados sobre la secuencia genómica humana.** En color rojo se muestra la zona de delección de Ea. Se utilizaron los oligonucleótidos EadelF2/R2 y Ta3/Ta4 (Tabla 1).

Para la realización de las qPCR de los genotipados se usó 1  $\mu$ L de cada lisado en un volumen total de 10  $\mu$ L de reacción que llevaba 5  $\mu$ L de la mezcla “Bio-rad iTaqUniversal Sybr Green Supermix” 2X (1725122, Bio-Rad) y 0,5  $\mu$ L de cada oligonucleótido 10 mM. Se realizaron duplicados de qPCR para cada lisado y se analizaron en el termociclador Bio-Rad CFX-96 System en las condiciones mencionadas en el apartado correspondiente a la técnica qPCR.

El ADN genómico de los lisados celulares fue purificado mediante una extracción con fenol (0945-400mL, Amresco), fenol/cloroformo y cloroformo (1.02445.1000, Merck) seguido de una precipitación en etanol a 4°C y disolviéndose finalmente en 100  $\mu$ L de agua Milli-Q. Para visualizar la delección del *enhancer* en geles de agarosa se realizaron reacciones de PCR convencionales en reacciones de 50  $\mu$ l con 8  $\mu$ L de ADN genómico, 0,4  $\mu$ L de la polimerasa de ADN Dream Taq (EP0702, Thermo Fisher Scientific), 5  $\mu$ L de tampón 10X, 1  $\mu$ L de dNTPs 10 mM y 4  $\mu$ L a 10 mM de cada oligonucleótido Eadel-F2/R2 que flanquean a Ea. Las condiciones de la PCR fueron las que se muestran a continuación en la tabla 3.

2  $\mu$ L de cada PCR fue analizados en geles de agarosa al 1% y el ADN genómico de los clones con la delección fue secuenciado con los oligonucleótidos Eadel-F2 y R2. Los clones Ea<sup>+/-</sup> obtenidos fueron testados por citometría de flujo usando el anticuerpo para FITC-TCR $\beta$  humano y clasificados en 2 grupos, los que expresaban y los que no expresaban el TCR $\alpha\beta$ .

Paso	Temperatura	Tiempo
------	-------------	--------

<b>Desnaturalización inicial</b>	95°C	3 minutos
<b>40 ciclos</b>	95°C	30 segundos
	60°C	30 segundos
	72°C	1 minuto
<b>Extensión final</b>	72°C	10 minutos
<b>Mantenimiento</b>	4°C	∞

**Tabla 3. Pasos de PCR convencional.**

## 8.2 Generación de clones Jurkat E $\alpha^{-/-}$

Para generar clones E $\alpha^{-/-}$  se partió de  $10^6$  células de uno de los clones E $\alpha^{+/-}$ , clon IV.16, con expresión del TCR $\alpha\beta$  y obtenido usando las gARN más externas (gARN3 y gARN4). Para deletar E $\alpha$  en el alelo intacto se usó el plásmido pX458 con las guías más internas, gARN1 y gARN2, ya que así solo podían reconocer el alelo que contiene el *enhancer* restante. Se usó la misma estrategia de transfección descrita mediante el uso del electroporador Nucleofector (Lonza), siguiendo los pasos descritos anteriormente. 24 horas después de la transfección se sortearon el 3% de las células GFP<sup>+</sup> con más brillo, que fueron cultivadas un día para que se recuperaran antes de realizar diluciones límite en placas de 96 pocillos de fondo plano usando medio condicionado y medio fresco en una proporción 50/50.

Los clones que crecieron fueron lisados y analizados para localizar la ausencia de E $\alpha$  mediante qPCR con los oligonucleótidos Ta3/Ta4 siguiendo los mismos pasos y en las mismas condiciones descritas anteriormente. Los clones con la delección fueron analizados para la expresión del TCR $\alpha\beta$  realizando una tinción con el anticuerpo FITC-TCR $\beta$  humano y observando la pérdida de la expresión de TCR $\alpha\beta$  mediante citometría de flujo. La estructura exacta de la delección fue determinada mediante secuenciación de los amplicones de PCR de mayor tamaño obtenidos con ADN genómico y los oligonucleótidos Eadel-F2/R2

## 9. Transformación bacteriana y purificación de plásmidos

Se transformaron bacterias competentes *Escherichia coli* cepa DH5 $\alpha$  comerciales (XL-1-BLUE-Competent *E. coli*, Agilent Technologies) para el crecimiento del plásmido px458 (Addgene, referencia 48138) y sus derivados conteniendo las gARN de interés. Antes de empezar se descongeló el medio SOC (*Super Optimal Broth with Catabolic repressor*) y se pusieron en marcha dos termobloques con movimiento (Eppendorf), uno a 37°C y otro a 42°C. Se incubó 1  $\mu$ L (aproximadamente 1 $\mu$ g) de cada plásmido junto con las bacterias competentes durante 15 minutos. Después se sometió a las bacterias a un choque térmico de 45 segundos a 42°C, seguido de una incubación de 2 minutos en hielo (a 42°C se abren los poros y entra el plásmido, en hielo se cierran y se recuperan las bacterias del choque térmico). Después se añadieron 300  $\mu$ L de SOC. A continuación, se incubaron las bacterias durante 1 hora a 37°C en movimiento. 10  $\mu$ L y 100  $\mu$ L de bacterias transformadas fueron sembradas con un asa de Henle en placas de agar con ampicilina a 100  $\mu$ g/mL. Las placas se incubaron invertidas a 37°C durante toda la noche.

Al día siguiente se picó una colonia aislada que se sembró en un vial con 5 mL de medio LB y 100  $\mu$ g/mL de ampicilina y se incubó durante 1 - 2 horas a 37°C en movimiento (180 rpm). Este cultivo puede usarse para hacer una minipreparación de plásmido dejándolo crecer hasta el día siguiente (*miniprep*). Si lo que se quiso era una preparación a media escala (*midiprep*), cuando el vial de 5 mL se vio turbio por crecimiento bacteriano se pasó el contenido a un matraz de 2 L con 500 mL de medio LB con 100  $\mu$ g/mL de ampicilina que se dejó incubar toda la noche a 37°C a 180 rpm.

### 9.1 Midipreps

Para la realización de *midipreps* se utilizó el kit “NucleoBind Xtra Midi Prep” (Macherey-Nagel). Para ello, se centrifugaron las bacterias durante 30 minutos a 6.000 rpm a 4°C en una centrífuga J2-MC con un rotor JA-14 (Beckman). Tras terminar la centrifugación se retiró el sobrenadante y se

siguió el protocolo recomendado. Primero se resuspendió el precipitado de bacterias en 8 mL de tampón de resuspensión RES. Posteriormente se añadieron 8 mL de tampón de lisis LYS y se incubó 5 minutos a TA (el lisado adquirió un tono azul). En estos 5 minutos se preparó la columna para el aislamiento del plásmido mediante la adición de 12 mL del reactivo de equilibración EQU. Al tubo con la mezcla azul se le añadieron 8 mL de la solución de neutralización NEU y se echó el contenido del lisado a la columna lentamente por los bordes para evitar su saturación. Después que hubo pasado el lisado se añadieron 5 mL de reactivo EQU. A continuación, se añadieron 8 mL del tampón de lavado WASH para lavar el filtro. A continuación, se puso un tubo de recogida bajo la columna y se eluyó el plásmido con 5 mL de reactivo de elución ELU. Al plásmido eluido se le añadieron 3,5 mL de isopropanol para precipitarlo, y se centrifugó durante 1 hora a 4.000 rpm a 4°C en una centrífuga de mesa. Se eliminó el sobrenadante con cuidado, se lavó el precipitado con 2 mL de etanol al 70% y se volvió a centrifugar durante 10 minutos a 4.000 rpm a 4°C. Se dejó secar el precipitado y se resuspendió en 500 µL de agua MilliQ en un tubo de 1,5 mL. La concentración del plásmido se cuantificó mediante espectrofotometría en un nanodrop 2000 y se guardó a -20°C.

## 9.2 Minipreps

Con el objetivo de purificar y obtener una pequeña cantidad de plásmido para la comprobación de su correcta estructura mediante secuenciación se realizaron *minipreps*. Para la realización de estas preparaciones de plásmido a pequeña escala se usó el kit comercial “QUIaprep spin miniprep” (Qiagen).

Para ello, se centrifugaron las bacterias en tubos de 1,5 mL a 8.000 rpm a TA durante 3 minutos en una minifuga. Se eliminó el sobrenadante y se repitió este paso una vez más en los mismos tubos. El precipitado resultante de las dos centrifugaciones se resuspendió en 250 µL del tampón P1. Después se añadieron 250 µL del tampón P2 y se mezcló por inversión 4-6 veces. Tras esto se añadió el tampón N3 y se mezcló inmediatamente por inversión 4-6 veces hasta que tomó un aspecto nuboso. El siguiente paso fue centrifugar 10 minutos a 13.000 rpm a TA en una minifuga. El sobrenadante se añadió a las

columnas del kit y se realizó una centrifugación de 30-60 segundos a 13.000 rpm (*spin*) para que el líquido pasase a través de la columna. Se añadió 0,5 mL de tampón PE para lavar y se hizo un *spin* similar al anterior. Se añadió 0,75 mL de tampón PE de nuevo, se dio otro *spin* y se tiró el tubo usado con el líquido. Se puso otro tubo de 1,5 mL bajo la columna, se añadieron 50 µL de tampón de elución EB, se incubó 1 minuto y se centrifugó otro minuto a 13.000 rpm a TA en una minifuga. En el líquido que quedó en el tubo estaba el plásmido purificado.

## 10. Preparación de muestras para secuenciación

Para secuenciar el inserto con las guías en los plásmidos pX458 se utilizó el oligonucleótido U6-F (tabla 1), que hibrida justo antes de la zona que se quiere secuenciar.

- |                            |             |
|----------------------------|-------------|
| 1. Primer U6-F             | 0.7 µL      |
| 2. 500 ng de ADN           | x µL        |
| 3. H <sub>2</sub> O milliQ | hasta 12 µL |

Para la secuenciación de los clones el procedimiento fue el mismo, pero se sustituyó el oligonucleótido U6-F por los oligonucleótidos Eadel-F2 y R2 (Tabla 1), por lo que obtuvimos 2 secuencias (una desde cada extremo del amplicón) para cada clon seleccionado. La secuenciación Sanger por capilares se realizó en la Unidad de Genómica del IPBLN.

## 11. 4C-seq

La técnica 4C-seq (captura de conformación cromosómica circular) deriva del método 3C (captura de conformación cromosómica) que se usa para detectar todas las regiones genómicas que interactúan con una región de interés o “anzuelo”. Para ello se emplean PCR inversas y la secuenciación masiva para amplificar, identificar y cuantificar los fragmentos de ADN asociados a una región genómica determinada. Para realizar el método se siguió el protocolo detallado y descrito por de Laat, W., *et al.* (146).

Se obtuvieron  $10^7$  timocitos de ratones *Rag2<sup>-/-</sup>* que habían sido tratados 10 días antes del sacrificio con una inyección intraperitoneal con 150  $\mu$ L de PBS con 150  $\mu$ g del anticuerpo CD3 $\epsilon$  (clon 145-2C11) (100363, Biolegend). Por otro lado, se aislaron  $10^7$  linfocitos CD3<sup>+</sup> purificados por citometría de flujo desde esplenocitos totales de ratones de la cepa C57BL6.

Ambas muestras, por separado, fueron resuspendidas en 5 mL de medio RPMI-1640 con un 10% de iFBS y tratadas con 5 mL de formaldehído (15710, Electron Microscopy Sciences) al 4% recién preparado en RPMI-1640 con 10% de iFBS (quedando una concentración final de 2% de formaldehído) para entrecruzar las proteínas y los ácidos nucleicos de las células. Se mezcló la suspensión celular invirtiendo el tubo y se incubó 10 minutos a TA en movimiento. Las reacciones se pararon añadiendo glicina 1M fría, quedando a una concentración final de 0,125 M y se pasaron las células a hielo. Las células se centrifugaron a 4°C durante 5 minutos a 1.800 rpm en una y posteriormente se resuspendieron en 1 mL de PBS frío para pasarlas a tubos de 1,5 mL. Se volvieron a centrifugar las células a 4°C durante 5 minutos a 1.500 rpm y se lisaron añadiendo 1 mL de un tampón de lisis frío recién preparado (150 mM de NaCl, 50 mM de Tris pH: 7,5, 0,5% de NP-40, 1% de Triton X-100, 5 mM de EDTA e inhibidores de la proteasa “Complete EDTA Free” (Roche) e incubado 20 minutos en hielo. De esta forma se lisaron las membranas celulares, pero no sus núcleos.

Para recoger los núcleos se centrifugó el lisado a 4°C durante 5 minutos a 1.500 rpm en una centrífuga de mesa, quedando los núcleos en el fondo del tubo. Se realizó un lavado con 0,5 mL de tampón de enzima de restricción al 1,2X (RE1): tampón #2 para HindIII y tampón #3.1 para DpnII (New England Biolabs). Después se resuspendieron los núcleos en 0,5 mL del tampón de la enzima de restricción correspondiente al 1,2X, se calentó hasta 37°C y se añadió SDS al 10% quedando con una concentración final de 0,3% para lisar los núcleos. Los lisados se incubaron en movimiento a 37°C durante 1 hora y después se añadió Triton X-100 con una concentración final de 2,5% y se incubó de nuevo 1 hora a 37°C en movimiento para eliminar el efecto desnaturizante del SDS. Posteriormente, se añadieron 10  $\mu$ g de HindIII (200U) o de DnpII (100U) (R0104 y R0543, New England Biolabs) para digerir

la cromatina en una reacción que se incubó durante 3 a 4 horas a 37°C en movimiento. Se volvió a añadir las enzimas y se repitió la incubación durante toda la noche. Las enzimas se inactivaron con incubación a 65°C (HindIII) u 80°C (DnpII) durante 20 minutos.

Para ligar los productos genómicos que permanecen cercanos y generar una librería 3C, las muestras se pasaron a tubos de 50 mL y se añadieron 700 µL de tampón 10x de la ligasa T4 (660mM de Tris pH: 7,5, 50 mM de MgCl<sub>2</sub>, 50 mM de DTT y 10 mM de ATP), 10 µL de ligasa del ADN T4 (4,000 U) (MO20, New England Biolabs) y agua libre de nucleasas hasta los 7 mL. Las muestras se incubaron toda la noche a 16°C. Al día siguiente se añadieron a las muestras 7 µL de ATP 0.1M y 10 µL de ligasa del ADN T4 tras lo cual se volvieron a incubar 6 horas a 16°C. Después se añadieron 70 µL de proteinasa K 20 mg/mL y se incubaron toda la noche a 65°C.

Para recuperar el ADN se realizó una extracción con fenol, fenol/cloroformo y cloroformo como se ha explicado previamente, pero dejando precipitar el ADN a -20°C durante toda la noche después de añadir 700 µL de acetato de sodio 3M y pH4,5, 8mL de isopropanol y 40 µg de GlycoBlue (Thermo Fisher Scientific), un coprecipitante de ácidos nucleicos que visibiliza el precipitado gracias a su colorante azul. Al día siguiente se realizó una centrifugación de 1 hora a 4°C y 12.000 rpm en una minifuga para terminar de precipitar el ADN y lavarlo posteriormente con etanol frío al 70%. El ADN resultante se resuspendió en 200 µL de agua libre de nucleasas y se purificó con columnas Anmicon 30k (Millipore), resuspendiendo el producto final en 450 µL de agua libre de nucleasas.

Después se añadieron 50 µL de tampón Cutsmart 10X y 15 µL de enzima BfaI (150 U) (R0568, New England Biolabs) a cada muestra y se incubaron 6 horas a 37°C y en movimiento. Tras ese tiempo se volvieron a añadir 15 µL de BfaI y se dejó incubando toda la noche a 37°C en movimiento. La enzima se inactivó al día siguiente dejando las muestras a 80°C durante 20 minutos.

En tubos de 50 mL se realizó una segunda ligación para circularizar los fragmentos añadiendo 500 µL de tampón de ligasa de ADN 10X, 10 µL de ligasa de ADN T4 y agua libre de nucleasas hasta los 5 mL de volumen final dejando la reacción en incubación toda la noche a 16°C. El siguiente paso fue

añadir 5  $\mu$  de ATP 0,1 M y 10  $\mu$ L de ligasa de ADN T4, dejando las muestras de nuevo en incubación durante 6 horas a 16°C. El ADN de las muestras se extrajo con fenol, fenol/cloroformo y cloroformo, se precipitó añadiendo 500  $\mu$ L de acetato de sodio 3M a pH 4,5, 6 mL de isopropanol y 40  $\mu$ g de GlycoBlue, incubándose toda la noche a -20°C. Al día siguiente se centrifugaron las muestras durante 1 hora a 4°C y 12.000 rpm en una minifuga, se lavó con etanol frío al 70%, se resuspendió la muestra en 200  $\mu$ L de agua libre de nucleasas y se purificó el ADN con columnas Amicon 30K, resuspendiendo el ADN purificado en 450  $\mu$ L de tampón TE (Tris 10 mM a pH 7,5 y EDTA 0.1 mM).

Pasos	Temperatura	Tiempo
<b>Incubación</b>	94°C	2 minutos
<b>16 ciclos</b>	94°C	10 segundos
	60°C	1 minuto
	68°C	3 minutos
<b>Incubación</b>	68°C	5 minutos
<b>Incubación</b>	4°C	$\infty$

**Tabla 4. Condiciones 1ª tanda PCR 4C-seq.**

Cada conjunto de muestras se amplificó en dos rondas de PCR inversa para cada una de las 4 regiones usadas como anzuelo. Los oligonucleótidos se obtuvieron de IDT y se muestran en la tabla 1. La primera tanda de PCR se realizó en 8 reacciones independientes de 50  $\mu$ L que contenían 200 ng de ADN molde, 1  $\mu$ L de dNTPs 10 mM, 5  $\mu$ L del oligonucleótido “*reading*” 5 mM, 5  $\mu$ L de oligonucleótido “*nonreading*” 5 mM, 0,7  $\mu$ L ADN polimerasa (Expanding Long Template Polymerase) (11681842001, Sigma-Aldrich), 5  $\mu$ L de tampón 1 10x y agua libre de nucleasas hasta 50  $\mu$ L. Las condiciones de esta PCR se muestran en la Tabla 4.

Las muestras de la PCR se juntaron y se hicieron alícuotas de 50  $\mu$ L en tubos para purificar el ADN usando las columnas “Quiaprep 2.0 spin” (Qiagen) y eluyéndose posteriormente en 100  $\mu$ L de Tris 10 mM y pH 8,5.

La segunda tanda de PCR se realizó en 4 reacciones independientes de 50  $\mu$ L con 5  $\mu$ L de ADN purificado de la PCR realizada anteriormente, 1  $\mu$ L de dNTP 10 mM, 5  $\mu$ L de “*universal-index primer*”, conteniendo el oligonucleótido Universal A o B y uno de los 16 oligonucleótidos Index (Tabla 1), 0,7  $\mu$ L de ADN polimerasa (Expanding Long Template Polymerase), 5  $\mu$ L de tampón 1 10x y agua libre de nucleasas hasta los 50  $\mu$ L. Las condiciones de esta segunda tanda de PCR se muestran a continuación en la tabla 5:

Pasos	Temperatura	Tiempo
Incubación	94°C	2 minutos
20 ciclos	94°C	10 segundos
	60°C	1 minuto
	68°C	3 minutos
Incubación	68°C	5 minutos
Incubación	4°C	$\infty$

**Tabla 5. Condiciones 2ª tanda PCR 4C-seq.**

Las muestras de la segunda PCR se juntaron en una sola, se mezclaron y se purificaron usando columnas “Quiaprep 2.0 spin” y el ADN resultante se eluyó en 100  $\mu$ L de agua libre de nucleasas. La cantidad de ADN purificado se midió con el espectrofotómetro NanoDrop 2000 y por medio de un gel de agarosa al 1,5%. Las muestras se juntaron en una sola con una concentración similar de cada una y se secuenciaron usando el secuenciador NextSeq 500 (Illumina) con lecturas de 75 pb.

El análisis de las lecturas se realizó en el Servicio de Bioinformática del IPBLN. Se creó un archivo *fastq* para cada combinación de oligonucleótidos. Antes de realizar el estudio bioinformático se realizó un análisis de la calidad usando el software fastqc (147). Este análisis mostró una media de calidad de 33,1 y un

valor de un 88% de las muestras con un índice  $Q > 30$ . Se usó *cutadapt* (148) para eliminar las lecturas con una calidad inferior a 20 (parámetro mínimo de calidad). Los archivos que quedaron se analizaron con el programa “*pipe4C analysis pipeline*” (146). El programa *pipeline* necesita lecturas en formato FASTQ (comprimidas) y experimentos con parámetros específicos para cada experimento 4C-seq organizada en archivos para cada anzuelo. Para ello, las lecturas primero se escanearon desde el lado 5´ hasta la presencia del primer motivo RE. Por lo general, más del 90% tendrá este motivo y en nuestro caso fue del 100%. Posteriormente, todas las lecturas se recortaron desde su extremo 5´, dejando intacto el motivo RE. Después se recortaron las lecturas y se alinearon con el genoma de referencia de ratón C57BL6 mm10 de UCSC usando *Bowtie2* (149). Una vez alineados, como los extremos de los fragmentos de restricción son las unidades más pequeñas que se pueden mapear de los experimentos 4C, el *pipeline* cuenta las lecturas por extremos de fragmento. Lo que hace el *pipe4C* es crear *in-silico* una biblioteca de extremos de fragmentos para determinar que fragmentos se asignan de forma única al genoma de referencia con el tamaño proporcionado y después cuenta las lecturas del mapeo hasta los extremos de dicho fragmento. Las lecturas por extremos de fragmento se normalizaron para tener en cuenta la profundidad de la secuenciación de la biblioteca 4C-seq, superponiendo todas las lecturas asignadas al cromosoma que contiene cada anzuelo. La superposición asegura que la suma total de las lecturas realizadas sea de 1 millón. Después de la normalización, los datos 4C mapeados se suavizaron promediando una media de extremos de fragmentos.

## 12. Análisis estadísticos

Los análisis estadísticos se realizaron usando el programa Prism 5.0 (Graphpad). Se han realizado un mínimo de tres experimentos independientes para cada caso de estudio. El número de experimentos independientes realizados (n) se indica en las leyendas de las figuras. Se usó la prueba t-Student no apareado de dos colas no paramétrica con corrección de Welch como se indica usando el programa GraphPad Prism. Los valores de *p* que

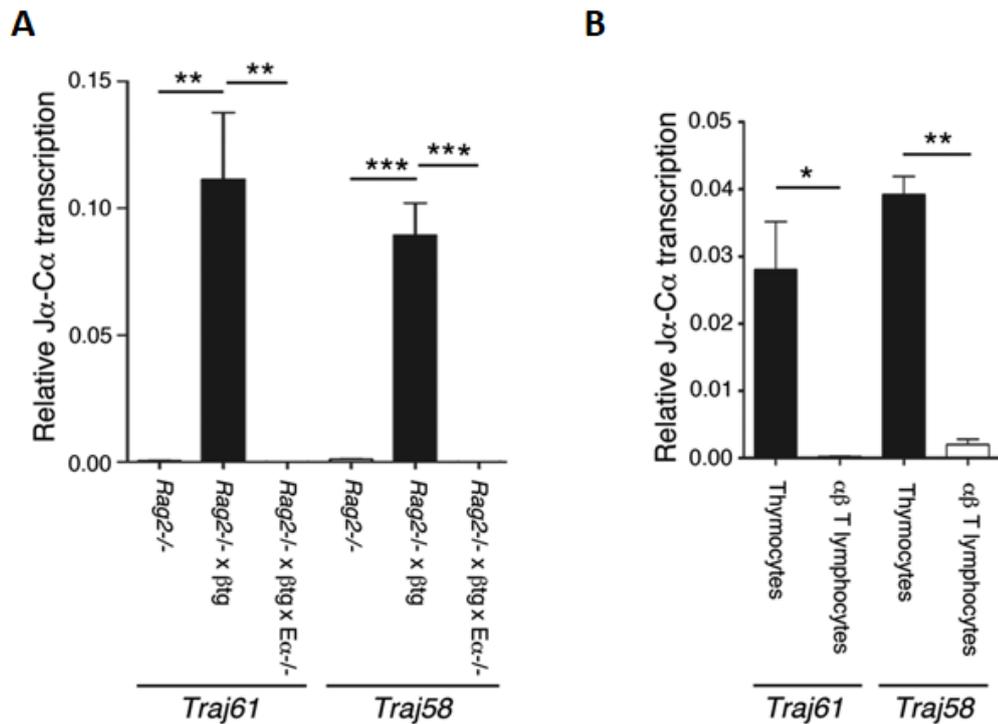
indican significación estadística se representan mediante asteriscos de la siguiente manera:  $p < 0,05$  (\*),  $p < 0,005$  (\*\*) y  $p < 0,0005$  (\*\*\*). La ausencia de asterisco indica que el cambio relativo al control no fue estadísticamente significativo.

## RESULTADOS

## RESULTADOS

### 1. La transcripción del gen reordenado de la cadena TCR $\alpha$ depende de E $\alpha$ en linfocitos T $\alpha\beta$

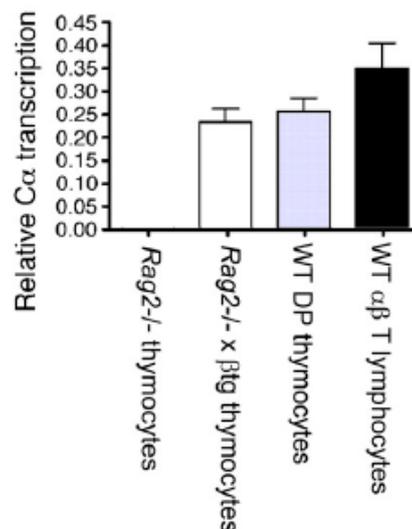
La señal mediada por el pre-TCR en timocitos DN $\beta$ a durante la selección- $\beta$  es responsable de la activación de la transcripción no codificante dependiente de E $\alpha$  de los segmentos génicos J $\alpha$  en timocitos DP, mientras que la señal mediada por el TCR $\alpha\beta$  en timocitos DP durante la selección positiva produce la reducción de estos transcritos en timocitos SP y linfocitos T  $\alpha\beta$  (60,94).



**Figura 15. Transcripción no codificante dependiente de E $\alpha$  del gen *Tcr $\alpha$*  durante el desarrollo de los timocitos.** (A) Análisis de los transcritos J $\alpha$  - C $\alpha$  indicados en timocitos de ratones Rag2<sup>-/-</sup>, Rag2<sup>-/-</sup> x  $\beta$ tg y Rag2<sup>-/-</sup> x  $\beta$ tg x E $\alpha$ <sup>-/-</sup>. (B) Análisis por RT-qPCR de los transcritos J $\alpha$ -C $\alpha$  indicados en timocitos y linfocitos T de ratones Rag2<sup>-/-</sup> x DO11.10 (n=4). Los resultados se normalizaron usando la transcripción del gen *Actb* como normalizador y representan la media  $\pm$ SEM de los duplicados de la RT-qPCR de n experimentos independientes. Se usó la prueba t-Student no paramétrica con datos no pareados y corrección del Welch; los valores de p se representan con asteriscos de la siguiente manera: p<0,05 (\*), p<0,005 (\*\*), y p<0,0005 (\*\*\*).

Como se esperaba, teniendo en cuenta los datos previos (22,47,59,87), se observaron altos niveles de transcritos de *Traj61 - Ca* y *Traj58 - Ca* en timocitos DP de ratones *Rag2*<sup>-/-</sup> x  $\beta$ tg, mientras que estos mismos transcritos estaban ausentes en timocitos DN3a de ratones *Rag2*<sup>-/-</sup> y en timocitos DP de ratones *Rag2*<sup>-/-</sup> x  $\beta$ tg x *Ea*<sup>-/-</sup> (Figura 15 A). Estos transcritos constituyen la mayoría de transcritos  $J\alpha$ -Ca y Ca derivados del gen *Tera* no reordenado en timocitos DP (47,59,87).

En consonancia con datos previos de inhibición de *Ea* en linfocitos T  $\alpha\beta$  (94), hemos observado una reducción del 85% en la transcripción no codificante de *Traj61 - Ca* y *Traj58 - Ca* en células T  $\alpha\beta$  comparada con la transcripción observada timocitos DP de ratones *Rag2*<sup>-/-</sup> x DO11.10 (Figura 15 B). Estos resultados confirman los datos previos (23,94), indicando que la actividad transcripcional de *Ea* estudiada en el contexto del gen *Tera* no reordenado está fuertemente regulada durante el desarrollo de las células T, estando activada en timocitos DP y fuertemente inhibida en linfocitos T  $\alpha\beta$ . De acuerdo con los datos previos (94), la magnitud de la transcripción total del gen *Tera* reordenado en timocitos DP y en linfocitos T  $\alpha\beta$  de ratones WT, medidos como transcritos de Ca, son similares a los del gen *Tera* no reordenado en timocitos DP de ratones *Rag2*<sup>-/-</sup> x  $\beta$ tg (Figura 16).

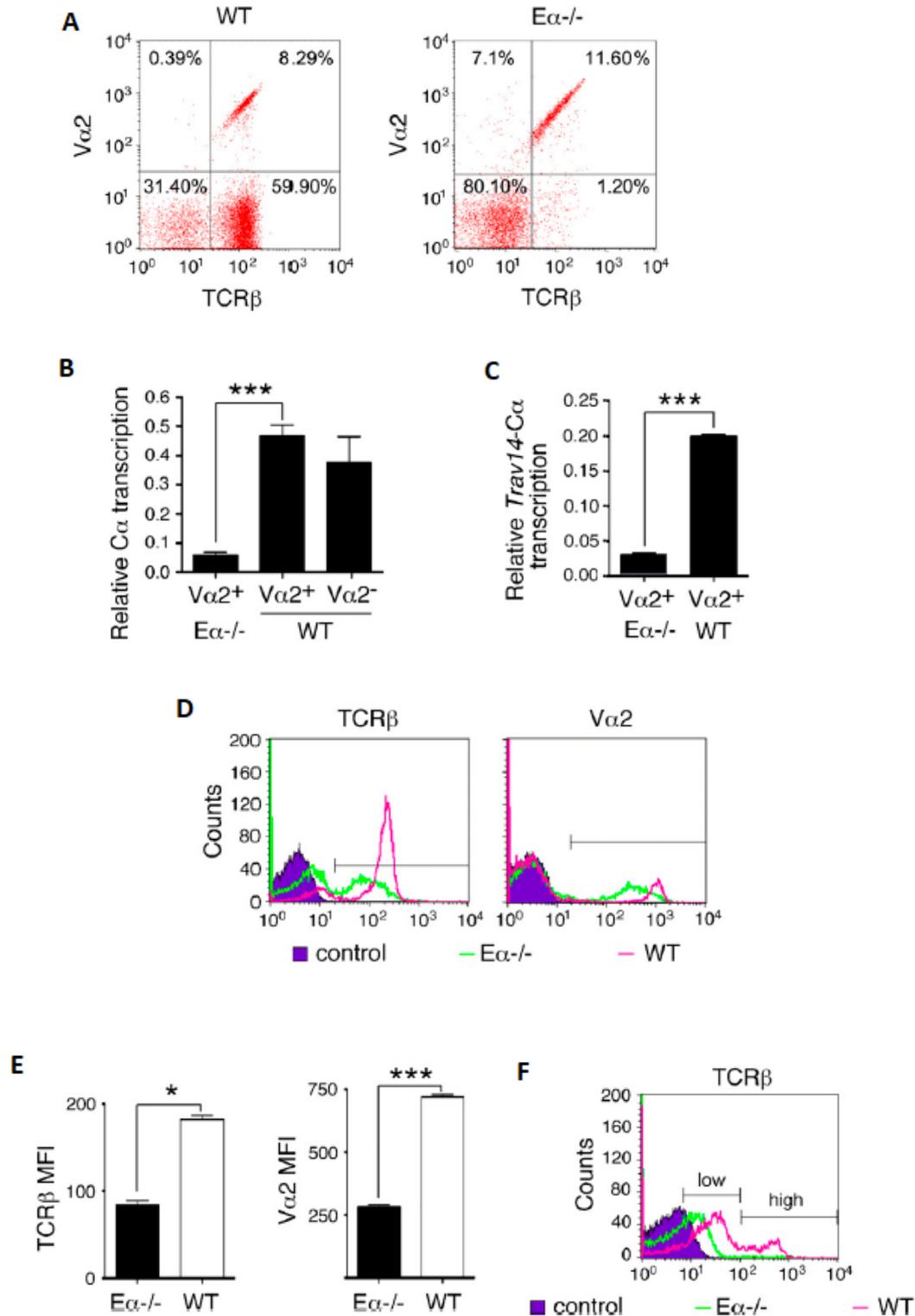


**Figura 16. Transcripción del gen *Tera* en timocitos y linfocitos T de ratones *wild-type* (WT).** Transcripción de Ca en timocitos de ratones WT, *Rag2*<sup>-/-</sup> y *Rag2*<sup>-/-</sup> x  $\beta$ tg y en linfocitos de ratones WT. (n=6). Los resultados se normalizaron usando la transcripción del gen *Actb* como normalizador y representan la media  $\pm$  SEM de los duplicados de la RT-qPCR de n experimentos independientes. Se usó la prueba t-Student no paramétrica con datos no pareados y corrección del Welch.

Estos resultados indican que la fuerte inhibición de E $\alpha$  en linfocitos T  $\alpha\beta$  no impide que el gen *Tcra* reordenado se transcriba eficientemente. Estos datos previos sugirieron la posibilidad de que E $\alpha$  podría no ser necesario para la normal transcripción del gen *Tcra* reordenado en células T  $\alpha\beta$  maduras (94).

E $\alpha$  es un regulador en *cis* que promueve la transcripción no codificante de J $\alpha$  y la recombinación V $\alpha$ J $\alpha$  en timocitos DP (22). Debido a su importante papel en la expresión de la cadena TCR $\alpha$ , en los ratones E $\alpha^{-/-}$  la generación de las células T  $\alpha\beta$  queda comprometida, observándose una reducción del número de estas células de unas 20 veces en bazo (22). Las pocas células T  $\alpha\beta$  presentes en los ratones E $\alpha^{-/-}$  provienen de unos reordenamientos que siempre comprenden segmentos génicos de la familia *Trav14*, células T V $\alpha$ 2<sup>+</sup> (22), y se deben a unos escasos reordenamientos V $\alpha$ J $\alpha$  dependientes de E $\delta$  que ocurren en timocitos DN3a, y que solo son observables en ratones E $\alpha^{-/-}$  (130). Aunque hay diversos reordenamientos V $\alpha$ J $\alpha$  dependientes de E $\delta$  en timocitos DN3a, solo sobreviven los que contienen reordenamientos productivos *Trav14* - J $\alpha$  en timocitos DP E $\alpha^{-/-}$ , ya que tienen la suficiente expresión de la cadena TCR $\alpha$  para pasar la selección positiva y permitir su maduración a linfocitos T  $\alpha\beta$ . La transcripción del gen *Tcra* reordenado en linfocitos T E $\alpha^{-/-}$  se debe a una activación independiente de *enhancer* de los promotores de los segmentos *Trav14*, ya que tanto E $\alpha$  como E $\delta$  están ausentes en ambos alelos del gen *Tcra* en los ratones E $\alpha^{-/-}$ .

Para analizar con claridad la importancia de E $\alpha$  en el contexto de los reordenamientos *Trav14* - J $\alpha$  en ratones E $\alpha^{-/-}$ , aislamos las células V $\alpha$ 2<sup>+</sup> presentes ratones WT y E $\alpha^{-/-}$  (Figura 17A) y comparamos la transcripción de C $\alpha$  y de los reordenamientos *Trav14* - J $\alpha$  en linfocitos V $\alpha$ 2<sup>+</sup> y V $\alpha$ 2<sup>-</sup> de ratones WT y E $\alpha^{-/-}$  (Figuras 17B y C).



**Figura 17. Transcripción del gen *Tcrα* en linfocitos T  $\alpha\beta$  de ratones WT y E $\alpha^{-/-}$  y expresión de TCR $\alpha\beta$  y Va2 en timocitos y linfocitos T. (A) Expresión en superficie del TCR $\alpha\beta$  y Va2 en esplenocitos de ratones WT y E $\alpha^{-/-}$ . (B) Transcripción de Ca en linfocitos T Va2<sup>+</sup> y Va2<sup>-</sup> de ratones WT y en linfocitos T Va2<sup>+</sup> de ratones E $\alpha^{-/-}$  (n=4-10). (C) Transcripción del reordenamiento *Trav14 - Ca* en linfocitos T de ratones WT y E $\alpha^{-/-}$  (n=4). (D) Solapamiento de los análisis de la expresión en superficie de TCR $\alpha\beta$  y Va2 en esplenocitos de ratones WT y E $\alpha^{-/-}$ . (E) Medida de la expresión en superficie de TCR $\alpha\beta$  y Va2 en esplenocitos de ratones WT y E $\alpha^{-/-}$  (n=3). (F) Solapamiento de los análisis de la expresión en superficie de TCR $\alpha\beta$  y Va2 en timocitos de ratones WT y E $\alpha^{-/-}$ . Los resultados se**

normalizaron mediante el análisis de los transcritos del gen *Actb* como normalizador y representan la media  $\pm$  SEM de los duplicados de la RT-qPCR de *n* experimentos independientes. Se usó la prueba t-Student no paramétrica con datos no pareados y corrección del Welch, los valores de *p* se representan con asteriscos de la siguiente manera:  $p < 0,05$  (\*) y  $p < 0,0005$  (\*\*\*)

Los niveles de transcritos  $V\alpha$  fueron similares en linfocitos T  $V\alpha 2^+$  y  $V\alpha 2^-$  de ratones WT (Figura 17B), indicando que la transcripción del gen *Tcra* reordenado mediada por los promotores de *Trav14* en presencia de Ea es comparable a la mediada por otros promotores  $V\alpha$ . A pesar de la expresión lo suficientemente alta de la cadena TCR $\alpha$  para pasar la selección positiva y permitir la maduración de los timocitos DP a linfocitos T  $\alpha\beta$  maduros en ausencia de Ea, los niveles de  $V\alpha$  presentes en linfocitos T de bazo de ratones  $E\alpha^{-/-}$  se encuentran altamente reducidos según estudios previos (22). De acuerdo con estos datos, nuestros datos confirmaron que los niveles de los transcritos de  $V\alpha$  en células T  $V\alpha 2^+$  de ratones  $E\alpha^{-/-}$  mostraban un descenso del 90% comparados con los de ratones WT (Figura 17B) indicando que la contribución de Ea para la transcripción del gen *Tcra* reordenado mediada por los promotores de *Trav14* es necesaria para la expresión normal de este gen, incluso en el caso de estar iniciada por acción de los promotores *Trav14*. Para analizar inequívocamente la contribución de Ea a la transcripción del gen *Tcra* reordenado, sin interferencias de transcritos de los alelos no reordenados presentes en las células T  $V\alpha 2^+$   $E\alpha^{-/-}$ , las cuales quedan incluidas al analizar los transcritos totales de  $V\alpha$ , comparamos los transcritos *Trav14* -  $V\alpha$  en linfocitos T  $V\alpha 2^+$  de ratones WT y  $E\alpha^{-/-}$  (Figura 17C). Acorde a los resultados que se obtuvieron del análisis de los transcritos  $V\alpha$ , encontramos que la abundancia de los transcritos *Trav14* -  $V\alpha$  de linfocitos T  $V\alpha 2^+$   $E\alpha^{-/-}$  estaba muy reducida al compararla con la presente en los linfocitos T  $V\alpha 2^+$  de ratones WT. Estos datos amplían los anteriormente publicados (22,130), indicando que Ea también es necesario para la transcripción normal de los alelos del gen *Tcra* derivados de los reordenamientos productivos *Trav14* -  $J\alpha$ .

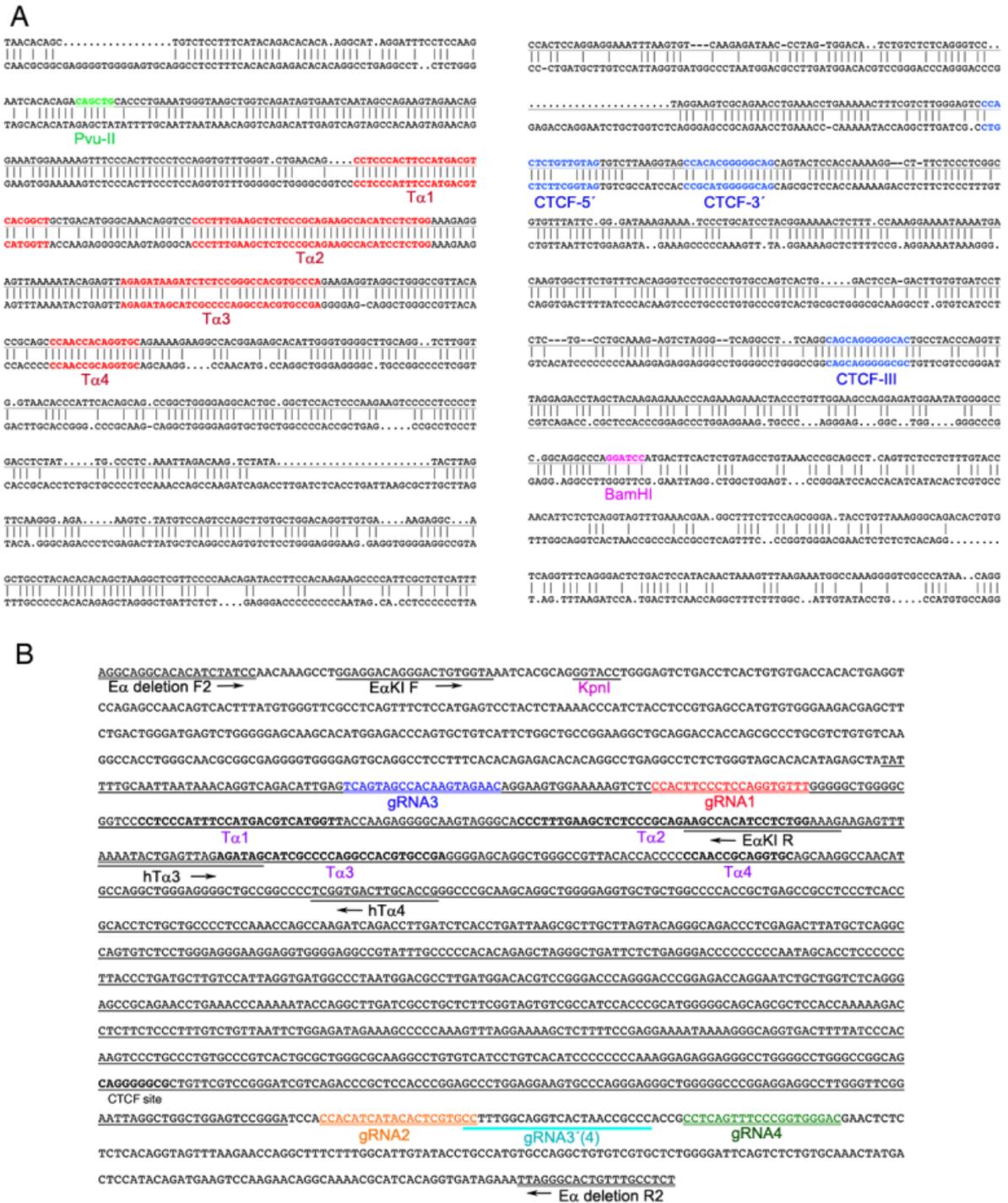
La presencia de las subunidades TCR $\alpha$ , TCR $\beta$  y CD3 es necesaria para la expresión eficiente y regulada del TCR $\alpha\beta$  en la membrana celular durante el desarrollo de las células T (150–152). De acuerdo con la presencia reducida de los transcritos *Trav14* reordenados y  $V\alpha$  en células T  $E\alpha^{-/-}$   $V\alpha 2^+$  comparada con las células T WT  $V\alpha 2^+$ , la expresión en superficie del TCR $\alpha\beta$  y de  $V\alpha 2$  en células T  $E\alpha^{-/-}$   $V\alpha 2^+$  se encuentra reducida comparada con la observada en

células T WT  $V\alpha 2^+$  (Figuras 17D y E). Asimismo, también se observó una reducción de la expresión de  $TCR\alpha\beta$  en los timocitos con altos niveles de expresión (los cuales representan a los timocitos SP que han pasado la selección positiva), así como de los que expresan niveles medios de  $TCR\alpha\beta$  (que representan a los timocitos DP) en ratones  $E\alpha^{-/-}$  comparados con ratones WT (Figura 17F). Estos resultados concuerdan con datos previos (22) e indican que la transcripción reducida de los segmentos génicos *Trav14* en ratones  $E\alpha^{-/-}$  se manifiesta con un número reducido de complejos  $TCR\alpha\beta$  en la membrana de los timocitos DP, SP y linfocitos T  $\alpha\beta$ . Estos datos confirman el papel fundamental de *E\alpha* en la transcripción y expresión de los alelos *Tcra* reordenados, incluyendo aquellos que contienen el segmento génico *Trav14* en timocitos y linfocitos T  $\alpha\beta$ .

Los genes de la cadena  $TCR\alpha$  de humano y ratón tienen una estructura muy similar, con la secuencia de *E\alpha* muy conservada (Figura 18A). Para ampliar el análisis de la función de *E\alpha* en la transcripción del gen reordenado de la cadena  $TCR\alpha$  en el contexto de otro segmento génico  $V\alpha$  y en células humanas, delecionamos el *enhancer* en la línea celular Jurkat usando la técnica de edición génica CRISPR-Cas9 (Figura 18B). El fragmento de *E\alpha* delecionado corresponde con el eliminado previamente en la generación del ratón  $E\alpha^{-/-}$  (22).

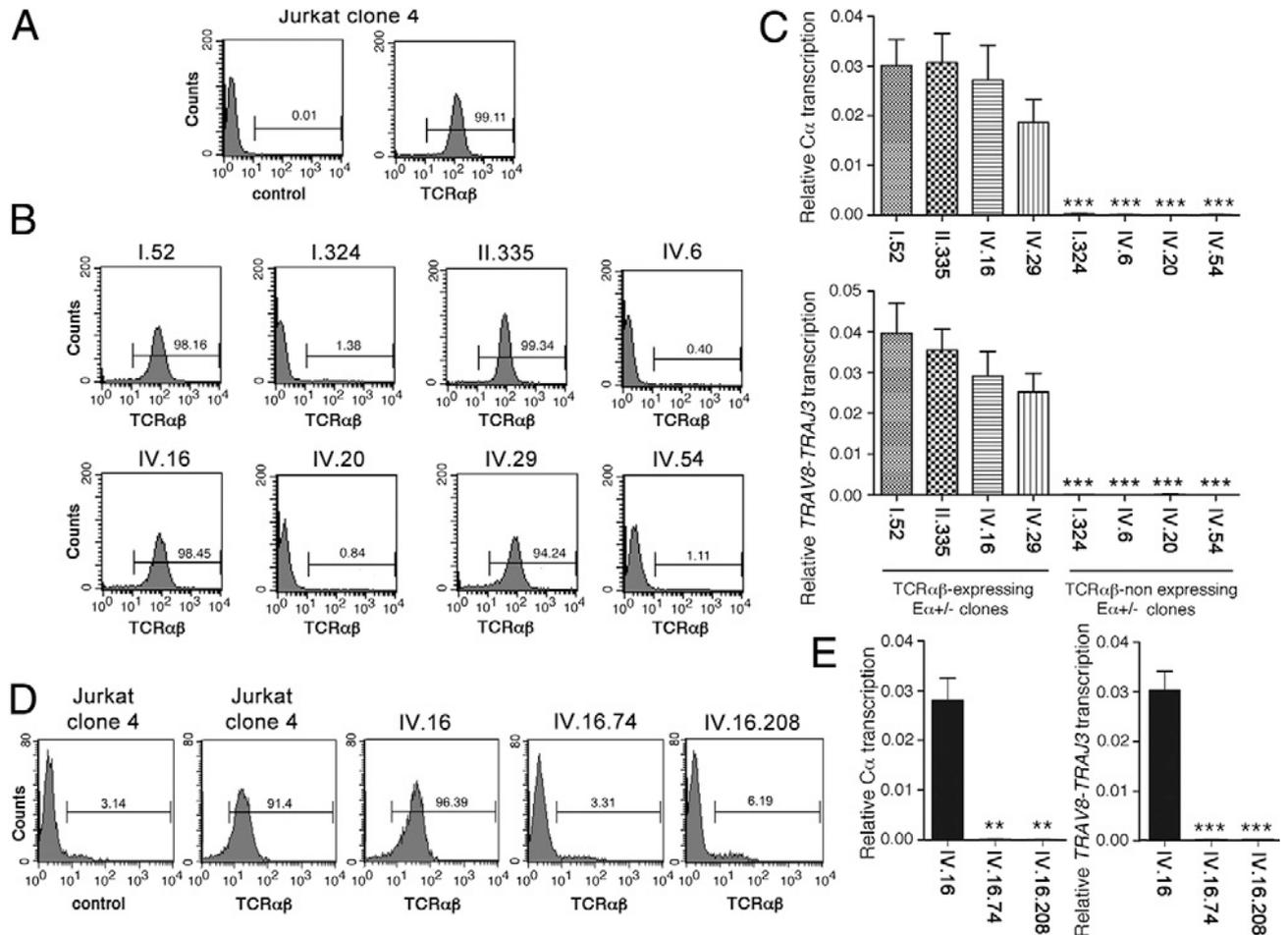
Elegimos las células Jurkat porque son un modelo celular ampliamente usado durante décadas para el estudio de la señalización mediada por el  $TCR\alpha\beta$ , además estas células pueden sobrevivir sin expresar el complejo  $TCR\alpha\beta$  en la superficie celular (153–158). Como las células Jurkat presentan una expresión variable del  $TCR\alpha\beta$  en superficie (159), primero aislamos un clon Jurkat, llamado clon 4, basado en la expresión media del  $TCR\alpha\beta$  medido por citometría de flujo para poder confirmar las consecuencias fenotípicas de la deleción de *E\alpha* en clones individuales (Figura 19A). Para deleccionar *E\alpha*, las células Jurkat del clon 4 se transfectaron con un plásmido derivado del px458 que lleva simultáneamente guías de ARN que flanquean a *E\alpha* en 5' y en 3' usando el electroporador Nucleofector (Lonza). Generamos 4 versiones diferentes del plásmido px458 con dos gARN flanqueantes cada una: plásmido

I con los gARNs 1 y 3; II con los gARNs 1 y 4; III con los gARNs 2 y 3 y IV con los gARNs 2 y 4 (Figura 18B y Tabla 1).



**Figura 18. Comparación de las secuencias de Eα humano y de ratón y estrategia de edición CRISPR-Cas9 para la eliminación de Eα en células Jurkat. (A)** Homología entre la secuencia superior (ratón) e inferior (humana) de Eα y las zonas circundantes. Los elementos Ta1-Ta4, definidos por la unión de factores de transcripción, se muestran en rojo. Los sitios CTCF se señalan en azul. La secuencia eliminada en ratones Eα<sup>-/-</sup> se muestra subrayada. Se muestran los sitios de restricción Pvu-II (verde) y BamHI (rosa) usados para eliminar Eα en ratones. **(B)** Secuencia de la región de Eα humano indicando las regiones Ta1, Ta2, Ta3, Ta4, los sitios de unión de distintos factores de transcripción y las

gARN usadas (gARN1 en rojo, gARN2 en naranja, gARN3 en azul y gARN4 en verde). Se muestran los lugares de anillamiento de los oligonucleótidos hTα3, hTα4, Eadel-F2 y R2 usados. La región subrayada indica la región humana que corresponde con la deletada en el ratón  $E\alpha^{-/-}$ .



**Figura 19. Expresión en superficie del TCRαβ y transcripción del gen *TCRA* reordenado en clones Jurkat  $E\alpha^{+/+}$  y  $E\alpha^{-/-}$ .** (A) Expresión en superficie del TCRαβ en el clon Jurkat 4 analizada por citometría de flujo: tinción de anticuerpo de control de isotipo (control) y tinción de anticuerpo TCRβ (TCRαβ). (B) Expresión del TCRαβ en varios clones  $E\alpha^{+/+}$  derivados del clon Jurkat 4 analizados por citometría de flujo: tinción con anticuerpo TCRβ (TCRαβ). (C) Transcripción de *Ca* y del reordenamiento *TRAV8-2 - TRAJ33* en clones  $E\alpha^{+/+}$  que expresan y que no expresan TCRαβ, analizados por RT-qPCR (n=6). Los resultados se normalizaron con los transcritos del gen *ACTB* y representan la media ± SEM de los duplicados de la RT-qPCR de n experimentos independientes. (D) Expresión del TCRαβ en los clones  $E\alpha^{-/-}$  IV.16.75 y IV.16.208 derivados del clon  $E\alpha^{+/+}$  IV.16 analizados por citometría de flujo: tinción con el anticuerpo TCRβ (TCRαβ). (E) Transcripción de *Ca* y del reordenamiento *TRAV8-2 - TRAJ33* en el clon  $E\alpha^{+/+}$  IV.16 y en los clones  $E\alpha^{-/-}$  IV.16.74 y IV.16.208 analizados por RT-qPCR (n=4-8). Los resultados se normalizaron con los transcritos del gen *ACTB* y representan la media ± SEM de los duplicados de la RT-qPCR de n experimentos independientes. Se usó la prueba t-student no paramétrica con datos no pareados y corrección del Welch, los valores de *p* se representan con asteriscos de la siguiente manera: *p* < 0,005 (\*\*) y *p* < 0,0005 (\*\*\*).

Aislamos un total de 18 clones  $E\alpha^{+/+}$ , de entre los 300-600 clones derivados de las distintas combinaciones de gARN (Figura 18), con diferentes deleciones monoalélicas del *enhancer* (Figura 19). La identificación de los clones en los que ocurrió la deleción del *enhancer* se hizo mediante PCR con

los oligonucleótidos Eadel-F2/R2 (Tabla 1). Las secuencias de las deleciones en los distintos clones Ea<sup>+/-</sup> obtenidos se presentan a continuación.

**Ea<sup>+/-</sup> I-CLONES: gARN1/gARN2**

**I-47**

GGCCACCTGGGCAACGCGGGGAGGGGTGGGGAGTGCAGGCCTCCTTTACACAGAGACACACAGGCC  
TGAGGCCTCTCTGGGTAGCACACATAGAGCTATATTTTGCAATTAATAAACAGGTCAGACATTGAGTC  
AGTAGCCACAAGTAGAACAGGAAAGTGAAAAAGTCTCCACTTGCCCTTTGGCAGGTCCTAAACCG  
CCCACCGCCTCAGTTTCCCGGTGGGACGAACTCTCTCTCACAGGTAGTTTAAGAACCAGGCTTTCTT  
TGGCATTGTATACCTGCCATGTGCCAGGCTGTGTCGTGCTCTGGGGATTGAGTCTCTGTGCAAATAT  
GACTCCATACAGATGAAGTCCAAGAA

**I-52**

GGCCACCTGGGCAACGCGGGGAGGGGTGGGGAGTGCAGGCCTCCTTTACACAGAGACACACAGGCC  
TGAGGCCTCTCTGGGTAGCACACATAGAGCTATATTTTGCAATTAATAAACAGGTCAGACATTGAGTC  
AGTAGCCACAAGTAGAACAGGAAAGTGAAAAAGTCTCCATCATACTCGTGCCCTTTGGCAGG  
TCACTAACCGCCACCGCCTCAGTTTCCCGGTGGGACGAACTCTCTCTCACAGGTAGTTTAAGAAC  
CAGGCTTTCTTTGGCATTGTATACCTGCCATGTGCCAGGCTGTGTCGTGCTCTGGGGATTGAGTCTCT  
TGCAAATATGACTCCATACAGATGAAGTCCAAGAA

**I-253**

GGCCACCTGGGCAACGCGGGGAGGGGTGGGGAGTGCAGGCCTCCTTTACACAGAGACACACAGGCC  
TGAGGCCTCTCTGGGTAGCACACATAGAGCTATATTTTGCAATTAATAAACAGGTCAGACATTGAGTC  
AGTAGCCACAAGTAGAACAGGAAAGTGAAAAAGTCTCCACTTCCCTCCAGGTGAGCCTTTG  
GCAGGTCCTAACCGCCACCGCCTCAGTTTCCCGGTGGGACGAACTCTCTCTCACAGGTAGTTTA  
AGAACCAGGCTTTCTTTGGCATTGTATACCTGCCATGTGCCAGGCTGTGTCGTGCTCTGGGGATTGAG  
TCTCTGTGCAAATATGACTCCATACAGATGAAGTCCAAGAA

**I-279**

GGCCACCTGGGCAACGCGGGGAGGGGTGGGGAGTGCAGGCCTCCTTTACACAGAGACACACAGGCC  
TGAGGCCTCTCTGGGTAGCACACATAGAGCTATATTTTGCAATTAATAAACAGGTCAGACATTGAGTC  
AGTAGCCACAAGTAGAACAGGAAAGTGAAAAAGTCTCCACTTCCCTCCAGGTGCATCATAACA  
CTCGTGCCCTTTGGCAGGTCCTAACCGCCACCGCCTCAGTTTCCCGGTGGGACGAACTCTCTCTCA  
CAGGTAGTTTAAGAACCAGGCTTTCTTTGGCATTGTATACCTGCCATGTGCCAGGCTGTGTCGTGCT  
TGGGGATTGAGTCTCTGTGCAAATATGACTCCATACAGATGAAGTCCAAGAA

**I-283**

GGCCACCTGGGCAACGCGGGGAGGGGTGGGGAGTGCAGGCCTCCTTTACACAGAGACACACAGGCC  
TGAGGCCTCTCTGGGTAGCACACATAGAGCTATATTTTGCAATTAATAAACAGGTCAGACATTGAGTC  
AGTAGCCACAAGTAGAACAGGAAAGTGAAAAAGTCTCCACTTCCCTCCAGGTGCATCATAACA  
TCGTGCCCTTTGGCAGGTCCTAACCGCCACCGCCTCAGTTTCCCGGTGGGACGAACTCTCTCTCA  
CAGGTAGTTTAAGAACCAGGCTTTCTTTGGCATTGTATACCTGCCATGTGCCAGGCTGTGTCGTGCT  
TGGGGATTGAGTCTCTGTGCAAATATGACTCCATACAGATGAAGTCCAAGAA

**I-324**

GGCCACCTGGGCAACGCGGGGAGGGGTGGGGAGTGCAGGCCTCCTTTACACAGAGACACACAGGCC  
TGAGGCCTCTCTGGGTAGCACACATAGAGCTATATTTTGCAATTAATAAACAGGTCAGACATTGAGTC  
AGTAGCCACAAGTAGAACAGGAAAGTGAAAAAGTCTCCACTTCCCGTCGTGCCCTTTGGCAGG  
TCACTAACCGCCACCGCCTCAGTTTCCCGGTGGGACGAACTCTCTCTCACAGGTAGTTTAAGAAC  
CAGGCTTTCTTTGGCATTGTATACCTGCCATGTGCCAGGCTGTGTCGTGCTCTGGGGATTGAGTCTCT  
TGCAAATATGACTCCATACAGATGAAGTCCAAGAA

**Eα<sup>+/-</sup> II-CLONES: gARN1/gARN4**

**II.29**

GGCCACCTGGGCAACGCGGCGAGGGGTGGGGAGTGCAGGCCTCCTTTACACAGAGACACACAGGCC  
TGAGGCCTCTCTGGGTAGCACACATAGAGCTATATTTTGCAATTAATAAACAGGTCAGACATTGAGTC  
AGTAGCCACAAGTAGAACAGGAAAGTGAAAAAGTCTCCACTTCCCTCCAGGTGTTTCCCGGT  
GGGACGAACTCTCTCTCACAGGTAGTTTAAGAACCAGGCTTTCTTTGGCATTGTATACCTGCCATGTG  
CCAGGCTGTGTCGTGCTCTGGGGATTCACTCTCTGTGCAAATATGACTCCATACAGATGAAGTCCAA  
GAA

**II.30**

GGCCACCTGGGCAACGCGGCGAGGGGTGGGGAGTGCAGGCCTCCTTTACACAGAGACACACAGGCC  
TGAGGCCTCTCTGGGTAGCACACATAGAGCTATATTTTGCAATTAATAAACAGGTCAGACATTGAGTC  
AGTAGCCACAAGTAGAACAGGAAAGTGAAAAAGTCTCCACTTCCCTCCAGTTTCCCGGTGGG  
ACGAACTCTCTCTCACAGGTAGTTTAAGAACCAGGCTTTCTTTGGCATTGTATACCTGCCATGTGCCAG  
GCTGTGTCGTGCTCTGGGGATTCACTCTCTGTGCAAATATGACTCCATACAGATGAAGTCCAAGAA

**II.274**

GGCCACCTGGGCAACGCGGCGAGGGGTGGGGAGTGCAGGCCTCCTTTACACAGAGACACACAGGCC  
TGAGGCCTCTCTGGGTAGCACACATAGAGCTATATTTTGCAATTAATAAACAGGTCAGACATTGAGTC  
AGTAGCCACAAGTAGAACAGGAAAGTGAAAAAGTCTCCACTTCCCTCCAGGTGCAGTTTCCCG  
GGTGGGACGAACTCTCTCTCACAGGTAGTTTAAGAACCAGGCTTTCTTTGGCATTGTATACCTGCCAT  
GTGCCAGGCTGTGTCGTGCTCTGGGGATTCACTCTCTGTGCAAATATGACTCCATACAGATGAAGTC  
CAAGAA

**II.335**

GGCCACCTGGGCAACGCGGCGAGGGGTGGGGAGTGCAGGCCTCCTTTACACAGAGACACACAGGCC  
TGAGGCCTCTCTGGGTAGCACACATAGAGCTATATTTTGCAATTAATAAACAGGTCAGACATTGAGTC  
AGTAGCCACAAGTAGAACAGGAAAGTGAAAAAGTCTCCACTTCCCTCCAGTTTCCCGGTGGG  
ACGAACTCTCTCTCACAGGTAGTTTAAGAACCAGGCTTTCTTTGGCATTGTATACCTGCCATGTGCCAG  
GCTGTGTCGTGCTCTGGGGATTCACTCTCTGTGCAAATATGACTCCATACAGATGAAGTCCAAGAA

**Eα<sup>+/-</sup> IV-CLONES: gARN3/gARN4**

**IV.6**

GGCCACCTGGGCAACGCGGCGAGGGGTGGGGAGTGCAGGCCTCCTTTACACAGAGACACACAGGCC  
TGAGGCCTCTCTGGGTAGCACACATAGAGCTATATTTTGCAATTAATAAACAGGTCAGACATTGAGTC  
AGTAGCCACAAGTAGTTTAGTTTCCCGGTGGGACGAACTCTCTCTCACAGGTAGTTTAAGAACCA  
GGCTTTCTTTGGCATTGTATACCTGCCATGTGCCAGGCTGTGTCGTGCTCTGGGGATTCACTCTCTGTG  
CAAATATGACTCCATACAGATGAAGTCCAAGAA

**IV.8**

GGCCACCTGGGCAACGCGGCGAGGGGTGGGGAGTGCAGGCCTCCTTTACACAGAGACACACAGGCC  
TGAGGCCTCTCTGGGTAGCACACATAGAGCTATATTTTGCAATTAATAAACAGGTCAGACATTGAGTC  
AGTAGCCACAAGTAGTTTCCCGGTGGGACGAACTCTCTCTCACAGGTAGTTTAAGAACCAGGCTTT  
CTTTGGCATTGTATACCTGCCATGTGCCAGGCTGTGTCGTGCTCTGGGGATTCACTCTCTGTGCAAAT  
ATGACTCCATACAGATGAAGTCCAAGAA

**IV.14**

GGCCACCTGGGCAACGCGGCGAGGGGTGGGGAGTGCAGGCCTCCTTTACACAGAGACACACAGGCC  
TGAGGCCTCTCTGGGTAGCACACATAGAGCTATATTTTGCAATTAATAAACAGGTCAGACATTGAGTC  
AGTAGCCACAAGTAGGGTTTCCCGGTGGGACGAACTCTCTCTCACAGGTAGTTTAAGAACCAGG  
CTTTCTTTGGCATTGTATACCTGCCATGTGCCAGGCTGTGTCGTGCTCTGGGGATTCACTCTCTGTGCA  
AACTATGACTCCATACAGATGAAGTCCAAGAA

#### IV.16

GGCCACCTGGGCAACGCGGCGAGGGGTGGGGAGTGCAGGCCTCCTTTTCACACAGAGACACACAGGCC  
TGAGGCCTCTCTGGGTAGCACACATAGAGCTATATTTTGCAATTAATAAACAGGTCAGACATTGAGTC  
AGTAGCCACAAGTAGCAGTTTCCCGGTGGGACGAACTCTCTCTCACAGGTAGTTTAAGAACCAG  
GCTTTCTTTGGCATTGTATACCTGCCATGTGCCAGGCTGTGTCGTGCTCTGGGGATTTCAGTCTCTGTGC  
AACTATGACTCCATACAGATGAAGTCCAAGAA

#### IV.20

GGCCACCTGGGCAACGCGGCGAGGGGTGGGGAGTGCAGGCCTCCTTTTCACACAGAGACACACAGGCC  
TGAGGCCTCTCTGGGTAGCACACATAGAGCTATATTTTGCAATTAATAAACAGGTCAGACATTGAGTC  
AGTAGCCACAAGTAGTTTCCCGGTGGGACGAACTCTCTCTCACAGGTAGTTTAAGAACCAGGCTT  
TCTTTGGCATTGTATACCTGCCATGTGCCAGGCTGTGTCGTGCTCTGGGGATTTCAGTCTCTGTGCAAAC  
TATGACTCCATACAGATGAAGTCCAAGAA

#### IV.28

GGCCACCTGGGCAACGCGGCGAGGGGTGGGGAGTGCAGGCCTCCTTTTCACACAGAGACACACAGGCC  
TGAGGCCTCTCTGGGTAGCACACATAGAGCTATATTTTGCAATTAATAAACAGGTCAGACATTGAGTC  
AGTAGCCACAAGTAGTTTCCCGGTGGGACGAACTCTCTCTCACAGGTAGTTTAAGAACCAGGCTT  
TCTTTGGCATTGTATACCTGCCATGTGCCAGGCTGTGTCGTGCTCTGGGGATTTCAGTCTCTGTGCAAAC  
TATGACTCCATACAGATGAAGTCCAAGAA

#### IV.29

GGCCACCTGGGCAACGCGGCGAGGGGTGGGGAGTGCAGGCCTCCTTTTCACACAGAGACACACAGGCC  
TGAGGCCTCTCTGGGTAGCACACATAGAGCTATATTTTGCAATTAATAAACAGGTCAGACATTGAGTC  
AGTAGCCACAAGTAGAACTCAGTTTCCCGGTGGGACGAACTCTCTCTCACAGGTAGTTTAAG  
AACCAGGCTTTCTTTGGCATTGTATACCTGCCATGTGCCAGGCTGTGTCGTGCTCTGGGGATTTCAGTCT  
CTGTGCAAACATGACTCCATACAGATGAAGTCCAAGAA

#### IV.54

GGCCACCTGGGCAACGCGGCGAGGGGTGGGGAGTGCAGGCCTCCTTTTCACACAGAGACACACAGGCC  
TGAGGCCTCTCTGGGTAGCACACATAGAGCTATATTTTGCAATTAATAAACAGGTCAGACATTGAGTC  
AGTAGCCACAAGTAGCAGTTTCCCGGTGGGACGAACTCTCTCTCACAGGTAGTTTAAGAACCAG  
GCTTTCTTTGGCATTGTATACCTGCCATGTGCCAGGCTGTGTCGTGCTCTGGGGATTTCAGTCTCTGTGC  
AACTATGACTCCATACAGATGAAGTCCAAGAA

**Figura 20. Secuencias de los sitios de delección en los clones Jurkat E $\alpha^{+/-}$  y E $\alpha^{-/-}$ .** Las secuencias de las gARN identificadas por secuenciación están indicadas utilizando el código de color: gARN1 en rojo, gARN2 en naranja, gARN3 en azul y gARN4 en verde.

Los análisis de la expresión del TCR $\alpha\beta$  mediante citometría de flujo mostraron que 4 de los 18 clones E $\alpha^{+/-}$ , concretamente I.324, IV.6, IV.20 y IV.54, no expresaban TCR $\alpha\beta$  en la superficie celular (Figura 19B). De acuerdo con la regulación en *cis* de E $\alpha$  (22), estos resultados indican que la delección monoalélica de E $\alpha$  en los clones que no expresan TCR $\alpha\beta$  en superficie ha ocurrido en el alelo productivo, mientras que la delección de E $\alpha$  en los clones TCR $\alpha\beta^{+}$  ha ocurrido en el alelo no productivo. Las dos gARN 5' y 3' usadas para la generación de los clones I y IV fueron distintas, lo que descartó posibles efectos “*off target*” derivados del uso de las gARN usadas en la edición por

CRISPR-Cas9 (Tabla 1 y Figura 18B). Después comparamos la abundancia de los transcritos  $\text{Ca}$  totales y los transcritos *TRAV8-2* - *TRAJ13* productivos de los clones  $\text{E}\alpha^{+/-} \text{TCR}\alpha\beta^{+}$  y  $\text{E}\alpha^{+/-} \text{TCR}\alpha\beta^{-}$  (Figura 19C). Aleatoriamente, escogimos cuatro clones  $\text{E}\alpha^{+/-}$  de entre los catorce  $\text{TCR}\alpha\beta^{+}$ , I.52, II.335, IV.16 y IV.29, para compararlos con los cuatro clones  $\text{E}\alpha^{+/-} \text{TCR}\alpha\beta^{-}$  I.324, IV.6, IV.20 y IV.54. Los datos mostraron que tanto los transcritos  $\text{Ca}$  como los *TRAV8-2* - *TRAJ33* no se encontraban en los clones  $\text{E}\alpha^{+/-} \text{TCR}\alpha\beta^{-}$ , mientras que su expresión era muy alta (normal) en los clones  $\text{E}\alpha^{+/-} \text{TCR}\alpha\beta^{+}$ .

Para confirmar que la transcripción de  $\text{TCR}\alpha$  en los clones  $\text{E}\alpha^{+/-} \text{TCR}\alpha\beta^{+}$  se debe a la presencia de  $\text{E}\alpha$  en el alelo no productivo generamos clones  $\text{E}\alpha^{-/-}$  en un segundo proceso de edición por CRISPR-Cas9. Para hacerlo, cogimos aleatoriamente el clon  $\text{E}\alpha^{+/-} \text{TCR}\alpha\beta^{+}$  IV.16 que se consiguió usando las gARNs más externas, gARN2 y gARN4. Para la delección específica del alelo productivo usamos las gARN 1 y gARN3, más internas, puesto que las secuencias complementarias en el otro cromosoma ya no están presentes al haber sido delecionadas anteriormente (Figuras 18B). Aislamos dos clones  $\text{E}\alpha^{-/-}$  de los más de 700 que fueron analizados, IV.16.74 y IV.16.208, basándonos en los análisis mediante qPCR con los oligonucleótidos hT $\alpha$ 3/hT $\alpha$ 4 y posterior análisis por citometría de flujo para medir la expresión del  $\text{TCR}\alpha\beta$  (Figura 19D). Para identificar los clones diferentes obtuvimos amplicones de cada uno de ellos usando los oligonucleótidos E $\alpha$ del-F2/R2 (Tabla 1) y su posterior secuenciación (Figura 21).

### $\text{E}\alpha^{-/-}$ IV.16 CLONES: gARN1/gARN2

#### IV.16.74

GGCCACCTGGGCAACGCGGCGAGGGGTGGGGAGTGCAGGCCTCCTTTTCACACAGAGACACACAGGCC  
TGAGGCCTCTCTGGGTAGCACACATAGAGCTATATTTTGCAATTAATAAACAGGTCAGACATTGAGTC  
AGTAGCCACAAGTAGATGAAACAGGAAGTGAAAAAGTCTCCACTTCCCTCCAGGTGCATCAT  
ACACTCGTGCCTTTGGCAGGTCACCTAACCGCCCGCCGCTCGAACTCTCTCTCACAGGTAGTTTAAG  
AACCAGGCTTTCTTTGGCATTGTATACCTGCCATGTGCCAGGCTATGTCGTGCTCTGGGGATTTCAGTCT  
CTGTGCAAACCTATGACTCCATACAGATGAAGTCCAAGAA

#### IV.16.208

GGCCACCTGGGCAACGCGGCGAGGGGTGGGGAGTGCAGGCCTCCTTTTCACACAGAGACACACAGGCC  
TGAGGCCTCTCTGGGTAGCACACATAGAGCTATATTTTGCAATTAATAAACAGGTCAGACATTGAGTC  
AGTAGCCACAAGTAGATGAAACAGGAAGTGAAAAAGTCTCCTCTTTGGCAGGTCACCTAACCGCC  
ACCGCCTCGAACTCTCTCTCACAGGTAGTTTAAGAACCAGGCTTTCTTTGGCATTGTATACCTGCCAT  
GTGCCAGGCTGTGTCGTGCTCTGGGGATTTCAGTCTCTGTGCAAACCTATGACTCCATACAGATGAAGTC  
CAAGAA

**Figura 21. Secuencias de los sitios de delección en los clones Jurkat  $E\alpha^{-/-}$ .** Las secuencias de las gARN identificadas por secuenciación están indicadas utilizando el código de color: gARN1 en rojo, gARN3 en azul y gARN4 en verde.

Los análisis de transcripción de  $C\alpha$  y de TRAV8-2 - TRAJ33 mediante RT-qPCR en los dos clones  $E\alpha^{-/-}$  confirmados por citometría de flujo demostraron la ausencia de estos transcritos en estas células comparado con la expresión de estos transcritos en el clon parental IV.16 (Figura 19E). Estos resultados confirman los anteriormente obtenidos en los clones  $E\alpha^{+/-}$ , indicando que la transcripción del gen *TCRA* reordenado y del TCR $\alpha\beta$  depende de la función de  $E\alpha$  en células Jurkat.

## 2. Secuencias distales colaboran con $E\alpha$ para el mantenimiento de la transcripción del gen *Tcra* reordenado en linfocitos T $\alpha\beta$

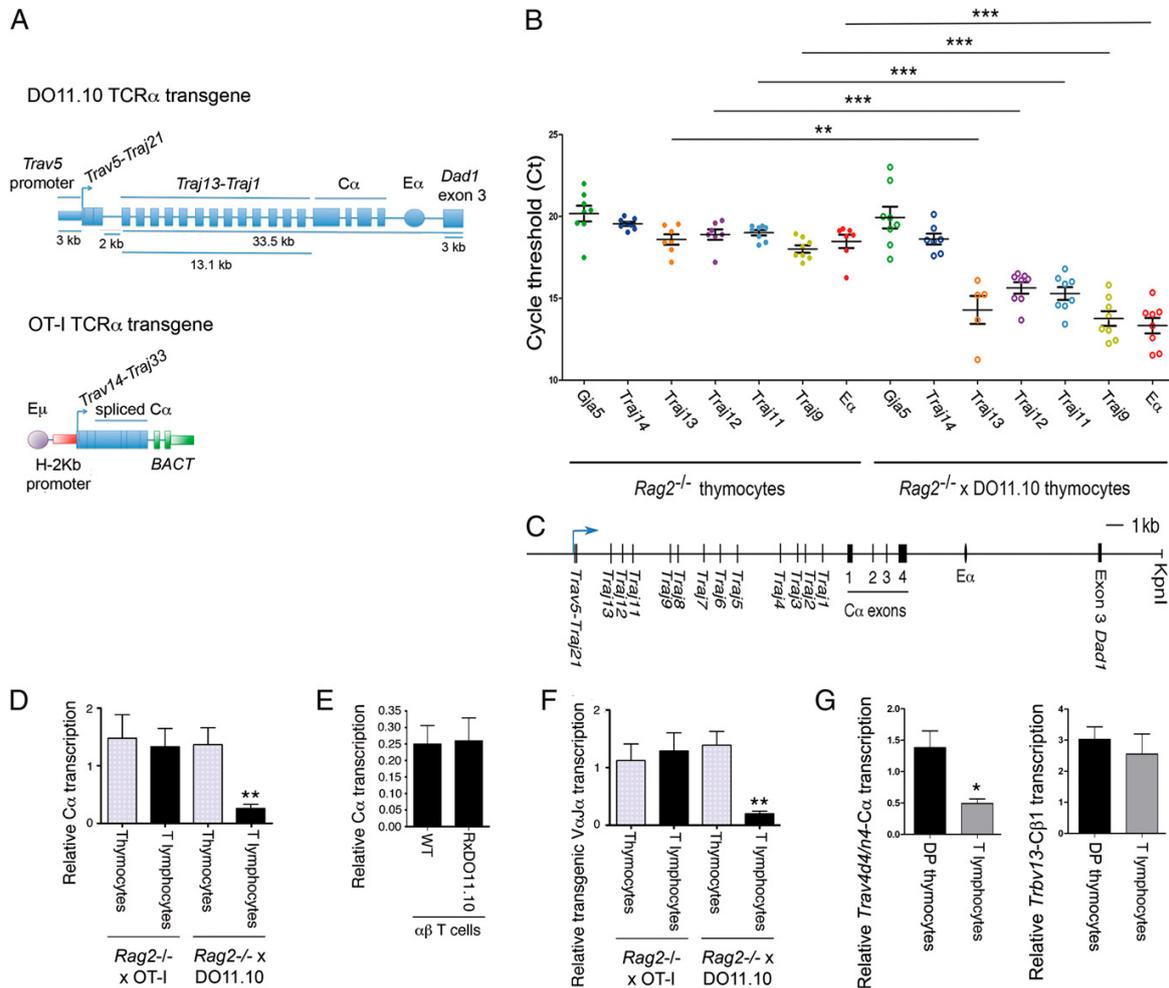
Los niveles de transcripción  $C\alpha$  derivados del gen *Tcra* reordenado en timocitos DP de ratones WT y en linfocitos T  $\alpha\beta$  son comparables a los del gen *Tcra* no reordenado que se encuentra en timocitos DP de ratones *Rag2<sup>-/-</sup> x  $\beta$ tg*, indicando que la fuerte inhibición que presenta  $E\alpha$  en linfocitos T  $\alpha\beta$  no impide que el gen *Tcra* reordenado se siga transcribiendo de forma eficiente (Figura 16B). La alta transcripción del gen *Tcra* reordenado que se da en linfocitos T  $\alpha\beta$  podría ser consecuencia de la cercanía de  $E\alpha$  a los segmentos  $V\alpha$  tras la recombinación  $V\alpha J\alpha$  o deberse a la colaboración del *enhancer* con una secuencia reguladora en *cis* que sea necesaria en el contexto del gen *Tcra* reordenado.

Experimentos previos con ratones transgénicos indicaban que  $E\alpha$  no es suficiente, por sí sólo, para mantener altos niveles de transcripción del gen *Tcra* reordenado en linfocitos T  $\alpha\beta$  (160), lo que indica que la configuración del gen *Tcra* reordenado en la que los promotores  $V\alpha$  y el *enhancer* se encuentran más próximos no es suficiente para prevenir un descenso de la transcripción. Estos resultados sugieren que debe existir una región reguladora en *cis*, que está ausente en los transgenes *Tcra* reordenados controlados por  $E\alpha$ , y que es necesaria para que se dé la transcripción eficiente del gen *Tcra* reordenado endógeno en linfocitos T  $\alpha\beta$ .

Con el fin de identificar esta secuencia colaboradora, realizamos un análisis transcripcional comparativo en timocitos DP y linfocitos T  $\alpha\beta$  en un transgén *Tcra* reordenado, el transgén *Tcra* DO11.10, en ratones *Rag2*<sup>-/-</sup> x DO11.10 (144). El transgén *Tcra* DO11.10 es el más largo que se ha diseñado nunca para recrear el contexto genómico del gen *Tcra* endógeno con el propósito de simular la regulación natural de este gen durante el desarrollo de los linfocitos T (144). Este transgén consiste en un reordenamiento natural *Trav5-d4/n4 - Traj21*, cuya transcripción está dirigida por su promotor *V $\alpha$*  natural; además contiene una región de 2 kb aguas abajo del segmento génico *Traj21* y una larga región genómica de 33,5 kb en 3' que incluye desde parte del *cluster* de segmentos génicos *J $\alpha$*  hasta el tercer exón de *Dad1*, incluyendo a *Ca* y el LCR completo (Figura 22A).

Para definir la región del *cluster* de segmentos *J $\alpha$*  presentes en el transgén, mapeamos la presencia de determinados segmentos génicos *J $\alpha$*  en el transgén *Tcra* DO11.10 mediante qPCR (Figura 22A y C). Para realizar el mapeo comparamos el número umbral de ciclos (Ct) obtenidos al ampliar la misma cantidad de ADN genómico de timocitos *Rag2*<sup>-/-</sup> (con dos copias endógenas del gen *Tcra*) y *Rag2*<sup>-/-</sup> x DO11.10 (con dos copias endógenas del gen *Tcra* y cuatro copias del transgén *Tcra* DO11.10) (Figura 22B). Como controles internos, comparamos la presencia del gen *Gja5* y de *E $\alpha$*  en ambas muestras (161). Nuestros datos indican que el segmento *Traj14* no se encuentra presente dentro del transgén *Tcra* DO11.10 mientras que los segmentos *Traj13*, *Traj12*, *Traj11* y *Traj9* están presentes, lo que indica que el fin de la secuencia incluida en el transgén se encuentra entre los segmentos génicos *Traj14* y *Traj13*. Como control de estos experimentos analizamos el transgén *Tcra* OT-I en ratones *Rag2*<sup>-/-</sup> x OT-I (162). La transcripción de este transgén está mediada por secuencias ectópicas como el promotor del gen *H2-K<sup>b</sup>*, el *enhancer* del gen *Igh* (*E $\mu$* ) y la región 3' no traducida del gen *HBB* humano (gen de la subunidad  $\beta$  de la hemoglobina) (Figura 22B) con el fin de garantizar una alta expresión constitutiva del transgén durante el desarrollo de los linfocitos T, incluyendo a los linfocitos T  $\alpha\beta$  maduros. Para analizar si el transgén *Tcra* DO11.10 contiene todas las secuencias necesarias para la alta expresión del gen *Tcra* endógeno reordenado, comparamos los niveles de transcripción de *Ca* en timocitos DP y linfocitos T  $\alpha\beta$  de ambos ratones

transgénicos en el contexto de ausencia de recombinación V(D)J, *Rag2*<sup>-/-</sup> x DO11.10 y *Rag2*<sup>-/-</sup> x OT-I (Figura 22D).



**Figura 22. Estructura de los transgenes *Tcr* DO11.10 y *Tcr* OT-I. Transcripción de los transgenes *Tcr* en timocitos DP y linfocitos T  $\alpha\beta$  de ratones *Rag2*<sup>-/-</sup> x DO11.10 y *Rag2*<sup>-/-</sup> x OT-I. (A) Representación esquemática de los transgenes *Tcr* DO11.10 y OT-I. Las secuencias derivadas del gen *Tcr* se representan en color azul, mientras que las secuencias que no derivan del gen *Tcr* se identifican con otros colores. Se muestra también la longitud de las secuencias presentes en las construcciones y los reordenamientos V $\alpha$ J $\alpha$ . (B) Mapeo por qPCR de la región genómica de 33,5 kb aguas abajo del reordenamiento *Trav5* – *Traj21* en el transgén *Tcr* DO11.10. Los Ct de las qPCR se han obtenido a partir de la misma cantidad de ADN genómico de ratones *Rag2*<sup>-/-</sup> (puntos rellenos) y de *Rag2*<sup>-/-</sup> x DO11.10 (puntos vacíos). Análisis de *Gja5* se ha usado como control negativo (2 copias en ambas cepas de ratón), mientras que *Ea* se ha usado como control positivo (2 copias en ratones *Rag2*<sup>-/-</sup> y 6 copias en ratones *Rag2*<sup>-/-</sup> x DO11.10). Los resultados representan la media de los duplicados  $\pm$  SEM (n=8). Se usó la prueba t-student no paramétrica con datos no pareados y corrección del Welch para comparar los datos de los puntos rellenos y huecos con el mismo color que corresponden a la misma secuencia, los valores de *p* se representan con asteriscos de la siguiente manera: *p* < 0,005 (\*\*) y *p* < 0,0005 (\*\*\*). (C) Representación a escala del transgén *Tcr* DO11.10. (D) Transcripción de C $\alpha$  en timocitos y en linfocitos T  $\alpha\beta$  de ratones *Rag2*<sup>-/-</sup> x OT-I y *Rag2*<sup>-/-</sup> x DO11.10 (n=5-8). (E) Transcripción de C $\alpha$  en linfocitos T  $\alpha\beta$  de ratones WT y *Rag2*<sup>-/-</sup> x DO11.10 (n=6). (F) Transcripción V $\alpha$ J $\alpha$  de los transgenes *Tcr* en timocitos y linfocitos T  $\alpha\beta$  de ratones de *Rag2*<sup>-/-</sup> x OT-I y *Rag2*<sup>-/-</sup> x DO11.10 (n=6). (G) Análisis comparativo de los transcritos V-C reordenados en los transgenes *Tcr* y *Tcrb* DO11.10 en timocitos DP y linfocitos T  $\alpha\beta$  de ratones *Rag2*<sup>-/-</sup> x DO11.10 (n=6). Los resultados se normalizaron con los transcritos del gen *Actb* y representan la media  $\pm$  SEM de los duplicados de la RT-qPCR de n experimentos independientes. Se usó la prueba t-Student no paramétrica con datos no pareados y corrección del Welch,**

los valores de  $p$  se representan con asteriscos de la siguiente manera:  $p < 0,05$  (\*),  $p < 0,005$  (\*\*) y  $p < 0,0005$  (\*\*\*)

Los resultados indican que el nivel de transcripción de Ca en linfocitos T  $\alpha\beta$  de los ratones transgénicos *Rag2*<sup>-/-</sup> x DO11.10 se encuentra muy disminuida comparada con la de los timocitos DP, pero no se encuentran diferencias entre los linfocitos T  $\alpha\beta$  respecto a los timocitos DP de los ratones *Rag2*<sup>-/-</sup> x OT-I. Estos resultados demuestran que la transcripción del gen *Tcra* en los ratones transgénicos *Rag2*<sup>-/-</sup> x DO11.10, pero no en los ratones transgénicos *Rag2*<sup>-/-</sup> x OT-I, sufre una fuerte disminución durante el paso de timocitos DP a linfocitos T  $\alpha\beta$ . La cuantificación de la transcripción de Ca indica que los niveles de estos transcritos en células T  $\alpha\beta$  son de aproximadamente un 15% de los que se encuentran en timocitos DP en los ratones *Rag2*<sup>-/-</sup> x DO11.10 (Figura 22D). Este nivel de reducción transcripcional es comparable al que se encuentra en los transcritos no codificantes dependientes de Ea derivados del gen *Tcra* endógeno en linfocitos T  $\alpha\beta$  en comparación con los timocitos DP en ratones WT (94). La abundancia relativa de los transcritos Ca que se encuentra en linfocitos T  $\alpha\beta$  de ratones WT y ratones *Rag2*<sup>-/-</sup> x DO11.10 fue similar (Figura 22E), lo que concuerda con la reducción de la transcripción del transgén *Tcra* DO11.10 en linfocitos T  $\alpha\beta$  respecto a timocitos DP, teniendo en cuenta que el ratón homocigótico *Rag2*<sup>-/-</sup> x DO11.10 tiene cuatro copias del transgén *Tcra* (144).

La abundancia de los transcritos Ca presente en timocitos DP de ratones *Rag2*<sup>-/-</sup> x DO11.10 y *Rag2*<sup>-/-</sup> x OT-I fue similar, siendo aproximadamente cuatro veces superior a la encontrada en timocitos DP de ratones WT y *Rag2*<sup>-/-</sup> x  $\beta$ tg (Figuras 16B y 21D), indicando que la mayoría de los transcritos proceden de los respectivos transgenes *Tcra*. Para comparar la transcripción del transgén *Tcra* DO11.10 frente al transgén *Tcra* OT-I, sin tener en cuenta la contribución de los alelos *Tcra* endógenos sin reordenar, comparamos transcritos VaJa específicos en timocitos DP y linfocitos T  $\alpha\beta$  de ratones *Rag2*<sup>-/-</sup> x DO11.10 y *Rag2*<sup>-/-</sup> x OT-I (Figura 22F). El transgén *Tcra* DO11.10 contiene reordenamientos *Trav*5-4/*n* - *Traj*21, mientras que el transgén *Tcra* OT-I contiene reordenamientos *Trav*14 - *Traj*33 (Figura 22A). Nuestros resultados mostraron que los transcritos *Trav*5-4/*n* - *Traj*21 se encontraban muy reducidos en linfocitos T  $\alpha\beta$  comparado con timocitos DP en ratones *Rag2*<sup>-/-</sup> x

DO11.10, mientras que no había diferencias significativas en la abundancia de los transcritos *Trav14 - Traj33* de linfocitos T  $\alpha\beta$  y timocitos DP de ratones *Rag2*<sup>-/-</sup> x OT-I (Figura 22F). De acuerdo con los datos obtenidos del análisis de transcritos Ca del gen *Tcra* endógeno no reordenado (94) y del transgén reordenado *Tcra* DO11.10 (Figura 22D), encontramos que los niveles de transcripción de *Travd5-4/n - Traj21* en linfocitos T eran alrededor del 15% de los que se encontraban en timocitos DP de ratones *Rag2*<sup>-/-</sup> x DO11.10 (Figura 22F).

Tanto el transgén *Tcra* DO11.10 como el transgén *Tcrb* DO11.10 se encuentran cointegrados en el mismo sitio genómico dentro del cromosoma 18 de ratón. El transgén *Tcrb* DO11.10 es el mismo que el transgén expresado en el ratón *Rag2*<sup>-/-</sup> x  $\beta$ tg, y consiste en un reordenamiento natural *Trbv13 - D $\beta$ -Trbj1-1*, cuya transcripción está dirigida por el promotor correspondiente al segmento génico *Trbv13* (140,144). Para evaluar si los efectos de inhibición del transgén *Tcra* observados en linfocitos T  $\alpha\beta$  de ratones *Rag2*<sup>-/-</sup> x DO11.10 se deben a efectos derivados del lugar de inserción de los transgenes en el ratón DO11.10, comparamos la abundancia de transcritos V - C derivados de los transgenes *Tcra* DO11.10 y *Tcrb* DO11.10 en timocitos DP y linfocitos T  $\alpha\beta$  (Figura 22G). Nuestros datos indican que, al contrario de lo observado en linfocitos T  $\alpha\beta$  en comparación con timocitos DP de ratones *Rag2*<sup>-/-</sup> x DO11.10 en los que se observa una reducción de los niveles de transcripción del transgén *Tcra*, los niveles de transcripción para el transgén *Tcrb* fueron similares en ambos tipos celulares. Por tanto, se puede descartar que la influencia del sitio de integración cromosómica sea causal en la inhibición de la transcripción del transgén *Tcra* DO11.10 en el paso de timocitos DP a linfocitos T  $\alpha\beta$ .

Estos datos indican que la colaboración funcional entre *Ea* y un elemento distal, que no se encuentra en el transgén *Tcra* DO11.10 y que se activa después de la recombinación *VaJa* y la selección positiva, sería necesaria para mantener la alta transcripción del gen *Tcra* reordenado en linfocitos T  $\alpha\beta$ .

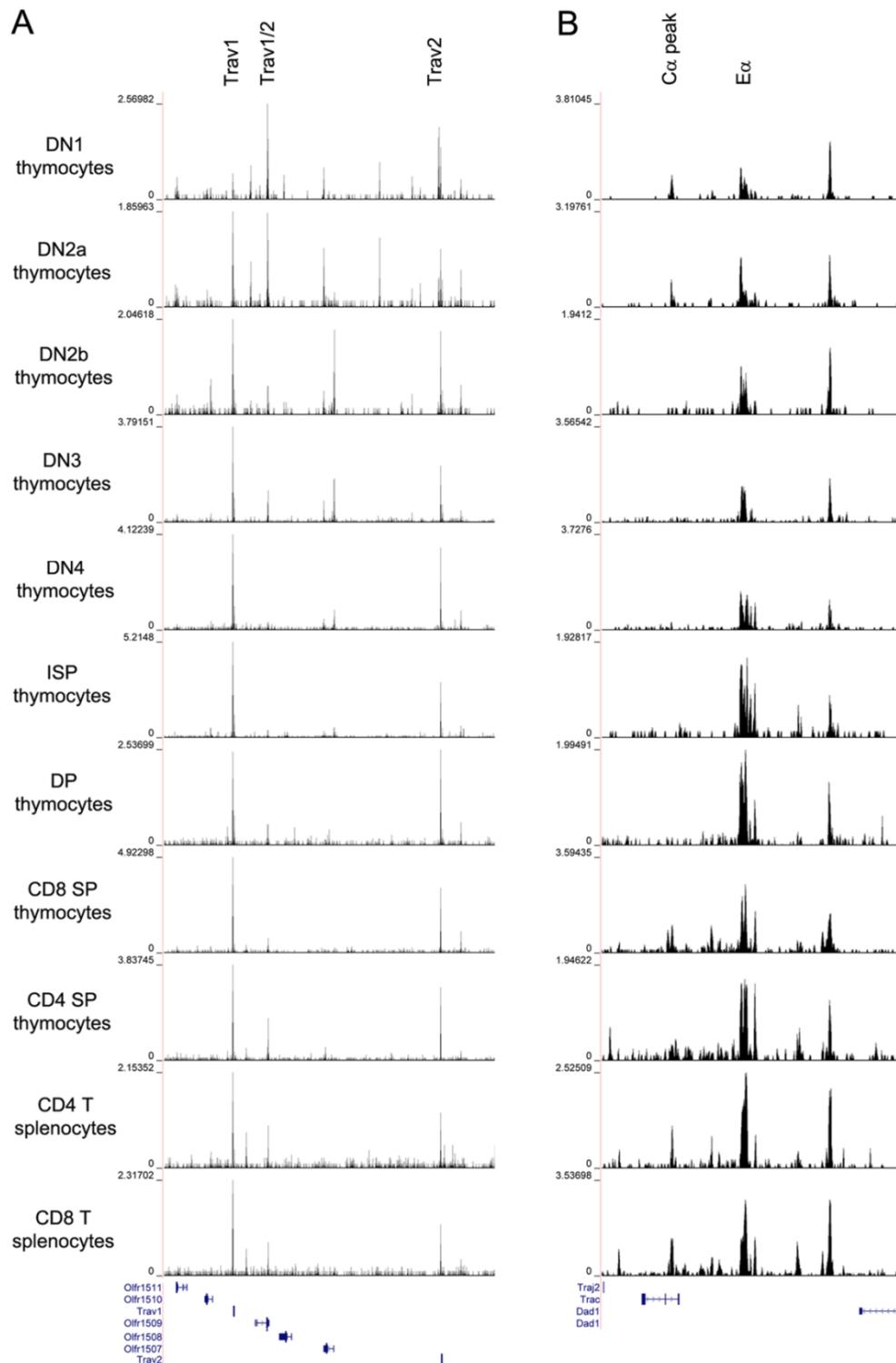
### 3. Análisis de las interacciones globales en el contexto del gen *Tcra* no reordenado en timocitos DP y del gen *Tcra* reordenado en linfocitos T $\alpha\beta$

Nuestros datos muestran que Ea es importante para activar la transcripción del gen *Tcra* reordenado en colaboración con otra región o regiones que no se encuentran en el transgén *Tcra* DO11.10. Con el fin de identificar dicha región colaboradora realizamos análisis comparativos por 4C-seq en timocitos DP de ratones *Rag2*<sup>-/-</sup> previamente inyectados con anticuerpo anti-CD3 para inducir la diferenciación de los timocitos DN3a a DP (con el gen *Tcra* en configuración no reordenada) y en linfocitos T  $\alpha\beta$  de ratones WT (con ambos alelos del gen *Tcra* en configuración reordenada). Cualquier secuencia que pueda interactuar con Ea debe permanecer integrada en el genoma después de que ocurra la recombinación V $\alpha$ J $\alpha$ . Por tanto, esta secuencia colaboradora podría estar localizada aguas arriba del segmento V $\alpha$  situado más 5' en el locus, *Trav1*, o aguas abajo del segmento J $\alpha$  funcional situado más 3', *Traj2*, para asegurar su retención en el contexto del gen *Tcra* reordenado.

Realizamos experimentos duplicados de 4C-seq para analizar de forma global las interacciones con cuatro secuencias diferentes del gen *Tcra* como anzuelos: 1) *Trav1*, que corresponde al promotor y al segmento génico *Trav1*; 2) *Trav1/2*, que corresponde a la secuencia génica del receptor olfativo 1509 localizada entre los segmentos génicos *Trav1* y *Trav2*, la cual se encuentra integrada en todos los reordenamientos V $\alpha$ J $\alpha$  excepto los que involucran al por segmento génico *Trav1*; 3) Ca peak, que corresponde a una región intrónica localizada entre el tercer y el cuarto exón de Ca; y 4) Ea.

La cromatina de las regiones *Trav1* y Ea permanecen accesibles durante el desarrollo de los timocitos DN2 hasta linfocitos T  $\alpha\beta$  según experimentos de ATAC-seq, mientras que las regiones de los anzuelos *Trav1/2* y Ca peak muestran un incremento de la accesibilidad en el paso de timocitos DN4/DP a células T  $\alpha\beta$  según se muestra en los resultados de experimentos de ATAC-seq con poblaciones purificadas de ratones WT ([www.immgen.org](http://www.immgen.org)) (Figura 23). Por tanto, las cuatro regiones seleccionadas consisten en regiones que están constitutivamente accesibles en timocitos DP y linfocitos T  $\alpha\beta$  (*Trav1* y Ea) o

que se abren durante la diferenciación de timocitos DN4/DP a linfocitos T  $\alpha\beta$  (*Trav1/2* y *Ca* peak).

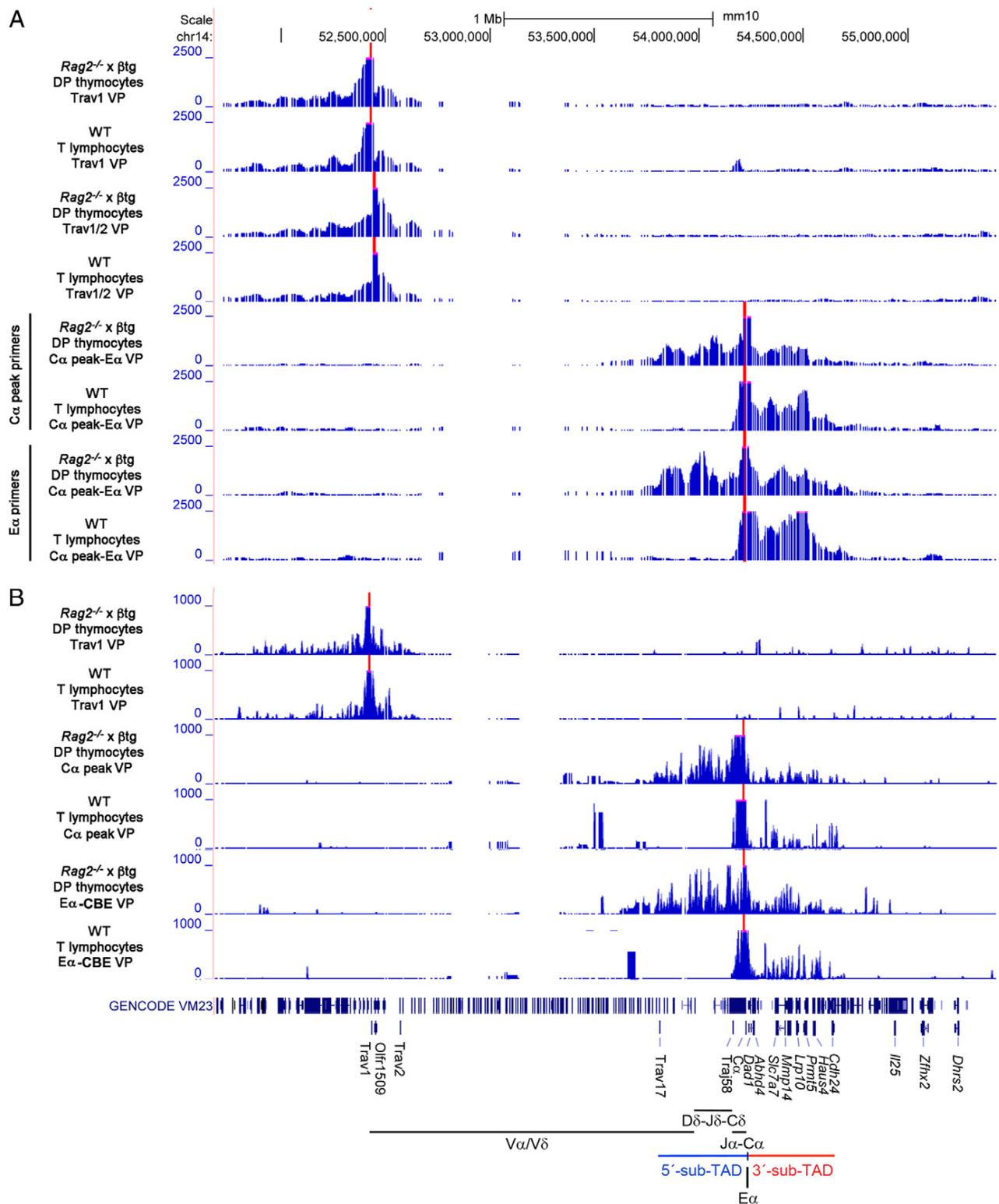


**Figura 23. Análisis de la cromatina mediante ATAC-seq durante la diferenciación de los timocitos a células T  $\alpha\beta$ .** (A) Análisis de la accesibilidad de la cromatina de la región genómica *Trav1–Trav2* en timocitos y linfocitos T  $\alpha\beta$  ([www.immgen.org](http://www.immgen.org)). Se indica la localización de las regiones *Trav1* y *Trav1/2* usadas como anzuelos en los experimentos de 4C-seq. (B) Análisis de la accesibilidad de la cromatina de la región *Ca–E $\alpha$ –Dad1* en timocitos y linfocitos T  $\alpha\beta$  ([www.immgen.org](http://www.immgen.org)). Se indica la localización de las regiones *Ca* peak y *E $\alpha$*  usadas como anzuelos en los experimentos de 4C-seq.

Primero realizamos los experimentos 4C-seq usando la enzima con un sitio de corte de 6 pares de bases HindIII como RE1 (Figura 24A). La digestión con HindIII separa las regiones *Trav1* y *Trav1/2* permitiendo diferenciar claramente entre las interacciones que establecen estas dos secuencias. Sin embargo, las regiones C $\alpha$  peak y E $\alpha$  permanecen unidas en un único fragmento C $\alpha$  peak-E $\alpha$  de 13.444 pb, con la región C $\alpha$  peak localizada en la región 5' y E $\alpha$  localizado en la región 3' de dicho fragmento. Para analizar interacciones específicas para C $\alpha$  peak y E $\alpha$ , se realizaron las PCR inversas utilizando oligonucleótidos orientados hacia cada extremo del fragmento, oligonucleótidos C $\alpha$  peak y E $\alpha$ .

Los experimentos de interacción con los anzuelos *Trav1* y *Trav1/2* no revelaron ninguna interacción relevante a larga distancia con ninguna secuencia, incluyendo a E $\alpha$ , ni en el contexto del gen *Tcra* no reordenado en timocitos DP, ni en el del gen *Tcra* reordenado en linfocitos T  $\alpha\beta$ . La interacción a larga distancia detectada entre *Trav1* y la región *Traj32-Traj23* no se consideró relevante por encontrarse en una región que es delecionada durante el proceso de recombinación. En consonancia con estos datos, los análisis de interacciones del fragmento C $\alpha$  peak-E $\alpha$  no revelaron ninguna interacción con las regiones correspondientes a *Trav1* y *Trav1/2*. Estos resultados son acordes con experimentos previos de 4C-seq usando los anzuelos *Trav1* y C $\alpha$  peak-E $\alpha$  en el contexto del gen *Tcra* no reordenado y reordenado en timocitos DP realizados anteriormente con un modelo *Knock-In* que integra un reordenamiento primario del gen *Tcra* (58).

De acuerdo con estos datos y con la formación de un lazo de cromatina derivado de interacciones entre E $\alpha$  y los promotores de los segmentos V $\alpha$  y J $\alpha$  proximales en timocitos DP (58,89), nuestro análisis 4C-seq mostró interacciones claras entre el fragmento C $\alpha$  peak-E $\alpha$  y los segmentos génicos V $\alpha$  proximales y J $\alpha$  del 5' sub-TAD en el contexto del gen *Tcra* no reordenado en timocitos DP. Como se esperaba (58,89), como consecuencia de la deleción de segmentos génicos V $\alpha$  y J $\alpha$  que ocurren durante los reordenamientos V $\alpha$ J $\alpha$ , las interacciones C $\alpha$  peak-E $\alpha$  con las regiones V $\alpha$  y J $\alpha$  del 5' sub-TAD se pierden en el contexto del gen *Tcra* reordenado en linfocitos T  $\alpha\beta$ .



**Figura 24. Contactos a larga distancia en el ADN en el contexto del gen *Tera* no reordenado en timocitos DP y *Tera* reordenado en linfocitos T  $\alpha\beta$ .** (A) Perfiles de interacción obtenidos mediante 4C-seq de los anzuelos indicados usando HindIII como RE1 en timocitos DP obtenidos de ratones *Rag2*<sup>-/-</sup> inyectados con un anticuerpo anti-CD3 $\epsilon$  y linfocitos T  $\alpha\beta$  de ratones WT. (B) Perfiles de interacción de 4C-seq de los anzuelos indicados usando DpnII como RE1 en timocitos DP obtenidos de ratones *Rag2*<sup>-/-</sup> inyectados con un anticuerpo anti-CD3 $\epsilon$  y linfocitos T  $\alpha\beta$  de ratones WT. Los datos se representan como lecturas por millón de lecturas mapeadas. La línea roja identifica el fragmento usado como anzuelo en cada caso. Los datos son representativos de dos experimentos independientes.

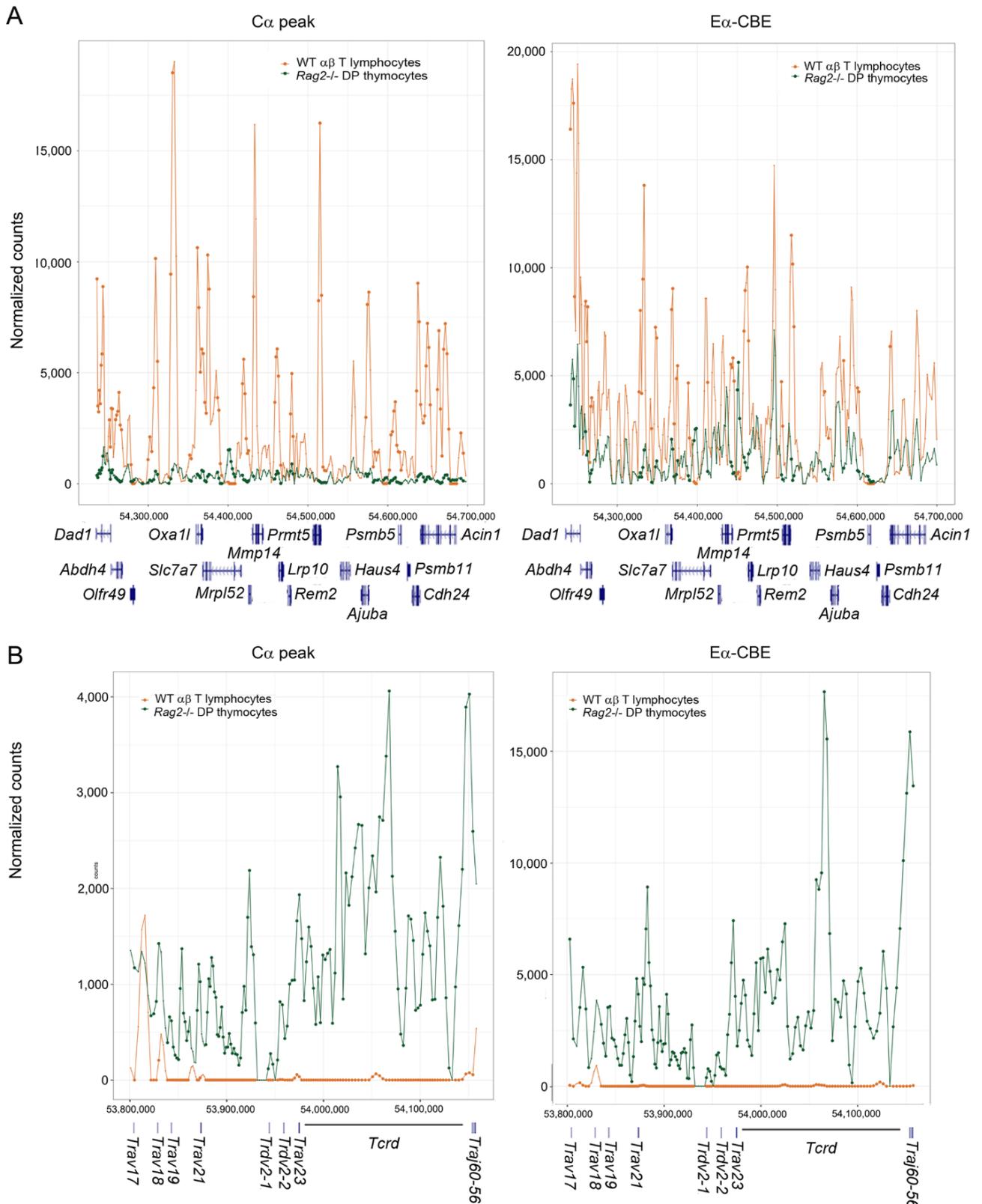
Nuestros análisis también detectaron claras interacciones entre la región Ca peak–Ea y secuencias localizadas en la región colindante del 3´ sub-TAD tanto en timocitos DP con el gen *Tcra* no reordenado como en linfocitos T  $\alpha\beta$  con el gen *Tcra* reordenado, sin encontrar diferencias entre los oligonucleótidos usados para identificar interacciones con la región Ca peak y los oligonucleótidos para identificar interacciones con Ea (Figura 24A). Por lo tanto, las interacciones que se establecen con el fragmento Ca peak–Ea se mantienen estables durante la recombinación VaJa y el desarrollo de timocitos DP a linfocitos T  $\alpha\beta$ .

Para identificar claramente qué interacciones se establecen de forma independiente con las regiones Ca peak y Ea, realizamos experimentos de 4C-seq usando como RE1 la enzima de corte DpnII que reconoce sitios de 4-pares de bases (Figura 24B). La digestión con la enzima DpnII permite separar la región Ca peak en un fragmento de 1.075 pb de la región Ea en un fragmento de 675 pb. Como era de esperar, los experimentos revelaron que Ea establece interacciones claras con los segmentos génicos Va proximales y Ja del 5´ sub-TAD en el contexto del gen *Tcra* no reordenado en timocitos DP que se pierden en el contexto del gen *Tcra* reordenado en linfocitos T  $\alpha\beta$  como resultado de la recombinación VaJa. Los experimentos con el anzuelo Ca peak mostraron interacciones comparables a las de Ea en ambos tipos celulares, indicando que tanto la región Ca peak como Ea contribuyen de forma independiente a la formación del lazo de cromatina en el 5´ sub-TAD presente en timocitos DP.

Igual que ocurrió con la enzima HindIII (Figura 24A), los experimentos con DpnII no revelaron interacciones a largas distancias de Ea o Ca peak con *Trav1* en timocitos DP y linfocitos T  $\alpha\beta$ , apoyando la idea de que *Trav1* no está involucrado en la regulación transcripcional del gen *Tcra* reordenado de forma dependiente de Ea (Figura 24B). Además, los análisis de los experimentos 4C-seq con DpnII mostraron interacciones independientes de las regiones Ca peak y Ea con secuencias del 3´ sub-TAD, tanto en el contexto del gen *Tcra* no reordenado en timocitos DP como en el del gen *Tcra* reordenado en linfocitos T  $\alpha\beta$  (Figura 24B).

Nuestros datos de 4C-seq indican la formación de un lazo de cromatina entre Ea y el 3´ sub-TAD (Figura 24), lo cual coincide con la detección de una

estructura anclada al *enhancer* que se extiende por el 3' sub-TAD en mapas de interacciones obtenidos mediante Hi-C (99). Para cuantificar los posibles cambios de interacción entre C $\alpha$  peak y E $\alpha$  con secuencias del 3' sub-TAD en timocitos DP de ratones *Rag2*<sup>-/-</sup> y linfocitos T  $\alpha\beta$  de ratones WT en nuestros experimentos de 4C-seq realizados con DpnII, realizamos un análisis comparativo usando el programa 4C-ker (163) (Figura 25). El perfil de picos que muestra este análisis indica que las interacciones entre C $\alpha$  peak y E $\alpha$  con múltiples regiones del 3' sub-TAD son más fuertes en linfocitos T  $\alpha\beta$  de ratones WT que en timocitos DP de ratones *Rag2*<sup>-/-</sup> (Figura 25A). Estos resultados indican las interacciones que se establecen entre C $\alpha$  peak y E $\alpha$  con el 3' sub-TAD están reguladas dinámicamente durante el desarrollo de los linfocitos T. De acuerdo con la formación del lazo de cromatina entre C $\alpha$  peak y E $\alpha$  con las regiones V $\alpha$  proximales y J $\alpha$  del 5' sub-TAD en timocitos DP y su pérdida como consecuencia de la recombinación V $\alpha$ J $\alpha$  en linfocitos T  $\alpha\beta$  de ratones WT (58,89), nuestro análisis de 4C-ker también cuantificó la pérdida de dichas interacciones de los dos anzuelos en linfocitos T  $\alpha\beta$  de ratones WT en comparación con timocitos DP de ratones *Rag2*<sup>-/-</sup> (Figura 25B). Estos análisis claramente demuestran que el tanto C $\alpha$  peak como E $\alpha$  establecen una mayor interacción con el 3' sub-TAD en el contexto del gen *Tcra* reordenado en linfocitos T  $\alpha\beta$ , que en el contexto del gen *Tcra* no reordenado en timocitos DP.



**Figura 25. Cuantificación de las interacciones de las regiones C $\alpha$  peak y E $\alpha$  con regiones del 5' sub-TAD y del 3' sub-TAD.** Las señales de los experimentos de C $\alpha$  peak y E $\alpha$  4C-seq realizados con DpnII en linfocitos T  $\alpha\beta$  de ratones WT (gráficos naranjas) y timocitos DP de ratones *Rag2*<sup>-/-</sup> inyectados con anti-CD3 $\epsilon$  (gráficos verdes) se cuantificaron mediante el programa 4C-ker de experimentos duplicados. Se muestran interacciones de los anzuelos indicados y el 3' sub-TAD (A) y con la región V $\alpha$  proximal en el 5' sub-TAD (B). Los círculos muestran las diferencias significativas.

## CONCLUSIONES



## CONCLUSIONES

1. E $\alpha$ , a pesar de estar fuertemente inhibido en linfocitos T  $\alpha\beta$ , con un 15% de su actividad en comparación con la actividad que tiene en timocitos DP, es necesario para la transcripción normal de todos los alelos reordenados del gen *Tcra*, incluyendo aquellos en los que la transcripción está dirigida por los promotores asociados con los segmentos génicos *Trav14* recombinados y responsables de presencia de linfocitos V $\alpha$ 2 en ratones E $\alpha^{-/-}$ .
2. E $\alpha$  es necesario para la transcripción del gen *TCRA* reordenado humano y la expresión del TCR $\alpha\beta$  en células Jurkat, y por extensión en linfocitos T  $\alpha\beta$  humanos.
3. El hecho que no se observe una bajada de la transcripción del gen *Tcra* cuando se compara en el contexto del gen *Tcra* no reordenado en timocitos DP y en el contexto del gen *Tcra* reordenado en linfocitos T  $\alpha\beta$  indica la necesidad de la colaboración funcional de E $\alpha$  con elementos reguladores adicionales tras la recombinación V $\alpha$ J $\alpha$  y la selección positiva para mantener alta la transcripción del gen *Tcra* reordenado en linfocitos T  $\alpha\beta$ .
4. La transcripción del gen *Tcra* en los contextos del transgén *Tcra* DO11.10 y del gen *Tcra* endógeno sin reordenar se inhibe de forma idéntica en el paso de timocitos DP a linfocitos T  $\alpha\beta$ , lo cual indica que la configuración génica resultante de los reordenamientos V $\alpha$ J $\alpha$  no evita la inhibición de E $\alpha$ . Por tanto, la configuración génica resultante de los reordenamientos V $\alpha$ J $\alpha$  no da lugar a un aumento de la transcripción desde los promotores V $\alpha$  reordenados por encontrarse relativamente cerca de E $\alpha$ .
5. La inhibición de la transcripción del transgén *Tcra* DO11.10 durante la transición de los timocitos DP a linfocitos T  $\alpha\beta$  no es consecuencia de su integración en un sitio de cromatina represivo de la transcripción.
6. La inhibición observada en la transcripción del transgén *Tcra* DO11.10 en el paso de los timocitos DP a linfocitos T  $\alpha\beta$  indica que la/s región/es que colaboradora/n con E $\alpha$  para la transcripción normal del gen *Tcra* reordenado en linfocitos T  $\alpha\beta$  no se encuentra dentro de la estructura de este transgén.

7. E $\alpha$  y C $\alpha$  peak interaccionan con las mismas regiones distantes dentro del 5' sub-TAD en el contexto del gen *Tcra* no reordenado en timocitos DP, contribuyendo de forma independiente a la formación de un lazo de cromatina.
8. E $\alpha$  y C $\alpha$  peak no establecen nuevas interacciones con regiones del 5' sub-TAD en el contexto del gen *Tcra* reordenado en linfocitos T  $\alpha\beta$ .
9. Tanto la región C $\alpha$  peak como E $\alpha$  contribuyen de forma independiente a la formación de lazo de cromatina con el 3' sub-TAD en el contexto del gen *Tcra* no reordenado en timocitos DP y en el contexto del gen *Tcra* reordenado en linfocitos T  $\alpha\beta$ .
10. Tanto la región de C $\alpha$  peak como E $\alpha$  establecen una mayor interacción con el 3' sub-TAD en el contexto del gen *Tcra* reordenado en linfocitos T  $\alpha\beta$ , que en el contexto del gen *Tcra* no reordenado en timocitos DP, la cual podría estar implicada en evitar la bajada de transcripción del gen *Tcra* reordenado en linfocitos T  $\alpha\beta$ .

## DISCUSIÓN



Los *enhancers* activan la transcripción de sus genes diana para la determinación de linajes celulares durante el desarrollo mediante interacciones físicas con sus promotores específicos localizados a largas distancias genómicas, estando estas interacciones restringidas normalmente dentro del sub-TAD donde se encuentran tanto el *enhancer* como los promotores. Así pues, la compartimentalización del genoma en sub-TAD favorece las interacciones promotor – *enhancer*. Sin embargo, esta situación puede ser diferente para los *enhancers* que se encuentran en el límite entre dos sub-TAD, como es el caso de *Ea*, situado entre el locus *Tcra/Tcrd* presente en el 5' sub-TAD y los genes *Dad1-Chd24* presentes en el 3' sub-TAD. Aunque este trabajo se enfoca en la regulación durante el desarrollo de la actividad de *Ea* sobre la activación de la transcripción del gen de la cadena TCR $\alpha$  situado en el 5' sub-TAD, nuestros datos indican la colaboración de este *enhancer* con regiones presentes en el 3' sub-TAD.

Nuestro trabajo identifica que la actividad de *Ea* es importante no sólo para la activación transcripción germinal y la recombinación V $\alpha$ J $\alpha$  del gen de la cadena TCR $\alpha$  en timocitos DP y para la transcripción del gen reordenado de la cadena TCR $\delta$  en linfocitos T  $\gamma\delta$  (22), sino también para la transcripción del gen reordenado de la cadena TCR $\alpha$ , tanto en ratones como en humanos. Además, este trabajo describe el requerimiento de una cooperación funcional entre *Ea* y secuencias distantes no presentes en el transgén *Tcra* DO11.10 para la transcripción alta y normal del gen reordenado en linfocitos T  $\alpha\beta$ . Por último, nuestro trabajo también identifica que *Ea* junto con la secuencia *Ca* peak establece fuertes interacciones con secuencias presentes en el 3' sub-TAD en el contexto del gen *Tcra* reordenado en linfocitos T  $\alpha\beta$ , de forma que dichas interacciones podrían impedir una bajada en la transcripción del gen reordenado de la cadena TCR $\alpha$  y de la expresión del TCR $\alpha\beta$  como consecuencia de la inhibición de *Ea* en linfocitos T  $\alpha\beta$ .

Los resultados previos indicaban que a pesar de la acusada disminución del 85% de la actividad de *Ea* detectada en timocitos SP y linfocitos T  $\alpha\beta$  (94), la transcripción del gen *Tcra* reordenado y la expresión del TCR $\alpha\beta$  se

mantienen muy altas en linfocitos T  $\alpha\beta$  (94,97,164). Estos datos sugirieron que la transcripción del gen *Tcra* reordenado en timocitos SP y linfocitos T  $\alpha\beta$  estuviera regulada por un mecanismo independiente de  $E\alpha$ , o bien que este *enhancer* necesitase colaborar con otra región distante que no se elimine durante el proceso de recombinación  $V\alpha J\alpha$  para asegurar dicha transcripción en estas células (94). Nuestros resultados demuestran claramente que  $E\alpha$ , a pesar de estar fuertemente inhibido con solo un 15% de su actividad comparada con la que tiene en timocitos DP, es necesario para la transcripción del gen reordenado de la cadena  $TCR\alpha$  y la expresión normal del  $TCR\alpha\beta$  en linfocitos T  $\alpha\beta$ , tanto en humanos como en ratones. Además, nuestros datos demuestran la existencia de una colaboración de  $E\alpha$  con regiones del 5' sub-TAD, como es el caso de  $C\alpha$  peak, y del 3' sub-TAD para la regulación transcripcional del gen *Tcra* durante el desarrollo de los linfocitos T  $\alpha\beta$ .

El análisis fenotípico inicial de los ratones  $E\alpha^{-/-}$  mostró que estos animales solo tienen linfocitos T  $V\alpha 2^+$ , derivados de reordenamientos con los segmentos *Trav14* (22). Estos ratones tienen una deficiencia en la transcripción germinal del gen *Tcra* no reordenado y la activación de la recombinación  $V\alpha J\alpha$  en timocitos DP, además de una inhibición en la transcripción del gen *Tcrd* reordenado. Experimentos subsecuentes revelaron que los reordenamientos *Trav14 - J\alpha* que dan lugar a la cadena  $TCR\alpha$   $V\alpha 2^+$  derivan de reordenamientos  $V\alpha J\alpha$  activados por  $E\delta$  en timocitos DN2b y DN3a (130). De hecho,  $E\delta$  es capaz de activar todos los reordenamientos  $V\alpha J\alpha$  en estas células de una forma muy minoritaria, pero solo aquellos que contienen *Trav14* pueden transcribir el gen *Tcra* reordenado de forma independiente de *enhancers* en estos ratones (estos alelos carecen de  $E\delta$  y  $E\alpha$  en los ratones  $E\alpha^{-/-}$ ) para permitir su selección positiva en el timo (22,130). Aunque los estudios previos habían comparado la transcripción del gen *Tcra* reordenado en linfocitos T  $\alpha\beta$  de ratones  $E\alpha^{-/-}$  y WT, esta comparación se realizó entre linfocitos T  $V\alpha 2^+$  de ratones  $E\alpha^{-/-}$  y linfocitos T  $\alpha\beta$  totales de ratones WT, los cuales la mayoría son  $V\alpha 2^-$  (22). Nuestra comparación de la transcripción del gen *Tcra* reordenado en linfocitos T  $V\alpha 2^+$  en ratones  $E\alpha^{-/-}$  y ratones WT es, por tanto, más apropiada y permite confirmar no solo que la transcripción del gen *Tcra* reordenado es comparable entre linfocitos T  $V\alpha 2^+$  y  $V\alpha 2^-$ , sino que  $E\alpha$  es

importante incluso para la transcripción normal de los alelos reordenados de *Trav14*.

Debido a que la mayoría de los linfocitos T  $\alpha\beta$  expresan solo una cadena TCR $\alpha$  (22,165), las diferencias observadas en la expresión en superficie de TCR $\alpha\beta$  y Va2<sup>+</sup> entre linfocitos T E $\alpha^{-/-}$  y WT no se explica por la expresión de dos cadenas Va2<sup>+</sup> en un pequeño porcentaje de linfocitos T  $\alpha\beta$  WT. Además, aunque los linfocitos T  $\alpha\beta$  WT contienen dos alelos TCR $\alpha$  reordenados, solo uno de ellos es funcional en la mayoría de las células (165), siendo los transcritos procedentes del reordenamiento VaJa improductivo degradados (166). Nuestros resultados muestran de forma inequívoca que E $\alpha$  es necesario para la transcripción del gen *Tcra* reordenado y la expresión adecuada del TCR $\alpha\beta$  en timocitos DP y linfocitos T Va2<sup>+</sup>. Por tanto, E $\alpha$  es necesario para la transcripción normal de todos los alelos del gen *Tcra* reordenados, incluyendo aquellos en los que la transcripción está dirigida por los promotores asociados con los segmentos génicos *Trav14*.

Nuestros experimentos para generar células Jurkat E $\alpha^{-/-}$  mediante el uso de técnicas de edición CRISPR/Cas9 para delecionar el fragmento del *enhancer* humano equivalente al fragmento de algo más de 1 kb previamente delecionado en los ratones E $\alpha^{-/-}$  fue exitoso. Aunque la aproximación inicial de introducir de forma independiente las dos gARN que flanquean E $\alpha$  en plásmidos de expresión de GFP fue infructuosa, tuvimos éxito cuando las dos gARN flanqueantes del *enhancer* fueron clonadas por parejas en el mismo plásmido de expresión de GFP. De esta manera se consiguieron distintos clones Jurkat E $\alpha^{+/-}$  mediante el uso de distintas combinaciones de gARN flanqueantes. El hecho que E $\alpha$  funcione en *cis* nos indicó que aquéllos clones E $\alpha^{+/-}$  en los que se observó expresión del TCR $\alpha\beta$  significaba que la delección había ocurrido en el alelo improductivo del gen *Tcra*, mientras que aquéllos en los que no se observó expresión del TCR $\alpha\beta$  significaba que la delección había ocurrido en el alelo productivo de este gen. La obtención de varios clones E $\alpha^{+/-}$  TCR<sup>-</sup> con distintas parejas de gARN descartó la aparición de fenómenos funcionales derivados del reconocimiento inespecífico de las gARN que afectasen al fenotipo observado. Para delecionar E $\alpha$  en el segundo alelo se eligió un clon E $\alpha^{+/-}$  TCR $\alpha\beta$ <sup>+</sup> obtenido con las gARN más externas, clon IV.16,

de forma que las gARN internas solo pudieran reconocer sus secuencias diana en el alelo productivo. De esta forma se consiguió de forma exitosa clones  $E\alpha^{-/-}$ , los cuales como era de esperar no expresaron TCR $\alpha\beta$ . Así pues, nuestros datos muestran una estrategia mediante el uso secuencial de gARN externas e internas que ha probado ser útil para la delección de secuencias en ambos alelos del gen *Tcra* mediante la tecnología CRISPR/Cas9.

En promedio, aproximadamente el 80% del total de transcritos TCR $\alpha$  que se encuentran en linfocitos T  $\alpha\beta$  maduros tienen reordenamientos V $\alpha$ J $\alpha$  con el correcto marco de lectura, alrededor del 16% tienen reordenamientos V $\alpha$ J $\alpha$  no tienen el correcto marco de lectura y sobre el 3-4% de los transcritos derivan de la configuración germinal no reordenada (167). La situación promedio de los linfocitos T  $\alpha\beta$  se ejemplifica en células Jurkat con prácticamente un 100% de los transcritos C $\alpha$  y V $\alpha$ J $\alpha$  derivados del alelo productivo. Nuestros datos en células Jurkat  $E\alpha^{+/-}$  y  $E\alpha^{-/-}$  demuestran que el *enhancer* es necesario para la transcripción del gen *TCRA* reordenado y la expresión normal del TCR $\alpha\beta$  en linfocitos T  $\alpha\beta$  humanos.

Nuestros resultados muestran además que, de forma similar a la reducción de los transcritos no codificantes observados en el contexto del gen *Tcra* endógeno no reordenado (94), la transcripción de C $\alpha$  y V $\alpha$ J $\alpha$  se encuentra muy reducida en el contexto del transgén *Tcra* DO11.10 reordenado, pero no en el contexto del transgén *Tcra* OT-I. El transgén *Tcra* DO11.10 está regulado por una región genómica de 33,5 kb que incluye desde *Traji13* al tercer exón de *Dad1*, mientras que el transgén *Tcra* OT-I está regulado por regiones reguladoras no relacionadas con el gen *Tcra*. Aunque los datos muestran que los niveles de los transcritos de C $\alpha$  presentes en linfocitos T  $\alpha\beta$  de ratones WT y *Rag2* $^{-/-}$  x DO11.10 son aparentemente similares (Figura 22E), hay que considerar que los linfocitos T  $\alpha\beta$  de ratones *Rag2* $^{-/-}$  x DO11.10 contienen cuatro copias del transgén del *Tcra* reordenado, además de los dos alelos no reordenados del gen *Tcra* endógeno (144). Teniendo en cuenta el número de copias del transgén *Tcra* DO11.10 presente en estos ratones, los datos indican que el nivel de transcripción de C $\alpha$  derivado de las cuatro copias del transgén *Tcra* DO11.10 es comparable al derivado del único alelo reordenado de forma productiva del gen *Tcra* endógeno presente en células T  $\alpha\beta$  WT (teniendo en

cuenta que los transcritos derivados del alelo *Tcra* endógeno reordenado de forma no productiva se degradan), lo cual es consistente con la reducción del 70-85% de la transcripción del transgén *Tcra* DO11.10 observada en el paso de timocitos DP a linfocitos T  $\alpha\beta$ . Estos resultados confirman que la configuración génica resultante de los reordenamientos *V $\alpha$ J $\alpha$*  no evita en sí la inhibición de *E $\alpha$*  (160), por tanto se puede descartar que el reordenamiento del gen *Tcra* no da lugar a una configuración especial que induzca una fuerte activación del promotor *V $\alpha$*  reordenado por encontrarse en una posición relativamente cercana a *E $\alpha$* . Además, el hecho de que se detecte una reducción de la transcripción similar en el transgén *Tcra* DO11.10 y en el gen *Tcra* endógeno sin reordenar en linfocitos T  $\alpha\beta$  implica la necesidad de una secuencia reguladora adicional presente en el gen endógeno y ausente en el transgén que es necesaria tras la recombinación *V $\alpha$ J $\alpha$*  para mantener los niveles normales de transcripción del gen *Tcra* reordenado en linfocitos T  $\alpha\beta$ . La reducción de la transcripción de *C $\alpha$*  dependiente de *E $\alpha$*  en linfocitos T  $\alpha\beta$  en el contexto del gen *Tcra* endógeno no reordenado (94) y en el transgén *Tcra* DO11.10 reordenado indica además que la secuencia colaboradora con *E $\alpha$*  se activa tras la recombinación *V $\alpha$ J $\alpha$*  y la selección positiva.

Se han realizado anteriormente intentos de identificar el elemento regulador capaz de proporcionar una transcripción alta y estable de transgenes *Tcra* reordenados en linfocitos T  $\alpha\beta$  en colaboración con *E $\alpha$* . Se ha descrito que el fragmento genómico *SacI-SalI* de 17,95-kb que va desde una región 5' de *Traj3* hasta el límite del gen *Dad1*, confiere niveles normales de expresión en un transgén *Tcra* *HY $\alpha$*  que se transcribe desde el promotor endógeno *Trav17* en ratones transgénicos (168). El elemento relevante dentro de esta región que pareció ser el responsable de esta regulación fue localizado en experimentos posteriores en un fragmento de 3,9-kb que incluye desde *Traj3*- al exón 1 de *C $\alpha$* , mediante el uso de un transgén reportero de CD2 humano en el contexto de un cromosoma bacteriano (169). En contraste con estos datos, nuestros resultados indican que la región genómica de 33,5-kb que incluye desde *Traj13* al exón 3 de *Dad1* que está presente en el transgén *Tcra* DO11.10, en la que están incluidos tanto el fragmento *Traj3 – Dad1* de 17,95 kb como el fragmento *Traj3 – exón 1* de *C $\alpha$*  de 3,9 kb descritos anteriormente, no es suficiente para evitar la reducción de la transcripción en la transición de timocitos DP a

linfocitos T  $\alpha\beta$  de dicho transgén. La razón de los resultados discrepantes observados entre los transgenes *Tcra* DO11.10 y *Tcra* HY $\alpha$  podría estar relacionada con la integración de este último en una localización cromosómica que impida su inhibición transcripcional en linfocitos T  $\alpha\beta$ . En el caso del transgén reportero de CD2 humano construido en el cromosoma bacteriano junto con el fragmento *Traj3* – exón 1 de *Ca* de 3,9-kb, la razón de los resultados discrepantes podría deberse al hecho de que la actividad de esta secuencia ha sido evaluada en el contexto de una construcción reportera artificial (169).

Puesto que la transcripción del transgén *Tcrb* DO11.10, el cual se encuentra cointegrado con el transgén *Tcra* DO11.10, es estable en timocitos DP y linfocitos T  $\alpha\beta$ , se puede descartar que la reducción de la transcripción del transgén *Tcra* DO11.10 se deba a un efecto inhibitorio derivado del sitio de integración de estos transgenes.

Por tanto, nuestros datos indican que dicha región reguladora, ha de encontrarse fuera de las secuencias presentes en el transgén *Tcra* DO11.10. La localización de la región colaboradora está restringida a regiones aguas arriba de *Trav1* (ya que toda la región entre *Trav1* y *TraJ2* es susceptible de ser eliminada por la recombinación  $V\alpha J\alpha$ ) o aguas abajo del tercer exón de *Dad1* en el 3' sub-TAD (ya que este es el límite 3' del transgén *Tcra* DO11.10). Con vistas a localizar la secuencia clave que colabora con *Ea* en la transcripción del gen *Tcra* reordenado, teniendo en cuenta la base de datos de Immgen ([www.immgen.org](http://www.immgen.org)) que identifica las regiones abiertas de cromatina identificadas mediante ATAC-seq, se seleccionaron cuatro anzuelos para ser utilizados en experimentos de 4C-seq: *Trav1*, *Trav1/2*, *Ca peak* y *Ea*. El anzuelo *Trav1* fue seleccionado por constituir la secuencia situada más 5' del gen que no se pierde tras la recombinación  $V\alpha J\alpha$  y que además se encuentra en configuración abierta durante todo el desarrollo de los linfocitos T  $\alpha\beta$ . El anzuelo *Trav1/2* fue seleccionado por estar presente en todos los reordenamientos  $V\alpha J\alpha$  excepto los que participa *Trav1*, y aumentar su accesibilidad en timocitos DP a T  $\alpha\beta$ , de forma coincidente con la presencia del gen *Tcra* en un estado reordenado. El anzuelo *Ca peak* fue elegido por estar presente en todos los alelos reordenados y aumentar su accesibilidad en el paso

de timocitos DP a SP. El anzuelo  $E\alpha$  fue elegido por ser la secuencia clave en la regulación transcripcional del gen *Tcra*, tanto en el contexto no reordenado como reordenado.

Nuestros experimentos de 4C-seq revelaron que tanto  $C\alpha$  peak como  $E\alpha$  son capaces de establecer interacciones de forma independiente con el 5' sub-TAD y el 3' sub-TAD. Así pues, hemos identificado una nueva secuencia,  $C\alpha$  peak, que junto con  $E\alpha$ , es capaz de establecer interacciones con ambos sub-TAD. Además, se observó que  $E\alpha$  y  $C\alpha$  peak no establecen nuevas interacciones a largas distancias con secuencias del gen *Tcra* que permanecen integradas tras la recombinación  $V\alpha J\alpha$ , sin embargo, se detectaron claras interacciones de estas dos regiones, tanto de forma independiente como conjunta, con regiones del 3' sub-TAD en timocitos DP y linfocitos T  $\alpha\beta$ . Por lo tanto, la región  $C\alpha$  peak- $E\alpha$  interacciona con ambos sub-TAD en timocitos DP y linfocitos T  $\alpha\beta$ . Estos datos, en consonancia con resultados previos en timocitos DP (99), indican que la región  $C\alpha$  peak- $E\alpha$  es capaz de superar la barrera formada por la región LCR HS2-6, la cual se sugirió previamente que aislaba el 3' sub-TAD de la influencia de  $E\alpha$ . Así como al locus *Tcra/Tcrd* de la influencia de los genes localizados en el 3' sub-TAD (126,128).

Nuestros análisis de 4C-ker indican además que las interacciones de la región la región  $C\alpha$  peak- $E\alpha$  con el 3' sub-TAD se intensifican muy fuertemente en el paso desde un contexto del gen *Tcra* no reordenado en timocitos DP a un contexto del gen *Tcra* reordenado en linfocitos T  $\alpha\beta$ . Estos datos sugieren que este refuerzo de las interacciones de  $E\alpha$  y  $C\alpha$  peak con regiones situadas en el 3' sub-TAD podría impedir que la bajada de actividad del *enhancer* que ocurre en esta transición se traduzca en una pérdida de expresión de la cadena TCR $\alpha$  en linfocitos T  $\alpha\beta$ . Actualmente se desconoce el mecanismo molecular que regularía este proceso. Un descenso en la función aislante de la región LCR HS2-6 durante el desarrollo de las células T podría ser una posible causa por la que se detecten mayores interacciones entre la región  $C\alpha$  peak- $E\alpha$  y el 3' sub-TAD en linfocitos T  $\alpha\beta$  que en timocitos DP.

Los mapas previos de Hi-C realizados en timocitos DP indicaron que los dos sub-TAD están físicamente segregados, mostrando interacciones de  $E\alpha$

que se extienden hasta los segmentos  $V\alpha$  proximales en el 5' sub-TAD y hasta *Chd24* en el 3' sub-TAD (99). Por tanto, estos experimentos confirman que la localización de este *enhancer* en el límite entre dos sub-TAD le permite la interacción hacia ambos lados. El rango de interacciones de  $E\alpha$  (y muy posiblemente también del fragmento  $C\alpha$  peak- $E\alpha$ ) se extiende a lo largo de toda la extensión del 5' sub-TAD que ocupa 525 kb y del 3' sub-TAD que ocupa 400 kb, en un área total cercana a 1 Mb, lo cual identifica a este *enhancer* como uno de los más potentes descritos en eucariotas en cuanto al área de influencia que abarca en su acción.

Los mapas de Hi-C en timocitos DP de ratones  $E\alpha CBE^{-/-}$  mostraron que la longitud e intensidad de las líneas que identifican las interacciones de  $E\alpha$  con el 5' sub-TAD y el 3' sub-TAD están debilitadas respecto a las observadas en los mapas de ratones WT (99). Estos resultados sugirieron que  $E\alpha$  core u otras regiones cercanas situadas 5' de los  $E\alpha CBE$  (presentes en los ratones  $E\alpha CBE^{-/-}$ ) participan en la organización de los sub-TAD a través de la extrusión asimétrica de la cromatina en timocitos DP. En consonancia con esta interpretación, nuestros datos indican que la región  $C\alpha$  peak participa en el establecimiento de interacciones físicas con ambos sub-TAD. La ubicación de la región  $C\alpha$  peak- $E\alpha$  en el límite entre ambos sub-TAD junto con su capacidad para de establecer interacciones específicas hacia ambos lados sugiere que esta región funciona como un conector entre ambos sub-TAD. El mecanismo molecular de cómo esta región opera para esta función podría estar relacionado con la identificación de posiciones variables en los límites entre dos sub-TAD en células individuales, lo cual podría ser la consecuencia de una diferente interacción entre dos sub-TAD vecinos en cada célula (170).

Nuestros datos de detección de interacciones de  $E\alpha$  con regiones del 3' sub-TAD abren la puerta a una posible regulación transcripcional de los genes presentes en esta región por el *enhancer*. Análisis recientes, incluyendo un estudio de nuestro laboratorio, han revelado que  $E\alpha$  es capaz de regular la transcripción de ciertos genes en el 3' sub-TAD en timocitos DP y linfocitos T  $\alpha\beta$  (99,100). En concreto, se ha demostrado que  $E\alpha$  es capaz de activar la transcripción de genes situados en los extremos del 3' sub-TAD: *Dad1* y *Abdh4* localizados en el extremo 5', y *Prmt5*, *Haus4*, *Ajuba* y *Chd24* localizados en el

extremo 3'; sin embargo, los genes localizados en el centro de este sub-TAD, como *Slc7a7*, *Mmp14* y *Lrp10*, son independientes del *enhancer*. Estos datos sugieren que los genes localizados en los extremos del 3' sub-TAD tienen una mayor probabilidad de interactuar con las regiones E $\alpha$  y C $\alpha$  peak. Este efecto podría derivarse de la extrusión alélica asimétrica mediada por los E $\alpha$ CBE, teniendo los bordes de la extrusión más probabilidades de interactuar que la zona interna del lazo extruido.

Datos recientes de nuestro laboratorio han ahondado en la regulación de los genes del 3' sub-TAD por E $\alpha$ , observándose el inesperado hallazgo de que el *enhancer* es capaz de activar la transcripción de *Dad1* y *Haus4* en timocitos DN3a y células no-T, cuando el *enhancer* está definido como "inactivo" ya que no es capaz de activar la transcripción del gen TCR $\alpha$  (60,100). El hecho de que los efectos observados ocurran solo sobre genes específicos apoya la idea que sea consecuencia de la activación dependiente de E $\alpha$  de determinados promotores en estas células. Estos resultados indican que la definición previa de E $\alpha$  como "inactivo" en timocitos DN3a, "activo" en timocitos DP e "inhibido" en linfocitos T  $\alpha\beta$ , basada exclusivamente en su capacidad para activar la transcripción germinal no codificante del gen *Tcra* no reordenado durante el desarrollo de las células T (23,94), debe ser revisada.

E $\alpha$  posee una configuración de cromatina abierta durante todo el desarrollo de los linfocitos T (Figura 23) ([www.immgen.org](http://www.immgen.org)). La formación del *enhanceosoma* de E $\alpha$  en timocitos DN3a es consecuencia de la unión cooperativa de múltiples factores de transcripción constitutivos a la cromatina del *enhancer*, antes de su activación para inducir la transcripción del gen *Tcra* durante la selección  $\beta$  (23,135,136,138,171). La señal del pre-TCR activa la transcripción germinal del gen *Tcra* en timocitos DP a través del reclutamiento de factores de transcripción inducibles y la disociación de los factores supresores HOXA, lo que resulta en el inicio de esta transcripción no codificante del gen *Tcra* que desencadena los reordenamientos V $\alpha$ J $\alpha$  (60,138). Por lo tanto, el reclutamiento de factores de transcripción inducibles mediado por el pre-TCR y la pérdida de los factores HOXA del *enhancesosoma* presente en timocitos DP activa E $\alpha$ . Esta activación es necesaria para la activación de la transcripción de los promotores V $\alpha$  proximales y J $\alpha$  en el 5' sub-TAD (60,138),

y es también probablemente necesaria para la activación de los promotores *Abdh4* y *Chd24* en el 3' sub-TAD en timocitos DP, pero no para los promotores *Dad1* y *Haus4* del 3' sub-TAD que están constitutivamente activados por el *enhancer* en timocitos DN3a. Actualmente se desconoce cómo se controla la distinta regulación temporal de *Eα* para la activación transcripcional de sus genes diana en ambos sub-TAD.

Es interesante que varias de las proteínas codificadas por los genes dependientes de *Eα* en los extremos del 3' sub-TAD son relevantes durante el desarrollo y la función de los linfocitos T. Entre estos genes destacan *Dad1*, *Prmt5* y *Ajuba*. DAD1 regula la N-glicosilación, la cual es importante para la prevención de la muerte celular por apoptosis y su sobreexpresión está relacionada con la activación de los linfocitos T (157–159). PRMT5 es una arginina metiltransferasa que cataliza la dimetilación simétrica de argininas de muchas proteínas, entre las que se encuentran las presentes en el residuo 3 de la histona H4 y el residuo 8 de la histona H3 (160,161). Esta proteína promueve la proliferación y supervivencia de células cancerosas, células madre y linfocitos (162–164), y tiene un papel importante en la diferenciación y migración de células leucémicas en la leucemia aguda mieloide (Acute Myeloid Leukemia, AML) por lo que podría ser una diana terapéutica para este tipo de cáncer (166); podría ser relevante que su delección en linfocitos T impide el desarrollo de los timocitos a células T, reduciéndose su número incluso en ratones heterocigotos (165). AJUBA funciona como una proteína andamiaje para el ensamblaje de proteínas involucradas en funciones fisiológicas como la adhesión celular, mitosis, meiosis, supervivencia, expresión génica o procesamiento de microRNA (167). Esta información permite especular sobre una posible función de *Eα* en la regulación de la proliferación y la supervivencia de los linfocitos T, y otras células, mediante la regulación de la expresión de estos genes del 3' sub-TAD (92).

Como resumen de las conclusiones de esta tesis, los datos mostrados demuestran que 1) *Eα* es fundamental para la transcripción del gen *Tcra* reordenado y la expresión del TCR $\alpha\beta$  en linfocitos T humanos y murinos, 2) *Eα* y *Cα* peak participan conjuntamente en esta regulación mediante el establecimiento de interacciones físicas con regiones del 5' sub-TAD y del

3' sub-TAD, y 3) nuestros análisis de 4C-ker indican un fortalecimiento de dichas interacciones con el 3' sub-TAD en la transición de timocitos DP a linfocitos T  $\alpha\beta$ , junto con la transición de un estado génico no reordenado a otro reordenado. Estas conclusiones permiten establecer la hipótesis de que el fortalecimiento de estas interacciones podría constituir el mecanismo molecular que impide una bajada de la expresión de la cadena TCR $\alpha$ , a pesar de la fuerte disminución de actividad de E $\alpha$  en linfocitos T  $\alpha\beta$ . Futuros experimentos serán necesarios para poder separar las funciones de E $\alpha$ core y E $\alpha$ CBE mediante la generación de nuevos modelos animales, entender la regulación dependiente de E $\alpha$  de los genes del 3' sub-TAD y analizar las consecuencias de la regulación de estos genes por el *enhancer*.



# BIBLIOGRAFÍA

## BIBLIOGRAFÍA

1. Goldsby RA, Kindt TJ, Kuby J, Osborne BA. Kuby Immunology (2002) <https://www.ncbi.nlm.nih.gov/nlmcatalog/101275636>
2. Schatz DG, Swanson PC. V(D)J recombination: Mechanisms of initiation (2011) *Annu Rev Genet* 45(1):167–202. <https://pubmed.ncbi.nlm.nih.gov/21854230/>
3. Mora T, Walczak AM. Quantifying lymphocyte receptor diversity (2018) *Systems Immunology* 183–98. <https://www.taylorfrancis.com/chapters/edit/10.1201/9781315119847-11/quantifying-lymphocyte-receptor-diversity-thierry-mora-aleksandra-walczak>
4. Rodríguez-Caparrós A, Álvarez-Santiago J, Valle-Pastor MJ, Suñé C, López-Ros J, Hernández-Munain C. Regulation of T-cell receptor gene expression by three-dimensional locus conformation and enhancer function (2020) *Int J Mol Sci* 21(22):8478. <https://pubmed.ncbi.nlm.nih.gov/33187197/>
5. Rothenberg E V, Taghon T. Molecular Genetics of T cell development (2005) *Annu Rev Immunol* 23(1):601–49. <https://pubmed.ncbi.nlm.nih.gov/15771582/>
6. Rothenberg E V., Zhang J, Li L. Multilayered specification of the T-cell lineage fate (2010) *Immunol Rev* 238(1):150–68. <https://pubmed.ncbi.nlm.nih.gov/10889068/>
7. Rothenberg E V., Moore JE, Yui MA. Launching the T-cell-lineage developmental programme (2008) *Nat Rev Immunol* 8:9–21. Available from: <https://www.nature.com/articles/nri2232/>
8. Rothenberg E V. Stepwise specification of lymphocyte developmental lineages (2000) *Curr Opin Genet Dev* 10(4):370–9. <https://pubmed.ncbi.nlm.nih.gov/10889068/>
9. Kumar B V., Connors TJ, Farber DL. Human T cell development, localization, and function throughout life (2018) *Immunity* 48(2):

- 202–13. <https://pubmed.ncbi.nlm.nih.gov/29466753/>
10. Taghon T, Yui MA, Pant R, Diamond RA, Rothenberg E V. Developmental and molecular characterization of emerging  $\beta$ - and  $\gamma\delta$ -selected pre-T cells in the adult mouse thymus (2006) *Immunity* 24(1):53–64. <https://pubmed.ncbi.nlm.nih.gov/16413923/>
  11. Hernández-Munain C, Casal Á, Juanes B, López-Ros J, Rodríguez-Caparrós A. Insights into the transcriptional regulation of the unrearranged and rearranged *Tcra* and *Tcrd* genes (2016) *J Clin Cell Immunol*. 07(02). doi: 10.4172/2155-9899.1000415
  12. Rodríguez-Caparrós A, García V, Casal Á, López-Ros J, García-Mariscal A, Tani-ichi S, Ikuta K., Hernández-Munain C. Notch signaling controls transcription via the recruitment of RUNX1 and MYB to enhancers during T cell development (2019) *J Immunol* 202(8):2460–72. <https://pubmed.ncbi.nlm.nih.gov/30877169/>
  13. Bentley GA, Mariuzza RA. The structure of the T cell antigen receptor (1996) *Annu Rev Immunol* 14:563–90. <https://pubmed.ncbi.nlm.nih.gov/8717525/>
  14. Clevers H, Alarcon B, Willeman T, Terhost C. The T cell receptor/CD3 complex: A dynamic protein ensemble (1988) *Annu Rev Immunol* 6:629–62. <https://pubmed.ncbi.nlm.nih.gov/3289580/>
  15. Li L-P, Lampert JC, Chen X, Leitao C, Popović J, Müller W, Blackestein T. Transgenic mice with a diverse human T cell antigen receptor repertoire (2010) *Nat Med* 16(9):1029–34. <https://pubmed.ncbi.nlm.nih.gov/20693993/>
  16. Singhal A, Mori L, De Libero G. T cell recognition of non-peptidic antigens in infectious diseases (2013) *Indian J Med Res* 138(5):620–31. <https://pubmed.ncbi.nlm.nih.gov/24434317/>
  17. Benveniste PM, Roy S, Nakatsugawa M, Chen ELY, Nguyen L, Millar DG, Ohashi PS, Hirano H, Adams EJ, Zúñiga-Pflücker JC. Generation and molecular recognition of melanoma-associated antigen-specific human T cells (2018) *Sci Immunol* 3(30):eaav4036.

<https://pubmed.ncbi.nlm.nih.gov/30552102/>

18. von Boehmer H, Teh HS, Kisielow P. The thymus selects the useful, neglects the useless and destroys the harmful (1989) *Immunol Today* 10(2):57–61. <https://pubmed.ncbi.nlm.nih.gov/2526642/>
19. Wucherpfennig KW, Gagnon E, Call MJ, Huseby ES, Call ME. Structural biology of the T-cell receptor: insights into receptor assembly, ligand recognition, and initiation of signaling (2010) *Cold Spring Harb Perspect Biol* 2 (4):a005140. <https://pubmed.ncbi.nlm.nih.gov/20452950/>
20. Thiel C, Otting N, Bontrop RE, Lanchbury JS. Generation and reactivation of T-cell receptor A joining region pseudogenes in primates (1995) *Immunogenetics* 43(1-2):57–62. <https://pubmed.ncbi.nlm.nih.gov/8537122/>
21. Glusman G, Rowen L, Lee I, Boysen C, Roach JC, Smit AFA, Wang K, Koop BF, Hood L. Comparative genomics of the human and mouse T cell receptor loci (2001) *Immunity* 15(3):337–49. <https://pubmed.ncbi.nlm.nih.gov/11567625/>
22. Sleckman BP, Bardon CG, Ferrini R, Davidson L, Alt FW. Function of the TCR $\alpha$  enhancer in  $\alpha\beta$  and  $\gamma\delta$  T cells (1997) *Immunity* 7(4):505–15. <https://pubmed.ncbi.nlm.nih.gov/9354471/>
23. Hernández-Munain C, Sleckman BP, Krangel MS. A developmental switch from TCR $\delta$  enhancer to TCR $\alpha$  enhancer function during thymocyte maturation (1999) *Immunity* 10(6):723–33. <https://pubmed.ncbi.nlm.nih.gov/10403647/>
24. Tani-ichi S, Satake M, Ikuta K. The pre-TCR signal induces transcriptional silencing of the TCR $\gamma$  locus by reducing the recruitment of STAT5 and Runx to transcriptional enhancers (2011) *Int Immunol* 23(9):553–63. <https://pubmed.ncbi.nlm.nih.gov/21750145/>
25. Hernández-Munain C. Recent insights into the transcriptional control of the *Tcra/Tcrd* locus by distant enhancers during the development of T-lymphocytes (2015) *Transcription* 6(4):65–73.

- <https://pubmed.ncbi.nlm.nih.gov/26230488/>
26. Monroe RJ, Sleckman BP, Monroe BC, Khor B, Claypool S, Ferrini R, Davidson L, Alt FW. Developmental regulation of TCR  $\delta$  locus accessibility and expression by the TCR  $\delta$  enhancer (1999) *Immunity* 10(5):503–13. <https://pubmed.ncbi.nlm.nih.gov/10367896/>
  27. Cobb RM, Oestreich KJ, Osipovich OA, Oltz EM. Accessibility Control of V(D)J Recombination (2006) *Adv Immunol* 91:45–109. <https://pubmed.ncbi.nlm.nih.gov/16938538/>
  28. Nishana M, Raghavan SC. Role of recombination activating genes in the generation of antigen receptor diversity and beyond (2012) *Immunology* 137(4):271–81. <https://pubmed.ncbi.nlm.nih.gov/23039142/>
  29. Van Gent DC, Ramsden DA, Gellert M. The RAG1 and RAG2 proteins establish the 12/23 rule in V(D)J recombination (1996) *Cell* 85(1):107–13. <https://pubmed.ncbi.nlm.nih.gov/8620529/>
  30. Tillman RE, Wooley AL, Hughes MM, Khor B, Sleckman BP. Regulation of T-cell receptor  $\beta$ -chain gene assembly by recombination signals: The beyond 12/23 restriction (2004) *Immunol Rev* 200:36–43. <https://pubmed.ncbi.nlm.nih.gov/15242394/>
  31. Drejer-Teel AH, Fugmann SD, Schatz DG. The beyond 12/23 rbestriction is imposed at the nicking and pairing steps of DNA cleavage during V(D)J recombination (2007) *Mol Cell Biol* 27(18):6288–99. <https://pubmed.ncbi.nlm.nih.gov/17636023/>
  32. Bettridge J, Na CH, Pandey A, Desiderio S. H3K4me3 induces allosteric conformational changes in the DNA-binding and catalytic regions of the V(D)J recombinase (2017) *Proc Natl Acad Sci U S A* 21;114(8):1904–9. <https://pubmed.ncbi.nlm.nih.gov/28174273/>
  33. Schlimgen RJ, Reddy KL, Singh H, Krangel MS. Initiation of allelic exclusion by stochastic interaction of *Tcrb* alleles with repressive nuclear compartments (2008) *Nat Immunol* 9(7):802–9. <https://pubmed.ncbi.nlm.nih.gov/18536719/>

34. Swanson PC. Fine Structure and activity of discrete RAG-HMG complexes on V(D)J recombination signals (2002) *Mol Cell Biol* 22(5):1340–51. <https://pubmed.ncbi.nlm.nih.gov/11839801/>
35. Thwaites DT, Carter C, Lawless D, Savic S, Boyes JM. A novel RAG1 mutation reveals a critical in vivo role for HMGB1/2 during V(D)J recombination (2019) *Blood* 133(8):820–9. <https://pubmed.ncbi.nlm.nih.gov/30538136/>
36. Cruz García A. Control de la procesividad de la resección por las proteínas CTIP, BRCA1 y las topoisomerasas del tipo II (2017). <https://digital.csic.es/handle/10261/166191>
37. Gellert M. V(D)J Recombination: RAG proteins, repair factors, and regulation (2002) *Annu Rev Biochem* 71(1):101–32. <https://pubmed.ncbi.nlm.nih.gov/12045092/>
38. Lescale C, Deriano L. The RAG recombinase: Beyond breaking (2017) *Mech Ageing Dev* 165:3–9. <https://pubmed.ncbi.nlm.nih.gov/27863852/>
39. Helmink BA, Sleckman BP. The response to and repair of RAG-mediated DNA double-strand breaks (2012) *Annu Rev Immunol* 30(1):175–202. <https://pubmed.ncbi.nlm.nih.gov/22224778/>
40. Schatz DG, Ji Y. Recombination centres and the orchestration of V(D)J recombination (2011) *Nat Rev Immunol* 11(4):251–63. <https://pubmed.ncbi.nlm.nih.gov/21394103/>
41. García-Mariscal A, del Blanco B, Hernández-Munain C. Generation of lymphocyte antigen receptor diversity: Validation of the “accessibility model” in controlling V(D)J recombination (2013) *Inmunología* 32(2):57-69. <https://www.elsevier.es/es-revista-inmunologia-322-articulo-generacion-diversidad-receptores-antigeno-linfocitos-S0213962612000856>
42. Yancopoulos GD, Alt FW. Developmentally controlled and tissue-specific expression of unrearranged VH gene segments (1985) *Cell* 40(2):271–81. <https://pubmed.ncbi.nlm.nih.gov/2578321/>

43. del Blanco B, Angulo Ú, Hernández-Munain C. Epigenetic Control of T-Cell Receptor Locus Rearrangements in Normal and Aberrant Conditions (2014) 295–329.  
<https://www.semanticscholar.org/paper/Epigenetic-Control-of-T-Cell-Receptor-Locus-in-and-Blanco-Angulo/ac93fdabb16620f304dc192e40426dfecc3acb3d>
44. Fondell JD, Marcu KB. Transcription of germ line V $\alpha$  segments correlates with ongoing T-cell receptor  $\alpha$ -chain rearrangement (1992) *Mol Cell Biol* 12(4):1480–9.  
<https://pubmed.ncbi.nlm.nih.gov/1312664/>
45. Goldman JP, Spencer DM, Raulet DH. Ordered rearrangement of variable region genes of the T cell receptor  $\gamma$  locus correlates with transcription of the unrearranged genes (1993) *J Exp Med* 177(3):729–39. <https://pubmed.ncbi.nlm.nih.gov/8382255/>
46. Schlissel MS, Corcoran LM, Baltimore D. Virus-transformed pre-B cells show ordered activation but not inactivation of immunoglobulin gene rearrangement and transcription (1991) *J Exp Med* 173(3):711–20.  
<https://pubmed.ncbi.nlm.nih.gov/1900081/>
47. Abarrategui I, Krangel MS. Regulation of T cell receptor- $\alpha$  gene recombination by transcription (2006) *Nat Immunol* 7(10):1109–15.  
<https://pubmed.ncbi.nlm.nih.gov/16936730/>
48. Hill L, Ebert A, Jaritz M, Wutz G, Nagasaka K, Tagoh H, Kostanova-Poliakova D, Schindler K, Sun Q, Bönelt P, Fischer M, Peters JM, Busslinger M. Wapl repression by Pax5 promotes V gene recombination by *Igh* loop extrusion (2020) *Nature* 584(7819):142–7.  
<https://pubmed.ncbi.nlm.nih.gov/32612238/>
49. Wilson A, Held W, MacDonald HR. Two waves of recombinase gene expression in developing thymocytes (1994) *J Exp Med* 179(4):1355–60. <https://pubmed.ncbi.nlm.nih.gov/8145048/>
50. Lin WC, Desiderio S. Cell cycle regulation of V (D) J recombination-activating protein RAG-2 (1994) *Proc Natl Acad Sci U S A* 91:2733–7.

<https://pubmed.ncbi.nlm.nih.gov/8146183/>

51. Irving BA, Alt FW, Killeen N. Thymocyte development in the absence of pre-T cell receptor extracellular immunoglobulin domains (1998) *Science* 280(5365):905–8.  
<https://pubmed.ncbi.nlm.nih.gov/9572735/>
52. Ferrero I, Mancini SJC, Grosjean F, Wilson A, Otten L, MacDonald HR. TCR $\gamma$  silencing during  $\alpha\beta$  T cell development depends upon pre-TCR-induced proliferation (2006) *J Immunol* 177(9):6038–43.  
<https://pubmed.ncbi.nlm.nih.gov/17056529/>
53. Erman B, Feigenbaum L, Coligan JE, Singer A. Early TCR $\alpha$  expression generates TCR $\alpha\gamma$  complexes that signal the DN-to-DP transition and impair development (2002) *Nat Immunol* 3(6):564–9.  
<https://pubmed.ncbi.nlm.nih.gov/12021779/>
54. Haks MC, Lefebvre JM, Lauritsen JPH, Carleton M, Rhodes M, Miyazaki T, Kappes DJ, Wiest DL. Attenuation of  $\gamma\delta$ TCR signaling efficiently diverts thymocytes to the  $\alpha\beta$  lineage (2005) *Immunity* 22(5):595–606. <https://pubmed.ncbi.nlm.nih.gov/15894277/>
55. Hayes SM, Li LQ, Love PE. TCR signal strength influences  $\alpha\beta/\gamma\delta$  lineage fate (2005) *Immunity* 22(5):583–93.  
<https://pubmed.ncbi.nlm.nih.gov/15894276/>
56. Joachims ML, Chain JL, Hooker SW, Knott-Craig CJ, Thompson LF. Human  $\alpha\beta$  and  $\gamma\delta$  thymocyte development: TCR gene rearrangements, intracellular TCR $\beta$  expression, and  $\gamma\delta$  developmental potential—differences between men and mice (2006) *J Immunol* 176(3):1543–52.  
<https://pubmed.ncbi.nlm.nih.gov/16424183/>
57. Huang CY, Sleckman BP, Kanagawa O. Revision of T cell receptor  $\alpha$  chain genes is required for normal T lymphocyte development (2005) *Proc Natl Acad Sci U S A* 102(40):14356–61.  
<https://pubmed.ncbi.nlm.nih.gov/16186502/>
58. Carico ZM, Roy Choudhury K, Zhang B, Zhuang Y, Krangel MS. *Tcrd* rearrangement redirects a processive *Tcra* recombination program to

- expand the *Tcra* Repertoire (2017) *Cell Rep* 19(10):2157–73.  
<https://pubmed.ncbi.nlm.nih.gov/28591585/>
59. Hawwari A, Bock C, Krangel MS. Regulation of T cell receptor  $\alpha$  gene assembly by a complex hierarchy of germline *J $\alpha$*  promoters (2005) *Nat Immunol* 6(5):481–9. <https://pubmed.ncbi.nlm.nih.gov/15806105/>
  60. del Blanco B, García-Mariscal A, Wiest DL, Hernández-Munain C. *Tcra* enhancer activation by inducible transcription factors downstream of pre-TCR signaling (2012) *J Immunol* 188(7):3278–93.  
<https://pubmed.ncbi.nlm.nih.gov/22357628/>
  61. Dauphars DJ, Mihai A, Wang L, Zhuang Y, Krangel MS. *Trav15-dv6* family *Tcrd* rearrangements diversify the *Tcra* repertoire (2022) *J Exp Med* 219(2)e20211581. <https://pubmed.ncbi.nlm.nih.gov/34910107/>
  62. Boudil A, Matei IR, Shih HY, Bogdanoski G, Yuan JS, Chang SG, Montpellier B, Kowalski PE, Voisin V, Bashir S, Bader GD, Krangel MS, Guidos CJ. IL-7 coordinates proliferation, differentiation and *Tcra* recombination during thymocyte  $\beta$ -selection (2015) *Nat Immunol* 16(4):397–405. <https://pubmed.ncbi.nlm.nih.gov/25729925/>
  63. Taghon T, Waegemans E, Van de Walle I. Notch Signaling During Human T cell Development (2012) *Curr Top Microbiol Immunol* 360:75-97. <https://pubmed.ncbi.nlm.nih.gov/22692833/>
  64. Reizis B, Leder P. Direct induction of T lymphocyte-specific gene expression by the mammalian notch signaling pathway (2002) *Genes Dev* 16(3):295–300. <https://pubmed.ncbi.nlm.nih.gov/11825871/>
  65. Rodríguez-Caparrós A, Tani-ichi S, Casal Á, López-Ros J, Suñé C, Ikuta K, Hernández-Munain C. Interleukin-7 receptor signaling is crucial for enhancer-dependent TCR $\delta$  germline transcription mediated through STAT5 recruitment (2022) *Front Immunol* 13:943510.  
<https://pubmed.ncbi.nlm.nih.gov/36059467/>
  66. Kang J, Coles M, Raulet DH. Defective development of  $\gamma/\delta$  T cells in interleukin 7 receptor- deficient mice is due to impaired expression of T cell receptor  $\gamma$  genes (1999) *J Exp Med* 190(7):973–82.

- <https://pubmed.ncbi.nlm.nih.gov/10510087/>
67. Yashiro-Ohtani Y, He Y, Ohtani T, Jones ME, Shestova O, Xu L, Fang TC, Chiang MY, Intlekofer AM, Blacklow SC, Zhuang Y, Pear WS. Pre-TCR signaling inactivates Notch1 transcription by antagonizing E2A (2009) *Genes Dev* 23(14):1665–76.  
<https://pubmed.ncbi.nlm.nih.gov/19605688/>
68. González-García S, García-Peydró M, Martín-Gayo E, Ballestar E, Esteller M, Bornstein R, de la Pompa JL, Ferrando AA, Toribio ML. CSL-MAML-dependent Notch1 signaling controls T lineage-specific IL-7R $\alpha$  gene expression in early human thymopoiesis and leukemia (2009) *J Exp Med* 206(4):779–791.  
<https://pubmed.ncbi.nlm.nih.gov/19349467/>
69. Wang H, Zang C, Taing L, Arnett KL, Wong YJ, Pear WS, Blacklow SC, Liu S, Aster JC. NOTCH1-RBPJ complexes drive target gene expression through dynamic interactions with superenhancers (2014) *Proc Natl Acad Sci U S A* 111(2):705–10.  
<https://pubmed.ncbi.nlm.nih.gov/24374627/>
70. Olins DE, Olins AL. Chromatin history: Our view from the bridge (2003) *Nat Rev Mol Cell Biol* 4(10):809–14.  
<https://pubmed.ncbi.nlm.nih.gov/14570061/>
71. Wolffe AP, Guschin D. Review: chromatin structural features and targets that regulate transcription (2000) *J Struct Biol* 129(2–3):102–22. <https://pubmed.ncbi.nlm.nih.gov/10806063/>
72. Tyagi M, Imam N, Verma K, Patel AK. Chromatin remodelers: We are the drivers!! (2016) *Nucleus* 7(4):388–404.  
<https://pubmed.ncbi.nlm.nih.gov/27429206/>
73. Jacobs SA, Fischle W, Khorasanizadeh S. Assays for the determination of structure and dynamics of the interaction of the chromodomain with histone peptides (2004) *Methods Enzymol* 376:131–48.  
<https://pubmed.ncbi.nlm.nih.gov/14975303/>
74. Zhang P, Torres K, Liu X, Liu C, E. Pollock R. An overview of

- chromatin-regulating proteins in cells (2016) *Curr Protein Pept Sci* 17(5):401–10. <https://pubmed.ncbi.nlm.nih.gov/26796306/>
75. Allfrey VG, Faulkner R, Mirsky AE. Acetylation and methylation of histones and their possible role in the regulation of DNA synthesis (1964) *Proc Natl Acad Sci U S A* 51(5):786–94. <https://pubmed.ncbi.nlm.nih.gov/14172992/>
76. Dion MF, Altschuler SJ, Wu LF, Rando OJ. Genomic characterization reveals a simple histone H4 acetylation code (2005) *Proc Natl Acad Sci U S A* 102(15):5501–6. <https://pubmed.ncbi.nlm.nih.gov/15795371/>
77. Zentner GE, Henikoff S. Regulation of nucleosome dynamics by histone modifications (2013) *Nat Struct Mol Biol* 20(3):259–66. <https://pubmed.ncbi.nlm.nih.gov/23463310/>
78. McGrath J, Trojer P. Targeting histone lysine methylation in cancer (2015) *Pharmacol Ther* 150: 1–22. <https://pubmed.ncbi.nlm.nih.gov/25578037/>
79. Shimazaki N, Lieber MR. Histone methylation and V(D)J recombination (2014) *Int J Hematol* 100(3): 230–7. <https://pubmed.ncbi.nlm.nih.gov/25060705/>
80. McMurry MT, Krangel MS. A Role for histone acetylation in the developmental regulation of V(D)J recombination (2000) *Science* 287(5452):495–8. <https://pubmed.ncbi.nlm.nih.gov/10642553/>
81. Shimazaki N, Tsai AG, Lieber MR. H3K4me3 stimulates the V(D)J RAG complex for both nicking and hairpinning in trans in addition to tethering in cis: implications for translocations (2009) *Mol Cell* 34(5):535–44. <https://pubmed.ncbi.nlm.nih.gov/19524534/>
82. Ramón-Maiques S, Kuo AJ, Carney D, Matthews AGW, Oettinger MA, Gozani O, Yang W. The plant homeodomain finger of RAG2 recognizes histone H3 methylated at both lysine-4 and arginine-2 (2007) *Proc Natl Acad Sci U S A* 104(48):18993–8. <https://pubmed.ncbi.nlm.nih.gov/18025461/>
83. Mota-Gómez I, Lupiáñez DG. A (3D-Nuclear) Space Odyssey: Making

- sense of Hi-C maps (2019) *Genes (Basel)*: 10(6):415.  
<https://www.mdpi.com/2073-4425/10/6/415/htm>
84. Canela A, Maman Y, Jung S, Wong N, Callen E, Day A, Kieffer-Kwon K-R, Pekowska A, Zhang H, Rao SSP, Huang S-C, Mckinnon PJ, Aplan PD, Pommier Y, Aiden EL, Casellas R, Nussenzweig A. Genome organization drives chromosome fragility (2017) *Cell* 170(3):507-521.e18. <https://pubmed.ncbi.nlm.nih.gov/28735753/>
  85. Lupiáñez DG, Kraft K, Heinrich V, Krawitz P, Brancati F, Klopocki E, Horn D, Kayserili H, Opitz JN, Laxova R, Santos-Simarro F, Gilbert-Dussadier B, Wittler L, Borschiwer M, Hass SA, Osterwalder M, Franke M, Timmermann B, Hecht J, Spielman M, Visel A, Mundlos S. Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions (2015) *Cell* 161(5):1012–25. <https://pubmed.ncbi.nlm.nih.gov/25959774/>
  86. Shih H-Y, Krangel MS. Chromatin architecture, CCCTC-binding factor, and V(D)J recombination: managing long-distance relationships at antigen receptor loci (2013) *J Immunol* 190(10):4915–21. <https://pubmed.ncbi.nlm.nih.gov/23645930/>
  87. Abarrategui I, Krangel MS. Noncoding transcription controls downstream promoters to regulate T-cell receptor  $\alpha$  recombination (2007) *EMBO J* 26(20):4380–90. <https://pubmed.ncbi.nlm.nih.gov/17882258/>
  88. Hawwari A, Krangel MS. Regulation of TCR  $\delta$  and  $\alpha$  repertoires by local and long-distance control of variable gene segment chromatin structure (2005) *J Exp Med* 202(4):467–72. <https://pubmed.ncbi.nlm.nih.gov/16087716/>
  89. Shih H-Y, Verma-Gaur J, Torkamani A, Feeney AJ, Galjart N, Krangel MS. *Tcra* gene recombination is supported by a *Tcra* enhancer-and CTCF-dependent chromatin hub (2012) *Proc Natl Acad Sci U S A* 109(50):E3493-502. <https://pubmed.ncbi.nlm.nih.gov/23169622/>
  90. Ji Y, Little AJ, Banerjee JK, Hao B, Oltz EM, Krangel MS, Schatz DG.

- Promoters, enhancers, and transcription target RAG1 binding during V(D)J recombination (2010) *J Exp Med* 207(13):2809–16.  
<https://pubmed.ncbi.nlm.nih.gov/21115692/>
91. Hawwari A, Krangel MS. Role for rearranged variable gene segments in directing secondary T cell receptor  $\alpha$  recombination (2007) *Proc Natl Acad Sci U S A* 104(3):903–7.  
<https://pubmed.ncbi.nlm.nih.gov/17210914/>
  92. Shih H-Y, Krangel MS. Distinct contracted conformations of the *Tcra/Tcrd* locus during *Tcra* and *Tcrd* recombination (2010) *J Exp Med* 207(9):1835–41. <https://pubmed.ncbi.nlm.nih.gov/20696701/>
  93. Chen L, Carico Z, Shih HY, Krangel MS. A discrete chromatin loop in the mouse *Tcra-Tcrd* locus shapes the TCR $\delta$  and TCR $\alpha$  repertoires (2015) *Nat Immunol* 16(10):1085–93.  
<https://pubmed.ncbi.nlm.nih.gov/26258942/>
  94. del Blanco B, Angulo Ú, Krangel MS, Hernández-Munain C. T-cell receptor  $\alpha$  enhancer is inactivated in  $\alpha\beta$  T lymphocytes (2015) *Proc Natl Acad Sci U S A* 112(14):E1744–53.  
<https://pubmed.ncbi.nlm.nih.gov/25831496/>
  95. Zhang Y, Zhang X, Ba Z, Liang Z, Dring EW, Hu H, Lou J, Kyritsis N, Zurita J, Shamim MS, Aiden AP, Aiden EL, Alt FW. The fundamental role of chromatin loop extrusion in physiological V(D)J recombination (2019) *Nature* 573(7775):600–4.  
<https://pubmed.ncbi.nlm.nih.gov/31511698/>
  96. Jain S, Ba Z, Zhang Y, Dai HQ, Alt FW. CTCF-binding elements mediate accessibility of RAG substrates during chromatin scanning (2018) *Cell* 174(1):102–116.e14.  
<https://pubmed.ncbi.nlm.nih.gov/29804837/>
  97. Loguercio S, Barajas-Mora EM, Shih H-Y, Krangel MS, Feeney AJ. Variable extent of lineage-specificity and developmental stage-specificity of cohesin and CCCTC-binding factor binding within the immunoglobulin and T cell receptor loci (2018) *Front Immunol*. 2018

- 9:425. <https://pubmed.ncbi.nlm.nih.gov/29593713/>
98. Carico Z, Krangel MS. Chromatin dynamics and the development of the TCR $\alpha$  and TCR $\delta$  repertoires (2015) *Adv in Immunol* 128:307–61. <https://pubmed.ncbi.nlm.nih.gov/26477370/>
  99. Zhao H, Li Z, Zhu Y, Bian S, Zhang Y, Qin L, Naik AK, He J, Zhang Z, Krangel MS, Hao B. A role of the CTCF binding site at enhancer Ea in the dynamic chromatin organization of the *Tcra–Tcrd* locus (2020) *Nucleic Acids Res* 48(17):9621–36. <https://pubmed.ncbi.nlm.nih.gov/32853367/>
  100. Rodríguez-Caparrós A, Álvarez-Santiago J, López-Castellanos L, Ruiz-Rodríguez C, Valle-Pastor MJ, López-Ros J, Angulo Ú, Andrés-León E, Suñé C, Hernández-Munain C. Differently regulated gene-specific activity of enhancers located at the boundary of subtopologically associated domains: TCR $\alpha$  Enhancer (2022) *J Immunol* 208(4):910–28. <https://pubmed.ncbi.nlm.nih.gov/35082160/>
  101. Maston GA, Evans SK, Green MR. Transcriptional regulatory elements in the human genome (2006) *Annu Rev Genomics Hum Genet* 7(1):29–59. <https://pubmed.ncbi.nlm.nih.gov/16719718/>
  102. Moreau P, Hen R, Waslyk B, Everett R, Gaub MP, Chambon P. The SV40 72 base repair repeat has a striking effect on gene expression both in SV40 and other chimeric recombinants (1981) *Nucleic Acids Res* 9(22):6047–68. <https://pubmed.ncbi.nlm.nih.gov/6273820/>
  103. Mercola M, Wang XF, Olsen J, Calame K. Transcriptional enhancer elements in the mouse immunoglobulin heavy chain locus (1983) *Science* 221(4611):663–5. <https://pubmed.ncbi.nlm.nih.gov/6306772/>
  104. Heinz S, Romanoski CE, Benner C, Glass CK. The selection and function of cell type-specific enhancers (2015) *Nat Rev Mol Cell Biol* 16(3): 144–54. <https://pubmed.ncbi.nlm.nih.gov/25650801/>
  105. Shlyueva D, Stampfel G, Stark A. Transcriptional enhancers: from properties to genome-wide predictions (2014) *Nat Rev Genet* 15(4):272–86. <https://pubmed.ncbi.nlm.nih.gov/24614317/>

106. Lewis MW, Li S, Franco HL. Transcriptional control by enhancers and enhancer RNAs (2019) *Transcription* 10(4-5):171–86.  
<https://pubmed.ncbi.nlm.nih.gov/31791217/>
107. Heintzman ND, Hon GC, Hawkins RD, Kheradpour P, Stark A, Harp LF, Ye Z, Lee LK, Stuart RK, Ching CW, Ching KA, Antosiewicz-Bourget JE, Liu H, Zhang X, Green RD, Lobanenkov VV, Stewart R, Thompson JA, Crawford GE, Kellis M, Ren B. Histone modifications at human enhancers reflect global cell-type-specific gene expression (2009) *Nature* 459(7243):108–12.  
<https://pubmed.ncbi.nlm.nih.gov/19295514/>
108. Creyghton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, Hanna J, Lodato MA, Frampton GM, Sharp PA, Boyer LA, Young RA, Jeanisch R. Histone H3K27ac separates active from poised enhancers and predicts developmental state (2010) *Proc Natl Acad Sci U S A* 107(50):21931–6. <https://pubmed.ncbi.nlm.nih.gov/21106759/>
109. Pekowska A, Benoukraf T, Zacarias-Cabeza J, Belhocine M, Koch F, Holota H, Imbert J, Andrau J-C, Ferrier P, Spicuglia S. H3K4 trimethylation provides an epigenetic signature of active enhancers (2011) *EMBO J* 30(20):4198–210.  
<https://pubmed.ncbi.nlm.nih.gov/21847099/>
110. Rosenfeld MG, Lunyak V V., Glass CK. Sensors and signals: A coactivator/corepressor/epigenetic code for integrating signal-dependent programs of transcriptional response (2006) *Genes Dev* 20(11):1405–28. <https://pubmed.ncbi.nlm.nih.gov/16751179/>
111. Melamed P, Yosefzon Y, Rudnizky S, Pnueli L. Transcriptional enhancers: Transcription, function and flexibility (2016) *Transcription* 7(1):26–31. <https://pubmed.ncbi.nlm.nih.gov/26934309/>
112. Osipovich O, Oltz EM. Regulation of antigen receptor gene assembly by genetic-epigenetic crosstalk (2010) *Semin Immunol* 22(6):313–22.  
<https://pubmed.ncbi.nlm.nih.gov/20829065/>
113. Haberle V, Stark A. Eukaryotic core promoters and the functional basis

- of transcription initiation (2018) *Nat Rev Mol Cell Biol* 19(10):621–37.  
<https://pubmed.ncbi.nlm.nih.gov/29946135/>
114. Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, Hawkins RD, Barrera LO, Calcar SV, Qu C, Ching KA, Wang W, Weng Z, Green RD, Crawford GE, Ren B. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome (2007) *Nat Genet* 39(3):311–8. <https://pubmed.ncbi.nlm.nih.gov/17277777/>
115. Whitehurst CE, Chattopadhyay S, Chen J. Control of V(D)J recombinational accessibility of the D $\beta$ 1 gene segment at the TCR  $\beta$  locus by a germline promoter (1999) *Immunity* 10(3):313–22. <https://pubmed.ncbi.nlm.nih.gov/10204487/>
116. Oestreich KJ, Cobb RM, Pierce S, Chen J, Ferrier P, Oltz EM. Regulation of TCR $\beta$  gene assembly by a promoter/enhancer holocomplex (2006) *Immunity* 24(4):381–91. <https://pubmed.ncbi.nlm.nih.gov/16618597/>
117. Ryu CJ, Haines BB, Lee HR, Kang YH, Draganov DD, Lee M, Whitehurst CE, Hong HJ, Chen J. The T-cell receptor  $\beta$  variable gene promoter is required for efficient V $\beta$  rearrangement but not allelic exclusion (2004) *Mol Cell Biol* 24(16):7015–23. <https://pubmed.ncbi.nlm.nih.gov/15282302/>
118. Villey I, Caillol D, Selz F, Ferrier P, De Villartay JP. Defect in rearrangement of the most 5' TCR-J $\alpha$  following targeted deletion of T early  $\alpha$  (TEA): Implications for TCR  $\alpha$  locus accessibility (1996) *Immunity* 5(4):331–42. <https://pubmed.ncbi.nlm.nih.gov/8885866/>
119. Danino YM, Even D, Ideses D, Juven-Gershon T. The core promoter: At the heart of gene expression (2015) *Biochim Biophys Acta* 1849(8):1116–31. <https://pubmed.ncbi.nlm.nih.gov/25934543/>
120. Wang Q, Carroll JS, Brown M. Spatial and temporal recruitment of androgen receptor and its coactivators involves chromosomal looping and polymerase tracking (2005) *Mol Cell* 19(5):631–42. <https://pubmed.ncbi.nlm.nih.gov/16137620/>

121. Mora A, Sandve GK, Gabrielsen OS, Eskeland R. In the loop: promoter–enhancer interactions and bioinformatics (2015) *Brief Bioinform* 17(6):980-95. <https://pubmed.ncbi.nlm.nih.gov/26586731/>
122. Bulger M, Groudine M. Looping versus linking: toward a model for long-distance gene activation (1999) *Genes Dev* 13(19):2465–77. <https://pubmed.ncbi.nlm.nih.gov/10521391/>
123. Boija A, Klein IA, Sabari BR, Dall’Agnese A, Coffey EL, Zamudio A V, Li CH, Shrinivas K, Manteiga JC, Hannett NM, Abraham BJ, Afeyan LK, Guo YE, Rimel JK, Fant CB, Schuijers J, Lee TI, Taajes DJ, Young RA. Transcription factors activate genes through the phase-separation capacity of their activation domains (2018) *Cell* 175(7):1842-1855.e16. <https://pubmed.ncbi.nlm.nih.gov/30449618/>
124. Shrinivas K, Sabari BR, Coffey EL, Klein IA, Boija A, Zamudio A V, Schuijers J, Hannett NM, Sharp PA, Young RA, Chakraborty AK. Enhancer features that drive formation of transcriptional condensates (2019) *Mol Cell* 75(3):549-561.e7. <https://pubmed.ncbi.nlm.nih.gov/31398323/>
125. Diaz P, Cado D, Winoto A. A locus control region in the T cell receptor  $\alpha/\delta$  locus (1994) *Immunity* 1(3):207–17. <https://pubmed.ncbi.nlm.nih.gov/7889409/>
126. Harrow F, Ortiz BD. The TCR $\alpha$  locus control region specifies thymic, but not peripheral, patterns of TCR $\alpha$  gene expression (2005) *J Immunol* 175(10):6659–67. <https://pubmed.ncbi.nlm.nih.gov/16272321/>
127. Ortiz BD, Cado D, Winoto A. A new element within the T-cell receptor  $\alpha$  locus required for tissue-specific locus control region activity (1999) *Mol Cell Biol* 19(3):1901–9. <https://pubmed.ncbi.nlm.nih.gov/10022877/>
128. Gomos-Klein J, Harrow F, Alarcón J, Ortiz BD. CTCF-independent, but not CTCF-dependent, elements significantly contribute to TCR- $\alpha$  locus control region activity (2007) *J Immunol* 179(2):1088–95.

- <https://pubmed.ncbi.nlm.nih.gov/17617601/>
129. Arden B, Clark SP, Kabelitz D, Mak TW. Human T-cell receptor variable gene segment families (1995) *Immunogenetics* 42(6):455–500. <https://pubmed.ncbi.nlm.nih.gov/8550092/>
  130. Aifantis I, Bassing CH, Garbe AI, Sawai K, Alt FW, Von Boehmer H. The E $\delta$  enhancer controls the generation of CD4<sup>+</sup>CD8<sup>-</sup>  $\alpha\beta$ TCR-expressing T cells that can give rise to different lineages of  $\alpha\beta$  T cells (2006) *J Exp Med* 203(6):1543–50. <https://pubmed.ncbi.nlm.nih.gov/16754716/>
  131. Balmelle N, Zamarreño N, Krangel MS, Hernández-Munain C. Developmental activation of the TCR  $\alpha$  enhancer requires functional collaboration among proteins bound inside and outside the core enhancer (2004) *J Immunol* 173(8):5054–63. <https://pubmed.ncbi.nlm.nih.gov/15470049/>
  132. Ho IC, Yang LH, Morle G, Leiden JM. A T-cell-specific transcriptional enhancer element 3' of C( $\alpha$ ) in the human T-cell receptor  $\alpha$  locus (1989) *Proc Natl Acad Sci U S A* 86(17):6714–8. <https://pubmed.ncbi.nlm.nih.gov/2788889/>
  133. Roberts JL, Lauzurica P, Krangel MS. Developmental regulation of VDJ recombination by the core fragment of the T cell receptor  $\alpha$  enhancer (1997) *J Exp Med* 185(1):131–40. <https://pubmed.ncbi.nlm.nih.gov/8996249/>
  134. Bassing CH, Tillman RE, Woodman BB, Canty D, Monroe RJ, Sleckman BP, Alt FW. T cell receptor (TCR)  $\alpha/\delta$  locus enhancer identity and position are critical for the assembly of TCR  $\delta$  and  $\alpha$  variable region genes (2003) *Proc Natl Acad Sci U S A* 100(5):2598–603. <https://pubmed.ncbi.nlm.nih.gov/12604775/>
  135. Hernández-Munain C, Roberts JL, Krangel MS. Cooperation among multiple transcription factors is required for access to minimal T-cell receptor  $\alpha$ -enhancer chromatin in vivo (1998) *Mol Cell Biol* 18(6):3223–33. <https://pubmed.ncbi.nlm.nih.gov/9584163/>

136. del Blanco B, Roberts JL, Zamarreño N, Balmelle-Devaux N, Hernández-Munain C. Flexible stereospecific interactions and composition within nucleoprotein complexes assembled on the TCR $\alpha$  gene enhancer (2009) *J Immunol* 183(3):1871–83.  
<https://pubmed.ncbi.nlm.nih.gov/19596981/>
137. Collins B, Clambey ET, Scott-Browne J, White J, Marrack P, Hagman J, Kappler JW. Ikaros promotes rearrangement of TCR  $\alpha$  genes in an Ikaros null thymoma cell line (2013) *Eur J Immunol* 43(2):521–32.  
<https://pubmed.ncbi.nlm.nih.gov/23172374/>
138. Cieslak A, Charbonnier G, Tesio M, Mathieu E-L, Belhocine M, Touzart A, Smith C, Hypolite G, Andrieu GP, Martens JHA, Janssen-Megens E, Gut M, Gut I, Boissel N, Petit A, Puthier D, Macintyre E, Stunnenberg HG, Spicuglia S, Asnafi V. Blueprint of human thymopoiesis reveals molecular mechanisms of stage-specific TCR enhancer activation (2020) *J Exp Med* 217(9):e20192360.  
<https://pubmed.ncbi.nlm.nih.gov/32667968/>
139. Shinkai Y, Rathbun oGary, Lam KP, Oltz EM, Stewart V, Mendelsohn M, Charron J, Datta M, Young F, Stall AM, Alt FW. RAG-2-deficient mice lack mature lymphocytes owing to inability to initiate V(D)J rearrangement (1992) *Cell* 68(5):855–67.  
<https://pubmed.ncbi.nlm.nih.gov/1547487/>
140. Shinkai Y, Koyasu S, Nakayama K, Murphy KM, Loh DY, Reinherz EL, Alt FW. Restoration of T cell development in RAG-2-deficient mice by functional TCR transgenes (1993) *Science* 259(5096):822–5.  
<https://pubmed.ncbi.nlm.nih.gov/8430336/>
141. Jackson A, Kondilis HD, Khor B, Sleckman BP, Krangel MS. Regulation of T cell receptor  $\beta$  allelic exclusion at a level beyond accessibility (2005) *Nat Immunol* 6(2):189–97.  
<https://pubmed.ncbi.nlm.nih.gov/15640803/>
142. Schneider U, Schwenk HU, Bornkamm G. Characterization of EBV-genome negative “null” and “T” cell lines derived from children with acute lymphoblastic leukemia and leukemic transformed non-Hodgkin

- lymphoma (1977) *Int J Cancer* 19(5):621–6.  
<https://pubmed.ncbi.nlm.nih.gov/68013/>
143. Clarke SRmK, Barnden M, Kurts C, Carbone FR, Miller JF, Heath WR. Characterization of the ovalbumin-specific TCR transgenic line OT-I: MHC elements for positive and negative selection (2000) *Immunol Cell Biol* 78(2):110–7. <https://pubmed.ncbi.nlm.nih.gov/10762410/>
144. Murphy KM, Heimberger AB, Loh DY. Induction by antigen of intrathymic apoptosis of CD4<sup>+</sup>CD8<sup>+</sup>TCR<sup>lo</sup> thymocytes in vivo (1990) *Science* 250(4988):1720–3.  
<https://pubmed.ncbi.nlm.nih.gov/2125367/>
145. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F. Multiplex genome engineering using CRISPR/Cas systems (2013) *Science* 339(6121):819–23.  
<https://pubmed.ncbi.nlm.nih.gov/23287718/>
146. Krijger PHL, Geeven G, Bianchi V, Hilvering CRE, de Laat W. 4C-seq from beginning to end: A detailed protocol for sample preparation and data analysis (2020) *Methods* 170:17–32.  
<https://pubmed.ncbi.nlm.nih.gov/31351925/>
147. Andrews S. FastQC: a quality control tool for high throughput sequence data (2010)  
<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
148. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads (2011) *EMBnet.journal* 17(1):10–12.  
<https://journal.embnet.org/index.php/embnetjournal/article/view/200>
149. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2 (2012) *Nat Methods* 9(4):357–9.  
<https://pubmed.ncbi.nlm.nih.gov/22388286/>
150. Ohashi PS, Pircher H, Bürki K, Zinkernagel RM, Hengartner H. Distinct sequence of negative or positive selection implied by thymocyte T-cell receptor densities (1990) *Nature* 346(6287):861–3.

- <https://pubmed.ncbi.nlm.nih.gov/1975427/>
151. Bendelac A, Matzinger P, Seder RA, Paul WE, Schwartz RH. Activation events during thymic selection (1992) *J Exp Med* 175(3):731–42.  
<https://pubmed.ncbi.nlm.nih.gov/1740662/>
  152. Kearsse KP, Takahama Y, Punt JA, Sharrow SO, Singer A. Early molecular events induced by T cell receptor (TCR) signaling in immature CD4<sup>+</sup> CD8<sup>+</sup> thymocytes: Increased synthesis of TCR- $\alpha$  protein is an early response to tcr signaling that compensates for TCR- $\alpha$  instability, improves tcr assembly, and parallels other indicators of positive selection (1995) *J Exp Med* 181(1):193–202.  
<https://pubmed.ncbi.nlm.nih.gov/7528767/>
  153. Abraham RT, Weiss A. Jurkat T cells and development of the T-cell receptor signalling paradigm (2004) *Nat Rev Immunol* 4(4):301–8.  
<https://pubmed.ncbi.nlm.nih.gov/15057788/>
  154. Arnaud J, Erard M, Martin E, Llobera R, Gouaillard C, Constans J, Rubin B. Molecular modelling and endoplasmic reticulum retention of mutated TCR/CD3 complexes (2001) *Scand J Immunol* 53(6):540–52.  
<https://pubmed.ncbi.nlm.nih.gov/11422902/>
  155. Markegard E, Trager E, Yang C wen O, Zhang W, Weiss A, Roose JP. Basal LAT-diacylglycerol-rasGRP1 signals in T cells maintain TCR $\alpha$  gene expression (2011) *PLoS One* 6(9):e25540.  
<https://pubmed.ncbi.nlm.nih.gov/21966541/>
  156. Weiss A, Stobo JD. Requirement for the coexpression of T3 and the T cell antigen receptor on a malignant human T cell line (1984) *J Exp Med* 160(5):1284–99. <https://pubmed.ncbi.nlm.nih.gov/6208306/>
  157. Saito T, Weiss A, Gunter KC, Shevach EM, Germain RN. Cell surface T3 expression requires the presence of both  $\alpha$ - and  $\beta$ -chains of the T cell receptor (1987) *J Immunol* 139(2):625–8.  
<https://pubmed.ncbi.nlm.nih.gov/2955045/>
  158. Geisler C, Schøller J, Wahi MA, Rubin B, Weiss A. Association of the human CD3- $\zeta$  chain with the  $\alpha\beta$ -T cell receptor/CD3 complex. Clues

- from a T cell variant with a mutated T cell receptor- $\alpha$  chain (1990) *J Immunol* 145(6):1761–7. <https://pubmed.ncbi.nlm.nih.gov/2144007/>
159. Jóźwik A, Soroczyńska M, Witkowski JM, Bryl E. CD3 receptor modulation in Jurkat leukemic cell line (2004) *Folia Histochem Cytobiol* 42(1):41–3. <https://pubmed.ncbi.nlm.nih.gov/15046399/>
160. von Boehmer H. Developmental biology of T cells in T cell-receptor transgenic mice (1990) *Annu Rev Immunol* 8(1):531–56. <https://pubmed.ncbi.nlm.nih.gov/2188673/>
161. Haurogné K, Bach JM, Lieubeau B. Easy and rapid method of zygosity determination in transgenic mice by SYBR Green real-time quantitative PCR with a simple data analysis (2007) *Transgenic Res* 16(1):127–31. <https://pubmed.ncbi.nlm.nih.gov/17080302/>
162. Hogquist KA, Jameson SC, Heath WR, Howard JL, Bevan MJ, Carbone FR. T cell receptor antagonist peptides induce positive selection (1994) *Cell* 76(1):17–27. <https://pubmed.ncbi.nlm.nih.gov/8287475/>
163. Raviram R, Rocha PP, Müller CL, Miraldi ER, Badri S, Fu Y, Swanzei E, Proudhon C, Snetkova V, Bonneau R, Skok JA. 4C-ker: A Method to Reproducibly Identify Genome-Wide Interactions Captured by 4C-Seq Experiments (2016) *PLoS Comput Biol* 12(3):e1004780. <https://pubmed.ncbi.nlm.nih.gov/26938081/>
164. Mingueneau M, Kreslavsky T, Gray D, Heng T, Cruse R, Ericson J, Bendall S, Spitzer MH, Nolan GP, Kobayashi K, von Boehmer H, Mathis D, Benoist C, Immunological Genome Consortium. The transcriptional landscape of  $\alpha\beta$  T cell differentiation (2013) *Nat Immunol* 14(6):619–32. <https://pubmed.ncbi.nlm.nih.gov/23644507/>
165. Padovan E, Casorati G, Dellabona P, Meyer S, Brockhaus M, Lanzavecchia A. Expression of two T cell receptor  $\alpha$  chains: Dual receptor T cells (1993) *Science* 262(5132):422–4. <https://pubmed.ncbi.nlm.nih.gov/8211163/>
166. Wilkinson MF, MacLeod CL. Induction of T-cell receptor- $\alpha$  and - $\beta$  mRNA in SL12 cells can occur by transcriptional and post-

- transcriptional mechanisms (1988) *EMBO J* 7(1):101–9.  
<https://pubmed.ncbi.nlm.nih.gov/3129288/>
167. Sun X, Saito M, Sato Y, Chikata T, Naruto T, Ozawa T, Kpbayshi E, Kishi H, Muraguchi A, Takiguchi M. Unbiased analysis of TCR $\alpha$ / $\beta$  chains at the single-cell level in human CD8<sup>+</sup> T-cell subsets (2012) *PLoS One* 7(7):e40386. <https://pubmed.ncbi.nlm.nih.gov/22792299/>
168. Kovalovsky D, Pezzano M, Ortiz BD, Sant'Angelo DB. A novel TCR transgenic model reveals that negative selection involves an immediate, Bim-dependent pathway and a delayed, Bim-independent pathway (2010) *PLoS One* 5(1):e8675.  
<https://pubmed.ncbi.nlm.nih.gov/20072628/>
169. Kučerová-Levisohn M, Knirr S, Mejia RI, Ortiz BD. The 3'-Ja region of the TCR $\alpha$  locus bears gene regulatory activity in thymic and peripheral T cells (2015) *PLoS One* 10(7):e0132856–e0132856.  
<https://pubmed.ncbi.nlm.nih.gov/26177549/>
170. Luppino JM, Park DS, Nguyen SC, Lan Y, Xu Z, Yunker R, Joyce EF. Cohesin promotes stochastic domain intermingling to ensure proper regulation of boundary-proximal genes (2020) *Nat Genet* 52(8):840–8.  
<https://pubmed.ncbi.nlm.nih.gov/32572210/>
171. Spicuglia S, Payet D, Tripathi RK, Rameil P, Verthuy C, Imbert J, Ferrier P, Hempel WM. TCR $\alpha$  enhancer activation occurs via a conformational change of a pre-assembled nucleo-protein complex (2000) *EMBO J* 19(9):2034.  
<https://pubmed.ncbi.nlm.nih.gov/10790370/>

# PUBLICACIONES



1. Rodríguez-Caparrós A., Álvarez-Santiago J., Valle-Pastor M.J., Suñé C., López-Ros J., Hernández-Munain C. (2020) Regulation of T-cell receptor gene expression by three-dimensional locus conformation and enhancer function. *Int. J. Mol. Sci.* 21(22): 8478.  
Doi: 10.3390/ijms21228478  
PMID: 33187197
2. Rodríguez-Caparrós A\*, Álvarez-Santiago J\*, López-Castellanos L., Ruiz-Rodríguez C., Valle-Pastor M.J., López-Ros J., Angulo Ú., Andrés-León E., Suñé C., Hernández-Munain C. (2022) Differently regulated gene-specific activity of enhancers located at the boundary of subtopologically associated domains: TCR $\alpha$  enhancer. *J. Immunol.* 208 (4): 910-924.  
Doi: 10.4049/jimmunol.2000864  
PMID: 35082160

\*Estos investigadores han contribuido de la misma forma en la ejecución de este trabajo y, por tanto, ambos son primeros autores de la publicación.