



Universitat Autònoma de Barcelona

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

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

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

Study of the role of the histone demethylase JMJD3 as a transcriptional regulator

PhD Thesis

Raquel Fueyo Arévalo

Doctoral program in Biochemistry, Molecular Biology and Biomedicine.
Universitat Autònoma de Barcelona.



Investigation performed at the
Institute of Molecular Biology of Barcelona, CSIC.

Directed by:
Dr. Marian Martínez-Balbás

Tutored by:
Dr. Pere Suau León

Dr. Marian Martínez-Balbás

Raquel Fueyo Arévalo

Barcelona 2018

*Strangers on this road we are on,
we are not two we are one.*

The Kinks, "Strangers"

A mi compañero vital Alberto Gamazo,
por hacerme más libre.

Acknowledgements

Acknowledgements

“Though no man can draw a stroke between the confines of day and night, yet light and darkness are upon the whole tolerably distinguishable” (Edmund Burke).

Esta pomposa frase resume lo que estos años de tesis doctoral me han aportado como investigadora: una inconmensurable resistencia al fracaso y toneladas de pragmatismo. Con esto quiero comenzar a dar las gracias a Marian Martínez-Balbás, por haberme formado en competencias científicas y en competencias transversales como la idiosincrasia de la ciencia y sus científicos. El hecho de que no me pueda imaginar dedicándome a otra cosa en el futuro te lo debo en gran medida a ti. Gracias por las discusiones regadas con café (quizás más café del que deberíamos) y por no estar siempre de acuerdo conmigo. Eres una de las personas que más ha marcado mi vida y si retrocediera en el tiempo, elegiría de nuevo tu laboratorio para hacer la tesis. Gracias por motivarme de manera extraordinaria, apoyarme y darme espacio para crear. Mis éxitos científicos siempre serán en gran medida tuyos.

Continuaré dando las gracias a mi familia. Primeramente, a mi madre, que desde mi nacimiento sacrificó su bienestar por el mío y luchó para que las dos saliéramos ilesas en una carrera llena de obstáculos. Gracias por enseñarme a ser generosa y solidaria y por transmitirme tu buen humor. Lo único que no te agradezco es el gen de la impaciencia, no ha sido un buen complemento para esta tesis, ¿o sí?, jaja. Aunque nuestros mundos sean ahora tan diferentes, siempre serás un orgullo para mí, nunca lo olvides. Continúo dando las gracias a mis abuelas y a mi abuelo, que aunque lleven años muertos siguen presentes en mi mente como en mi infancia. Por un lado, gracias a mi abuela María “la Andaluza” por haberme querido sin medida y por haber participado activamente en mi educación. Solo lamento no haber podido enseñarte a leer y aunque no sirva de nada, de manera simbólica prometo que todo lo que la lectura me enseñe lo aplicaré con valor y estrategia como tú hiciste durante tu vida. Gracias también a mi abuelo Juan por confiar ciegamente en mí, ¡cuánto me avergonzaba que presumieras de nieta!... Gracias por hacerme sentir que te fortalecía hasta el último día de tu vida, te echo extraordinariamente de menos. Gracias a mi otra abuela María. Con los años me he dado cuenta de todos los valores feministas que me inculcabas y lamento infinitamente que los tuvieras que aprender de una manera atroz. Nunca se te agradecerá lo suficiente el esfuerzo que hiciste por todos los nietos y el amor que nos diste a todos y a cada uno, eras una persona admirable. Sigo con este emocionante ejercicio de memoria para dar las gracias a mi padre. Tantos años escuchando que somos tan parecidos me han hecho darme cuenta de lo difícil que es mantenerse cuerdo siendo idealista. Gracias por inculcarme valores éticos y este ansia por saber. Perdono tus fallos y utilizo tus fracasos como motivación en mi manera de vivir, ojalá pueda seguir compartiendo contigo mis vivencias muchos años. No podría terminar esta sección de agradecimientos sin dar las gracias a mi tío José Luis, que a pesar de ostentar el record mundial de ahijadas ha tenido tiempo y ganas de cuidar de mí. También gracias a mis tías Mari Carmen y Juana por ser mucho más que mis tías y celebrar mis triunfos como si fueran suyos. Gracias a José por tener ese corazón de oro, valoro mucho lo que has hecho por mí y sobre todo lo que haces por mi madre, te pido perdón por las veces que como adolescente me puse un poco rabuda jaja. Ahora que somos familia, tengo que añadir aquí los agradecimientos a mis suegros Conchi y Nacho. Gracias por acogermme sin peros y por contribuir con vuestra experiencia vital a mi vida.

Acknowledgements

Paso ahora a dar las gracias a mis compañeras del laboratorio de Marian. ¡Cuántas horas hemos compartido! Gracias a todas las que habéis pasado por ahí durante estos años, no hay ni una sola de vosotras que no me haya hecho aprender. Empiezo por estricto orden cronológico. Gracias a Conchi por ser mi primera mentora, el paso del tiempo me ha mostrado el estrés que supone estar en el último año de tesis y perdono que alguna vez fueras un poco impaciente conmigo. Gracias por enseñarme las técnicas de laboratorio y la responsabilidad que ser investigadora conlleva, te deseo lo mejor en el mundo científico, porque tu pasión lo merece. Continúo dando las gracias a Alejandra, la persona de la que más aprendí en el laboratorio. Gracias por no mirar el reloj a la hora de discutir sobre mi proyecto, por despertar en mí la curiosidad por la biología evolutiva y por reafirmarme en mi visión del mundo, espero que la vida te depare muchas alegrías. Gracias Mari por reírte de todas las bobadas que he dicho durante los años y por hacerme reír con tu amor por Jaume y por el actor, te deseo lo mejor en la vida y que sueñes todo lo que te propongas, como Alejandra y yo jaja...Gracias a Miriam Balada por hacerme crecer como mentora y hacerme sentir que podía transmitir mi motivación por la ciencia. Ευχαριστώ a Stella, que a pesar de que siempre decimos mal su nombre nos quiere igual. Gracias por escucharme durante estos años y ayudarme con mi proyecto. Quiero decirte que envidio tu exactitud haciendo los experimentos... ¡ojalá algún día me salgan a mí así los controles!, me lo he pasado genial contigo este tiempo. Grazie mille a Simona por enseñarme que los postres pueden llevar pistacho... ¡es broma! Gracias por ser tan buena compañera y valorar mis consejos, ¡da gusto trabajar con gente como tú!, espero que no te hayan molestado mucho mis ganas constantes de hablar jaja. Gracias a Claudia por enseñarme tantos trucos técnicos y por enseñarme lo valioso que es tener las cosas claras, fue un placer trabajar contigo. Gracias a Marta por hacerme sentir tan orgullosa con sus logros durante el Máster y por estar tan loca como para querer trabajar con JMJD3 y los enhancers. Valoro muchísimo tu prudencia y madurez y te deseo lo mejor en esta aventura que comienzas. Por último, doy las gracias a Elia por ser tan risueña y alegrarse de mis triunfos a pesar de que nos conocemos poco, seguro que tendrás un futuro brillante.

No nos engañemos, en el IBMB todos nos debemos favores y eso es maravilloso. Agradezco aquí el buen ambiente que hay entre estudiantes y postdocs y ojalá que en mis futuros laboratorios sea la mitad de bueno. Empiezo dando las gracias a Sonia Najas, diosa de la corticogenesis. Gracias por valorarme tanto como científica y estar siempre ahí cuando te he necesitado. Espero que seas capaz de ver la gran investigadora que hay en ti y te confieso que cada vez que diseño un experimento intento ser paciente y reflexiva como tú. Gracias a Juan Arranz por ser tan buen compañero y por hacerme reír hasta el infinito, a Mariajo por ser tan pragmática, simpática y sobre todo, por odiar la homeopatía como yo jaja (aunque no está la cosa para reírse...). Gracias a Isa por quedarse siempre hasta tan tarde y no hacerme sentir una loca jaja, en serio, gracias por ser tan buena compañera. Muchas gracias a todos los del laboratorio de Elisa: Lucía, Elena, Murielle, José...y especialmente a Susana por ser la persona más amable que conozco y a Gwen por hacerme ser mejor científica y hacerme tener en cuenta la complejidad del sistema nervioso a la hora de interpretar mis resultados. Por supuesto gracias a Elisa que me ha hecho sentir muy apreciada en el laboratorio ¡nunca pierdas esa pasión que te caracteriza! Gracias al laboratorio de Albert Jordan por esos seminarios tan divertidos y especialmente gracias a Andrea por ser una tía *de puta madre* y a Carles por ser el adalid de la transmisión oral de cotilleos científicos de hoy, ayer y de siempre. Me encantan tus historias de como se hacía X técnica hace años o de cotilleos de los científicos, esos seres. Gracias a Ismael por ser

Acknowledgements

un derroche de bondad. Gracias a Carme Gallego por enseñarme a clonar, aunque quizás no te acuerdes. Gracias a Sara Gutiérrez, por enseñarme que hay muchas maneras de ser científica. Sobre todo, gracias a Raúl por todas las horas de conversación acerca de si los científicos estamos locos o no y por estar en el laboratorio días y horas intempestivas confirmándome que efectivamente, muy cuerdos no estamos. Ojalá nuestros caminos se vuelvan a cruzar, eres un gran amigo y compañero. Por último, quiero dar las gracias a los del laboratorio de Joaquín Roca, sois todos majísimos y siempre me habéis ayudado, sobre todo gracias a Toni que es una excelente persona y un gran conversador, ¡nos vemos en el Instituto de cromatina de Mallorca! También, aunque no haya participado directamente en mi tesis quiero dar las gracias a Leonor Calopa, la mujer que hace que el IBMB se mantenga en pie, gracias por toda la ayuda administrativa.

De mi estancia en Colonia quiero dar las gracias a Álvaro Rada-Iglesias, por darme la oportunidad de ir a su laboratorio a adquirir otra visión científica y por ayudarme a conseguir mis futuras metas. Muchísimas gracias a Sara de la Cruz por ayudarme con las cuestiones técnicas y por compartir conmigo la transición PhD-postdoc con tanto humor, te deseo lo mejor.

Menudo rollo me está quedando...aunque como no tengo Facebook quizás es una buena manera de recopilar a mis amigos (supongo que cuando en el futuro las generaciones cyborg lean esto sonreirán pícaramente).

Paso a dar las gracias a mis amigos, que como ellos saben son también mi familia. ¡Cómo no! primero gracias a Rubén y a Jona, haberos conocido ha sido de las mejores cosas que me han pasado en la vida y gracias a vosotros, soy mejor persona. Gracias a Rubén por compartir mi pasión por la música y por moldear mi pensamiento político. Gracias a Jona por integrar en frases cortas soluciones a mis problemas, os quiero. Gracias a Pablo y a Laia, que me acogieron en su círculo de amistades y me hicieron sentir súper querida. Especialmente gracias a Laia por ser una gran amiga y porque sé que estará ahí siempre. Gracias a Irene, con la que comparto opiniones y enfermedades, eres de mis mejores amigas y haberte conocido es de las mejores cosas que me pasó en los últimos años. Gracias a Cris por haber sido el cuarto miembro de Parchís y por hacerme sentir muy querida, eres una buena amiga y una madre estupenda. Gracias a Sònia por ser tan divertida en momentos inesperados y estar siempre ahí. Gracias a Uxio, por ser un gran amigo y por valorarme tanto, te aseguro que es mutuo. Gracias a Triple A, por haber entrado en mi vida y ser imposible de echar... ¡qué tío más *de puta madre* eres! Gracias a Ana Martínez, que aunque haya sido de mis últimas adquisiciones en amistad, es una tía admirable y como Mariajo (leer más arriba), odia la homeopatía. Gracias a todas las Mallorquettes, especialmente a Virginia, con la que me comunico por telepatía. De la maravillosa isla gracias también a Carles, ¡eres un ángel-demoninchi!. Gracias a mis amigos asturianos por horas interminables de risas. Gracias a Pet y a Mike por mostrarme el erotismo de la extravagancia, por creer siempre en mí y demostrármelo. Gracias a Jairo y a Guti por ser mis cómplices en las fechorías leonesas. Gracias a mis amigos Sergio y Asier por su larga fidelidad.

Gracias a todas las bandas que han creado y crean la música que me apasiona. Especialmente gracias a los Kinks, a Teenage Fanclub y a Daniel Romano. Gracias al programa de radio Hotel Arizona y a su creador por facilitarme el escuchar música con sus horas de podcasts descargables, ¡son la mejor compañía para mis experimentos!

Acknowledgements

Gracias a Paco, mi profesor de Biología de primero de Bachillerato por motivarme tanto. A Adolfo Figueiras, de la Universidad de Santiago de Compostela por imprimir en los alumnos el espíritu crítico y científico.

Desagradecimientos a los encargados de la toma de decisiones políticas acerca de la ciencia en España durante los años 2010-2018. Repulsa hacia los impulsores de la ignorancia y la falsa moral.

Tras haber agradecido a seres vivos e inertes, paso a agradecer todo lo que he conseguido estos años al que es mi two-of, mi compañero vital, amante y marido: Alberto Gamazo. Gracias por ser el mayor apoyo que he tenido en mi vida y por saber que nunca me dejarás caer. A ciencia cierta sé que nadie se ha reído jamás lo que nosotros nos hemos reído y saber que nos seguiremos riendo del mundo juntos es uno de los motivos más poderosos para existir. Esta tesis te la debo a ti más que a nadie que no sea yo. Comprendes los pormenores de este oficio tedioso más que los nativos y has facilitado cada uno de mis pasos con ayuda psicológica, logística o humorística. Gracias por hacerme ser más valiente y por retarme cada día. Espero que siempre seas el cómplice de todos mis éxitos.

Table of contents

Abbreviations	1
Introduction	3
1. Chromatin is the platform for DNA regulatory events.....	3
1.1 Linear components and aspects of the chromatin.....	4
1.1.1 Histones.....	4
1.1.1.1 Histone acetylation and deacetylation.....	5
1.1.1.2 Histone methylation and demethylation.....	5
1.1.1.2.1 JmJC domain KDMs: the KDM6 subfamily.....	7
1.1.1.2.1.1 JMJD3: a member of the KDM6 subfamily.....	9
1.1.1.3 Other post-translational modifications of histones.....	10
1.1.2 Regulatory regions on the DNA sequence.....	10
1.1.2.1 Gene promoters.....	11
1.1.2.2 Enhancers.....	11
1.1.2.2.1 Enhancer identification.....	12
1.1.2.2.2 Transcriptional activity at enhancers.....	13
1.1.2.2.3 Superenhancers.....	13
1.1.2.3 Silencers and insulators.....	14
1.1.3 Proteins that act on the linear chromatin.....	14
1.1.3.1 Transcription factors.....	14
1.1.3.1.1 Pioneer transcription factors.....	14
1.1.3.1.1.1 The pioneer factor ASCL1.....	15
1.1.3.2 Transcriptional coactivators.....	15
1.1.3.2.1 The chromatin remodeler CHD8.....	16
1.1.3.3 RNA-polymerase II.....	16
1.2 Tridimensional components and aspects of chromatin.....	17
1.2.1 Levels of 3D-chromatin organization.....	17
1.2.1.1 Chromosome territories.....	17
1.2.1.2 Topological associated domains.....	18
1.2.1.3 Chromosomal compartments.....	19
1.2.1.4 Loop domains.....	20
1.2.1.4.1 Loop maintenance complex of proteins.....	20
2. The cortex development as a model of study.....	21
2.1 Definition, function and development of the cortex.....	21
2.1.1 Developmental cues in corticogenesis.....	22
2.1.1.1 Extrinsic factors.....	23

Table of contents

2.1.1.2	Intrinsic factors.....	23
2.2	Neural stem cells as an <i>in vitro</i> model of study.....	24
2.2.1	Neural stem cells <i>in vitro</i> versus <i>in vivo</i>	25
3.	The TGF β pathway.....	26
3.1	The TGF β superfamily of ligands.....	26
3.2	Definition and characteristics of the TGF β pathway.....	26
3.3	TGF β -pathway signaling cascade.....	27
3.4	SMAD proteins as the effectors of the TGF β -pathway.....	28
3.4.1	Types of SMAD proteins.....	28
3.4.2	Structural aspects of the SMAD proteins.....	29
3.4.3	Functions of the SMAD proteins in the chromatin.....	29
3.5	Functions of the TGF β pathway.....	30
3.5.1	Physiological functions of the TGF β pathway.....	30
3.5.1.1	Role of the TGF β pathway in proliferation.....	30
3.5.1.2	Role of the TGF β pathway in differentiation.....	30
3.5.1.3	Role of the TGF β pathway in apoptosis and in homeostasis.....	31
3.5.2	Pathological functions of the TGF β pathway.....	32
3.5.2.1	The TGF β pathway in cancer.....	32
3.5.2.2	The TGF β pathway in other diseases.....	32
3.6	The TGF β pathway in neural development and its relation to JMJD3.....	33
3.6.1	SMAD3 and JMJD3 coregulate the neural developmental program.....	33
3.6.2	SMAD3 and JMJD3 colocalize on gene promoters.....	33
3.6.3	TGF β promotes neurogenesis in the spinal cord in a JMJD3-dependent manner.....	33
3.6.4	TGF β -induced transcription elongation is mediated by JMJD3.....	34
Objectives	36
Methods	37
1.	Materials.....	37
1.1	Models of study.....	37
1.1.1	<i>In vitro</i> models.....	37
1.1.1.1	Mouse neural stem cells.....	37
1.1.1.2	Human HEK293T cells 36.....	37
1.1.2	<i>In vivo</i> model.....	38
1.1.2.1	Chicken neural tube.....	38
1.2	Reagents.....	38
1.3	Plasmids.....	39

Table of contents

1.4	Antibodies.....	39
1.5	Primers.....	40
1.6	Gene expression omnibus accessions.....	41
2	Experimental methods.....	41
2.1	Cell culture and genetic manipulation of cell lines.....	42
2.1.1	Cell culture growth and maintenance.....	42
2.1.1.1	Mouse neural stem cells.....	42
2.1.1.1.1	Mouse neural stem cells expansion.....	43
2.1.1.1.2	Mouse neural stem cells thawing and freezing.....	43
2.1.1.1.3	Mouse neural stem cells differentiation experiments.....	43
2.1.1.2	HEK293T cells.....	44
2.1.1.2.1	HEK293T cells expansion.....	44
2.1.1.2.2	HEK293T cells freezing and thawing.....	44
2.1.2	Genetic manipulation of growing cells.....	44
2.1.2.1	Calcium phosphate transfection.....	44
2.1.2.2	Nucleofection.....	45
2.1.2.3	Lentiviral transduction.....	45
2.1.2.4	CRISPR-Cas9 cell line generation.....	45
2.2	Chicken embryo manipulation.....	46
2.2.1	Ovoelectroporation of chicken embryonic neural tubes.....	46
2.3	Molecular biology procedures.....	47
2.3.1	Nucleic acids-related.....	47
2.3.1.1	Genomic DNA extraction.....	47
2.3.1.2	Phenol chloroform extraction and ethanol precipitation.....	47
2.3.1.3	RNA extraction and DNase treatment.....	47
2.3.1.4	Retrotranscription of RNA into cDNA.....	48
2.3.1.5	Conventional PCR.....	48
2.3.1.6	qPCR.....	48
2.3.1.7	Cloning of plasmidic DNA.....	49
2.3.1.7.1	pLKO.1 constructs.....	49
2.3.1.7.2	Luciferase constructs.....	50
2.3.1.7.3	CRISPR-Cas9 constructs.....	50
2.3.1.8	Amplification of plasmidic DNA.....	50
2.3.1.8.1	Mini- and maxi- preparations of DNA.....	50
2.3.1.9	DNA and RNA electrophoresis.....	51
2.3.1.10	4C-seq.....	51
2.3.2	Protein-related techniques.....	54
2.3.2.1	Total protein extraction.....	54

Table of contents

2.3.2.2	Nuclear cytoplasmic extraction.....	54
2.3.2.3	Protein quantification: Bradford.....	54
2.3.2.4	SDS-Page electrophoresis.....	55
2.3.2.5	Western Blot.....	55
2.3.2.6	Size exclusion chromatography.....	56
2.3.2.7	Coimmunoprecipitation.....	57
2.3.2.8	Chromatin immunoprecipitation.....	57
2.3.2.8.1	Chromatin immunoprecipitation of JMJD3.....	58
2.3.2.9	Indirect immunofluorescence.....	59
3	Bioinformatic methods.....	59
3.1	ChIP-seq data acquisition and analysis.....	59
3.1.1	Enhancer identification with ChIP-seq data.....	60
3.1.2	Putative regulated gene assignation.....	60
3.1.3	Enhancer gene ontology analysis: GREAT.....	60
3.1.4	Venn diagrams construction.....	61
3.1.5	Heatmaps construction.....	61
3.1.6	ChIP-seq capture obtaining.....	61
3.2	4-C bioinformatic analysis.....	61
3.2.1	Box-plot construction.....	62
3.3	RNA-seq and microarray comparison.....	62
4	Statistical analysis.....	62
4.1	Sample size.....	62
4.2	Standard deviation and standard error of the mean.....	62
4.3	Student's t-test.....	63
4.4	Saphiro-Wilk test.....	63
4.5	Equal proportion tests.....	63
4.6	Paired Wilcoxon test.....	63
	Results.....	64
1.	Epigenetic identification of neural enhancers.....	64
2.	TGF β regulation of neural enhancers.....	65
2.1	SMAD3 is bound to neural enhancers upon TGF β treatment.....	65
2.1.1	Interaction between the proneural factor ASCL1 and SMAD3.....	66
2.2	TGF β -dynamic activation of neural enhancers.....	68
2.2.1	Activation evaluated with eRNA transcription.....	68
2.2.2	Activation evaluated with histone modification acquisition.....	70
2.2.3	Activation evaluated with in vivo experiments.....	71
3.	JMJD3 cooperation in the SMAD3 and ASCL1 bound enhancers.....	71

Table of contents

3.1 Interaction between ASCL1 and JMJD3.....	73
3.2 Role of JMJD3 in the TGF β activation of enhancers.....	74
3.2.1 Contribution of JMJD3 to enhancer activation evaluated by eRNA.....	75
3.2.2 Role of JMJD3 in the acquisition of active enhancers histone marks.....	76
3.2.3 Role of the H3K27me2/3 in enhancer activation.....	76
4. Contribution of CHD8 to the TGF β -responsive enhancers activation.....	79
5. TGF β re-organizes the tridimensional structure of the chromatin.....	82
5.1 Design of the 4C-seq experiments.....	83
5.2 TGF β induces enhancer-enhancer and enhancer-promoter contacts.....	84
5.3 JMJD3 is necessary for TGF β -triggered contact formation.....	86
5.3.1 Cohesin contribution to the 3D-chromatin structure.....	87
6. <i>Neurog2(-6)</i> enhancer and <i>Neurog2</i> promoter cooperate to fine-tune transcription.....	89
6.1 Experimental strategy.....	90
6.2 Hierarchy between enhancers and promoters in the TGF β -signaling.....	92
6.2.1 Contribution of the <i>Neurog2</i> promoter to transcription.....	92
6.3 Contribution of the <i>Neurog2(-6)</i> enhancer to transcription.....	93
6.3.1.1 <i>Neurog2(-6)</i> enhancer contribution to neurogenesis.....	95
Discussion.....	98
1. About the epigenetic identification of enhancers.....	98
1.1 About the identified poised enhancers.....	98
2. About the TGF β regulation of enhancers.....	99
2.1 About the interaction of ASCL1 and SMAD3 at enhancers.....	99
2.2 About the TGF β activation of enhancers.....	100
2.2.1 About the enhancer activation measured as eRNA transcription.....	100
2.2.2 About the enhancer activation measured as histone mark acquisition.....	101
2.2.3 About the enhancer activation <i>in vivo</i>	102
3. About the localization of JMJD3 at TGF β activated enhancers.....	103
3.1 About the interaction of JMJD3 and ASCL1.....	103
3.2 About the role of JMJD3 in the activation of enhancers.....	104
4. About the contribution of CHD8 to the TGF β enhancer activation.....	107
5. About the TGF β re-organization of the chromatin.....	108
5.1 About the design of the 4C-seq experiment.....	108
5.2 About the TGF β -induction of a functional superenhancer formation.....	109
5.3 About the contribution of JMJD3 to the functional superenhancer formation.....	111
6. About the <i>Neurog2(-6)</i> enhancer and <i>Neurog2</i> promoter cooperation in the fine-tuning of transcription.....	112
6.1 About the contribution of promoters to transcription.....	113

Table of contents

6.2 About the contribution of enhancers to transcription.....	114
6.2.1 About the contribution of the <i>Neurog2</i> (-6) enhancer to neurogenesis.....	115
7. Global analysis of the results.....	116
Conclusions.....	117
Bibliography.....	118
Addendums.....	119
1. Publications.....	119
1.1 Publications as a first author.....	119
1.2 Other publications.....	119

Abbreviations

Abbreviations

3D: tridimensional
A: adenine
AP: apical progenitors
ATP: adenosine triphosphate
BP: basal progenitors
bHLH: basic helix-loop-helix
BMP: bone morphogenetic proteins
bps: base pairs
BSA: bovine serum albumin
C: cytosine
ChIP-seq: chromatin immunoprecipitation coupled to genomic sequencing
CNT: chicken neural tube
CNS: central nervous system
CTD: carboxy-terminal domain
DNA: deoxyribonucleic acid
DNTP: deoxynucleotide triphosphate
DSG: di(N-succinimidyl) glutarate
EGF: epidermal growth factor
EMT: epithelial to mesenchymal transition
EP: electroporation / electroporated
ESCs: embryonic stem cells
E.X: day Xth of mouse embryonic development e.g. E11, E12.5
eRNA: enhancer RNA
FBS: fetal bovine serum
FGF: fibroblast growth factor
G: guanine
GEO: gene expression omnibus
GF: growth factor
GO: gene ontology
gRNA: guide-RNA
GTF: general transcription factors
H: histone
H3K27me2: histone 3 lysine 27 dimethylation
H3K27me3: histone 3 lysine 27 trimethylation
HAT: histone acetyl transferase
HDAC: histone deacetylase
HH: Hamburger and Hamilton
HKMTs: histone lysine methyl transferases
JDTA genes: genes regulated by TGF β in a JMJD3-dependent manner
JmJC: Jumonji-C domain containing proteins
K: lysine

Abbreviations

KDMs: histone lysine demethylases
KO: knockout
LB: lysogeny broth
LMC: loop maintenance complex of proteins
Mb: megabase
MLL: mixed lineage leukemia
MW: molecular weight
NE: neuroepithelial progenitors
nt: nucleotides
NSCs: neural stem cells
PBS: phosphate buffered saline
PcG: Polycomb group of proteins
PCR: polymerase chain reaction
PIC: pre-initiation complex
PNK: T4-polynucleotide kinase
PTMs: post-translational modifications
RCF: relative centrifugal force
RG: radial glial cells
RIPA: radioimmunoprecipitation assay buffer
RNAPII: RNA-polymerase II
RNAPII-Ser5p: RNA-polymerase II phosphorylated at serine 5
RNAPII-Ser2p: RNA-polymerase II phosphorylated at serine 2
RPM: reads per million
RT-qPCR: retrotranscription followed by quantitative polymerase chain reaction
SBE: SMAD binding elements
SD: standard deviation
SDS: sodium dodecyl sulfate
SEM: standard error of the mean
Shh: Sonic Hedhog
SVZ: subventricular zone
T: thymine
TBE: Tris-borate-EDTA
TBP: TATA-box binding protein
TF: transcription factors
TAD: topological associated domain
TGF β : transforming growth factor beta
TGFBR: transforming growth factor beta receptor
TPR: tetratricopeptide repeats
TSS: transcriptional start site
V: volts
VZ: ventricular zone

Introduction

During embryonic development, signaling pathways, transcription factors (TF) and epigenetic regulators are orchestrated in a spatial-temporal manner, in order to promote the required gene expression programs (Buecker and Wysocka, 2012). Specifically, we have elucidated how the TGF β pathway together with lineage specific TF and chromatin modifying enzymes regulate neural stem cells (NSCs) development. We have focused our efforts to understand the role that the histone demethylase JMJD3 plays on these events.

In this section, I will review the literature on chromatin, neural development and TGF β -signaling in a point-by-point fashion in order to introduce the context in which this thesis has been elaborated.

1. Chromatin is the platform for DNA regulatory events

In 1878, Dr. Walther Flemming discovered a fibrous structure inside the cell nucleus that could be stained with basophilic dyes. Due to its color acquisition, Dr. Flemming called this material **chromatin**, from the greek *chroma*, color (Flemming, 1880).

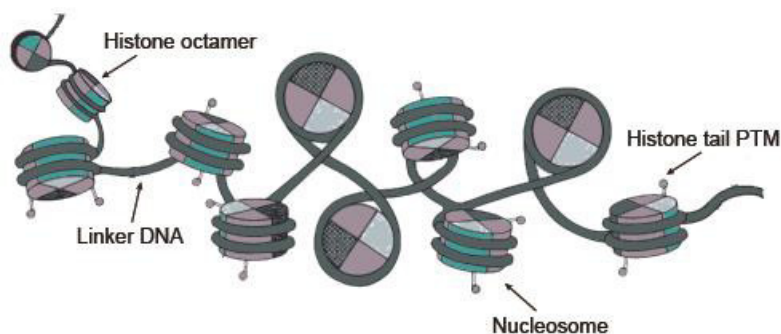


Figure I1. Schematic representation of nucleosomes and linker DNA conforming the chromatin. Adapted from (Rosa and Shaw, 2013).

Since those years, many physical, chemical and biological aspects of chromatin have been discovered. Chromatin is defined as the polymer in which deoxyribonucleic acid (DNA) is packed inside the nucleus. This polymer is composed of **nucleosomes**, that are formed by 147 base pairs (bps) of DNA wrapped around an octamer of histones (H) (two molecules of histone H3, two of H4, two of H2A and two of H2B) and of linker DNA bound by the histone H1 that holds the cores together (Kouzarides, 2007) (Figure I1). Observations with optical and electronic microscopy led to the classical bimodal description of eukaryotic chromatin: the gene-rich, permissive to transcription euchromatin, is positioned at the center of the nucleus, while heterochromatin, that corresponds to a gene-poor and compacted form of chromatin is generally positioned at the nuclear periphery (Ea et al., 2015).

Genome transcription occurs in the chromatin polymer. In this regard, chromatin coordinates transcriptional events at the tridimensional (3D) and the linear levels. In the following subsections I will deepen more in the molecules and mechanisms that operate at the different scales.

Introduction

1.1. Linear components and aspects of the chromatin

As I defined above, chromatin is composed of DNA and histones (Kornberg, 1974). These two elements bring with them layers of regulation for the genomic transcription.

1.1.1. Histones

Among eukaryotes, histones are the most conserved proteins. These alkaline molecules are composed of the histone fold domain and the **N-terminal tail**. The histone fold is also called globular domain, it interacts with the other core histones and is where DNA is wrapped around (Figure I2). On the other hand, N-terminal tails protrude from the core of nucleosomes and participate in gene regulation, nucleosome stability and condensation (Marino-Ramirez et al., 2005). Both the globular domain and the N-tail are subjected to post-translational modifications (PTMs) (Lawrence et al., 2016), nonetheless, in this thesis I will focus on the N-terminal tail PTMs.

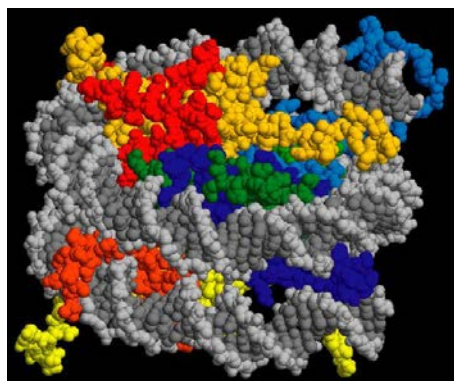


Figure I2. Ras Top representation of the core crystallographic structure of the nucleosomes. Histones H2A are represented in yellow, H2B in red, H3 in blue and H4 in green. DNA is represented in grey. Adapted from (Harp et al., 2000).

For many years it was believed that nucleosomes were an obstacle for DNA transcription (Wasylyk and Chambon, 1979). However, nowadays it is well accepted that they play an essential regulatory role in this process. In this regard, PTMs occurring in the N-tails might chemically modify the amino acids charge. The addition of different moieties to the N-tail residues elicits chemical changes that ultimately affect the interaction between the histones and the DNA, leading to compaction or decompaction (Kouzarides, 2007; Lee et al., 1993). Additionally, PTMs located at N-tails affect the inter-nucleosomal interactions and serve as a mechanism of protein targeting for remodelers and other chromatin acting proteins, endowing chromatin with high versatility and **regulatory capacity** (Bannister and Kouzarides, 2011). In this regard, the different PTMs exert different functions (Table I1).

1.1.1.1. Histone acetylation and deacetylation

Lysine acetylation and deacetylation is a dynamic process carried out by histone acetyl transferases (HATs) and histone deacetylases (HDACs). HATs utilize the acetyl CoA as a cofactor to catalyze the addition of an acetyl group to the ϵ -amino group of the lysine thus, neutralizing its positive charge. This reaction weakens the interaction between histones and DNA. For this reason, HATs usually work as coactivators and HDACs as corepressors (Hodawadekar and Marmorstein, 2007; Yang and Seto, 2007) (Table I1). One of the most studied HATs is **p300/CBP**. This enzyme catalyzes the acetylation of different lysines, being H3K27 acetylation the most relevant for this thesis due to its location at enhancers (Creyghton et al., 2010; Das et al., 2009).

Chromatin Modifications	Residues Modified	Functions Regulated
Acetylation	K-ac	Transcription, Repair, Replication, Condensation
Methylation (lysines)	K-me1 K-me2 K-me3	Transcription, Repair
Methylation (arginines)	R-me1 R-me2a R-me2s	Transcription
Phosphorylation	S-ph T-ph	Transcription, Repair, Condensation
Ubiquitylation	K-ub	Transcription, Repair
Sumoylation	K-su	Transcription
ADP ribosylation	E-ar	Transcription
Deimination	R > Cit	Transcription
Proline Isomerization	P-cis > P-trans	Transcription

Table I1. Summary of the main types of histone modifications, their target residues and the regulated cellular function. Adapted from (Kouzarides, 2007).

1.1.1.2. Histone methylation and demethylation

Histone methylation consists in the addition of one, two or three $-\text{CH}_3$ groups to lysines or arginines, and is performed by histone lysine methyl transferases (HKMTs) or arginine methyltransferases. This reaction does not change the electrical charge of the amino acid, but it does have functional consequences (Rea et al., 2000) (Table I1). There are many enzymes that catalyze methylation (Lan and Shi, 2009), but in this introduction I will focus on the most relevant for this thesis: EZH2/1 and MLL, that form part of the PRC2 and MLL complexes respectively.

During cellular identity acquisition, the antagonistic effect of **PcG** (Polycomb group of proteins) and **MLL** complexes control the appropriate expression of key developmental genes (Schuettengruber and Cavalli, 2009). PcG complexes mediate gene repression and chromatin compaction, and MLL complexes are gene activators. Concretely, from PcG, PRC2 is known to mediate H3K27 di- and tri-methylation (H3K27me2/3) repressing cell identity and developmental genes

Introduction

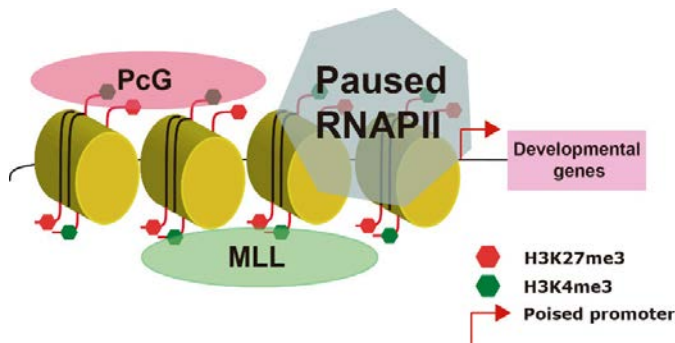


Figure I3. Schematic representation of a bivalent promoter.

enhancers, specifically, H3K4me1/2 is found at enhancers, and H3K4me2/3 at promoters (Bochynska et al., 2018). Interestingly, H3K27me3 and H3K4me3 coexist on some promoters called **bivalent promoters**. These are found on developmental genes in stem cell states and endow the cell with the capacity of a rapid activation upon developmental cues (Bernstein et al., 2006) (Figure I3).

During a long time, histone methylation was thought to be an irreversible modification whose loss was due to histone exchange. However, the identification of the first histone **lysine demethylase** (KDM) in 2004 changed this dogma (Shi et al., 2004). In the last years, many KDMs have been discovered, reviewed in (Fueyo et al., 2015), and according to their different mechanisms of action they have been classified in two families: the LSD family (or KDM1 family) and the JmJC (Jumonji-C domain containing proteins) family.

(Margueron and Reinberg, 2011). On the other hand, the activating MLL complexes operate on H3K4, and by methylating this residue facilitate gene expression activation. The different methylation levels on H3K4 have been correlated with the presence of gene promoters or

FAMILY	HUMAN HKDM	OTHER NAMES	HISTONE SUBSTRATE
KDM1	KDM1A	LSD1	H3K4me2/me1, H3K9me2/me1
	KDM1B	LSD2	H3K4me2/me1, H3K9me2/me1
JmJC	KDM2A	FBXL11	H3K36me2/me1
	KDM2B	FBXL10	H3K4me3, H3K36me2
	KDM3A	JMJD1A	H3K9me2/me1
	KDM3B	JMJD1B	
	JMJD1C	TRIP8	
	KDM4A	JMJD2A	H3K9me3/me2, H3K36me3/me2, H1.4K26me3
	KDM4B	JMJD2B	
	KDM4C	JMJD2C	
	KDM4D	JMJD2D	H3K9me3/me2/me1, H1.4K26me3/me2
	KDM5A	JARID1A	H3K4me3/me2
	KDM5B	JARID1B	
	KDM5C	JARID1C	
	KDM5D	JARID1D	
	KDM6A	UTX	H3K27me3/me2
	KDM6B	JMJD3	
UTY			
KDM7A	KIAA1718	H3K9me2/me1, H3K27me2/me1	
PHF8	KIAA1111	H3K9me2/me1, H4K20me1	
PHF2	KIAA0662	H3K9me2/me1, H4K20me3	

Table I2. Compendium of all the known KDMs, their aliases and histone substrates. Adapted from (Fueyo et al., 2015).

The LSD family was the first one to be discovered and comprises only two enzymes: LSD1 and LSD2. These enzymes demethylate H3K4me1/2 and H3K9me1/2 in a flavin-adenine dinucleotide-dependent reaction (Ciccone et al., 2009; Metzger et al., 2005; Shi et al., 2004). On the other hand, the **JmJC family** utilize Fe^{2+} and 2-oxoglutarate in the demethylation reaction and since the discovery of its first member in 2006 (Tsukada et al., 2006), the number of JmJC proteins has enormously grown (Fueyo et al., 2015; Kooistra and Helin, 2012) (Table I2).

1.1.1.2.1. *JmJC domain KDMs: the KDM6 subfamily*

In the following lines, I will describe more in depth the **JmJC family** of KDMs due to its importance in this thesis.

JmJC family of KDMs are a group of proteins that have been found in more than 100 species, from bacteria to eukaryotes (Clissold and Ponting, 2001). As I briefly mentioned above, their mechanism of action consists in a reaction in which the quaternary complex formed by the substrate bound to the catalytic domain, together with Fe^{2+} and α -ketoglutarate reacts with oxygen. The oxidative decarboxylation of α -ketoglutarate coordinated with the hydroxylation of the target methyl group creates a hydroxymethyl ammonium intermediate that gives rise to formaldehyde and the demethylation product (Tsukada et al., 2006) (Figure I4).

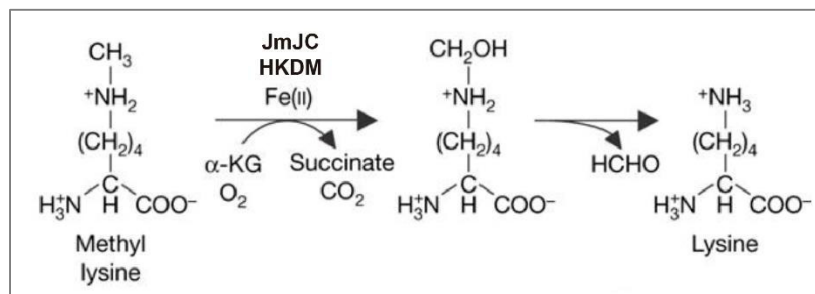


Figure I4. Chemical reaction of lysine demethylation by a JmJC-domain containing KDM. Adapted from (Tsukada et al., 2006)

KDMs can be subclassified into seven subfamilies according to their structural domains and specificity for histones (Table I2). In this thesis, I will particularly introduce the KDM6 subfamily due to its relevance in the experimental section.

The **KDM6 subfamily** of JmJC KDMs is composed of three members: KDM6A (UTX), KDM6B (JMJD3) and KDM6C (UTY) (De Santa et al., 2007; Lan et al., 2007; Lee et al., 2007). This subfamily appears in bilateral animals, even though some species only have one of the members (Zerbino et al., 2018). Besides, they are ubiquitously expressed (Uhlen et al., 2015). The difference between the three members relies on their structure. UTX and UTY contain a JmJC domain and several

Introduction

tetratricopeptide repeats that mediate protein-protein interaction. However, JMJD3 has only the JmJC domain (De Santa et al., 2007) (Figure I5). Another peculiarity of the KDM6 subfamily is their specificity. KDM6 KDMs demethylate the **H3K27me_{2/3}**, a facultative heterochromatin histone mark related to development and deposited by the PRC2 complex (Lan et al., 2007) (Margueron and Reinberg, 2011). These enzymes exert their activities in different contexts, and due to its ubiquitous expression they are involved in different physiological and pathological processes: hypoxia, inflammation, differentiation, cancer and others (Burchfield et al., 2015).

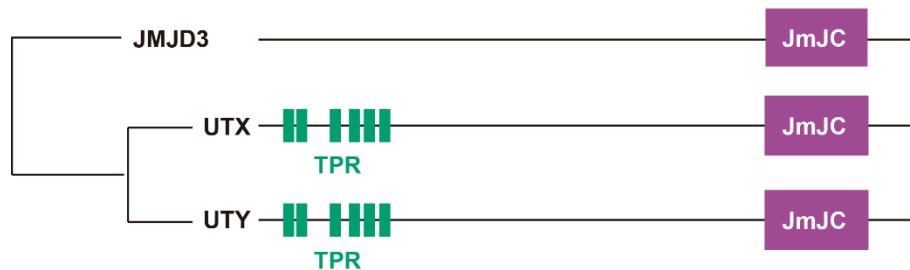


Figure I5. Schematic representation of the KDM6 subfamily of demethylases phylogenetic tree. JmJC: JumonjiC domain; TPR: tetratricopeptide repeats

1.1.1.2.1.1. JMJD3: a member of the KDM6 subfamily

As described above, the H3K27 methylation status is crucial for cell fate determination in stem cells. In concordance, H3K27me_{2/3} KDMs are involved in important aspects of development (Fueyo et al., 2015). Nonetheless, among the KDMs, JMJD3 is the protein that more frequently participates in **signaling** events. JMJD3, unlike UTX, is upregulated at the transcriptional level upon developmental, immune and oncogenic stress-related signals, being required stimulating cues for its activity (Agger et al., 2009; Burgold et al., 2008; De Santa et al., 2007). Even though it has been described that JMJD3 participates in proliferation, senescence, apoptosis and differentiation (Burchfield et al., 2015), in this thesis, I will focus on the participation of JMJD3 in developmental processes.

Different studies have analyzed the role of JMJD3 knockout (KO) experiments. In this regard, different results have been obtained depending on the KO strategy (Figure I6). Two different published works demonstrated that mutant mice carrying a homozygous deletion of the JmJC catalytic domain died perinatally due to premature alveolar development (Li et al., 2014; Satoh et al., 2010). In another research article, authors focused on the role of JMJD3 along neuronal maturation, and they found that JMJD3 is essential for the function of the Pre-Bötzinger complex, a cluster of interneurons that belong to the respiratory rhythm generator. Mice

lacking 50% of JMJD3 died peri- or neonatally (Burgold et al., 2012). Interestingly, the only study in which a total ablation of JMJD3 was generated claims that mice died at E6.5 (Ohtani et al., 2013), pointing to quintessential roles independently of the catalytic activity (Figure I6).



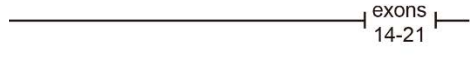


		Mutation	Phenotype
	JmJC	$Jmjd3^{-/-}$ in E16.5, 50% expression before	Perinatal death due to Pre-Bötzinger complex defects
	JmJC	Conventional deletion of JmJC catalytic domain	Perinatal death due to abnormal lung development
	exons 14-21	Conventional and conditional deletion of JmJC catalytic domain	Perinatal death due to abnormal lung development
	exons 15-21	Conventional and conditional deletion of JmJC catalytic domain	Perinatal death due to abnormal lung development
	exons 4-5 frameshift (total protein deletion)	$Jmjd3^{-/-}$	Embryonic lethality at E6.5

Figure I6. Phenotypical effects of *Jmjd3* KO experiments. Left panel shows a schematic representation of the followed strategy. Right table displays the obtained mutation and its correspondent phenotype. Figure based in (Burchfield et al., 2015).

Embryonic stem cells (ESCs) do not require JMJD3 for pluripotency maintenance (Ohtani et al., 2013). However, JMJD3 participates along the generation of the three germ layers. In all the cases described in the literature, JMJD3 demethylase activity is triggered upon a differentiation signal and by activating the promoters of lineage specific regulators switches on the endo- meso- or ectodermal developmental programs (Burgold et al., 2008; Kartikasari et al., 2013; Ohtani et al., 2013).

As this thesis is framed within the neurodevelopmental context, I will review the literature regarding **JMJD3 in the neural lineages** (Figure I7). Differentiation of ESCs into the neural lineage requires JMJD3 (Burgold et al., 2008), moreover, JMJD3 overexpression in NSCs triggers the transcription of neuronal genes in a KDM activity dependent manner (Jepsen et al., 2007). As I have previously mentioned, JMJD3 operates downstream developmental signals. It has been described that JMJD3 participates in retinoic acid neuronal generation by derepressing the promoter of *Ascl1*, a neurogenic protein (Dai et al., 2010).

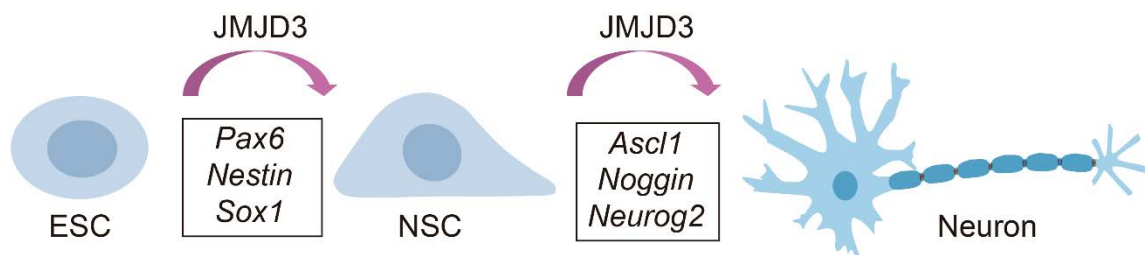


Figure I7. JMJD3 participates in neural commitment and neuronal differentiation. Names inside rectangles indicate some of the genes that are activated in each transition.

Introduction

In vivo experiments in the chicken neural tube (CNT) have demonstrated that TGF β -induced neuronal differentiation was dependent on JMJD3 KDM activity (Estaras et al., 2012). Moreover, upon BMP signal JMJD3 activates the promoter of *Noggin*, a negative regulator of BMP pathway thus, generating an inhibitory feedback regulatory loop that controls the generation of dorsal interneurons (Akizu et al., 2010) (Figure I7).

In the subsection 3.6 I will elaborate more on the results from Estarás et al. 2012 that served as a starting point for my doctoral thesis.

1.1.1.3 Other post-translational modifications of histones

Even though I have highlighted histone acetylation and methylation as the most relevant chromatin modifications in this thesis, there are many other PTMs with significant roles in transcriptional regulation. For example, phosphorylation of residues in the N-tail is a dynamic process carried out by kinases, this modification is triggered by extracellular signals, DNA damage and mitosis. It entails the addition of negative charge to the amino acid, regulating chromatin structure (Sawicka and Seiser, 2014). Other PTMs comprise the attachment of bigger chemical moieties. For instance, ubiquitylation of histones is a reaction in which a 76 amino acid residue (ubiquitin) is added to lysines having roles in histone cross-talks and DNA repair (Meas and Mao, 2015). On the other hand, sumoylation acts as ubiquitylation counterpart, adding ubiquitin-like modifiers to lysines preventing the joining of ubiquitines (Bannister and Kouzarides, 2011).

1.1.2 Regulatory regions on the DNA sequence

As a second relevant linear aspect of the DNA I will mention its sequence. The concatenation of adenines (A), guanines (G), cytosines (C) and thymines (T) form the so-called **genetic code** (Crick et al., 1961). Apart from being the template for the generation of RNAs and proteins, this combination of A, G, C and T generates a plethora of **motifs** that will be recognized by TFs in an orchestrated spatial-temporal manner (Slattery et al., 2014).

Frequently along the genome, there are **clusters of DNA binding motifs**, these clusters usually function as regulatory regions for the gene expression programs. In the following subsections, I will explain the main features of the different regulatory regions that can be found on the DNA.

1.1.2.1 Gene promoters

Promoters are cis-regulatory regions whose primary role is the control of the transcriptional output rate by correctly positioning the transcriptional initiation complex. In eukaryotes, the RNA-polymerase II (RNAPII) is the enzyme responsible of RNA transcription for the heterogeneous gene content. In this regard, promoters serve as a highly dynamic layer of transcriptional regulation by differing in their architecture and accessibility (Lenhard et al., 2012). Core promoters contain the transcriptional start site (TSS), and in some cases a TATA-box. These genomic regions serve as binding platforms for proteins of the pre-initiation complex (PIC), a complex formed by general transcription factors (GTFs) and RNAPII (Figure 18). In order to assemble a sufficient PIC to transcribe from DNA, the minimal required elements are: TBP (TATA box binding protein), TFIIB, TFIIF, TFIIE and TFIIH (Luse, 2014). Nonetheless, eukaryotes are versatile systems due to the different key DNA motifs that lead to lineage-specific innovations in terms of transcription required proteins (Lenhard et al., 2012).

Even though the recruitment of the PIC to core promoters is one of the most rate-limiting steps of transcription, there are additional steps of regulation in initiation from promoters. One of these steps is the **RNAPII pause-release** that occurs at a 70% of metazoan genes. After recruitment and transcription of approximately 50 nucleotides (nt), RNAPII pauses and remains in this state until cellular cues trigger productive elongation (release) (Liu et al., 2015).

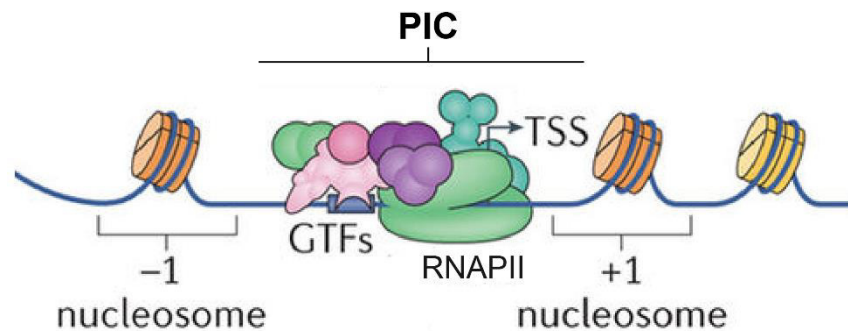


Figure 18. Schematic representation of the PIC bound to eukaryotic promoters. Adapted from (Soutourina, 2017).

1.1.2.2 Enhancers

The knowledge about enhancers has enormously increased in the last years, yet their definition remains the same as when the simian virus 40 (SV40) enhancer was discovered. Enhancers are DNA fragments with the ability to **positively drive target**

Introduction

gene expression independently of the genomic distance or orientation relative to their target gene promoter (Banerji et al., 1981).

During embryogenesis, the precise spatial-temporal control for gene expression has been shown to be mainly determined by cell-type specific enhancers. Enhancers have the capacity of being modularly regulated by transcription factors and signaling cascades in order to achieve different gene expression levels depending on the cellular context (Buecker and Wysocka, 2012).

1.1.2.2.1 Enhancer identification

Along the years, the identification of enhancers in a genome-wide manner was an arduous task due to the lack of general criteria to define them. With the forthcoming of the next generation sequencing (NGS) era and the use of chromatin immunoprecipitation coupled to sequencing (ChIP-seq), these cis-regulatory regions were discovered to display **specific chromatin features** that almost unequivocally identified bona fide enhancers (Heintzman et al., 2007) (Figure I9). Enhancers can be identified by using combinations of histone marks, cofactors, chromatin accessibility and DNA conservation data.

The first reports described that presence of H3K4me1 and absence or low levels of H3K4me3 was characteristic of enhancers (Heintzman et al., 2007; Koch et al., 2007), however, these histone marks do not describe their activity status. Later on, H3K27ac was attributed to active enhancers (Creyghton et al., 2010) (Figure I9) and H3K27me3 was discovered to be present in a ESCs specific set of inactive enhancers called poised enhancers (Rada-Iglesias et al., 2011).

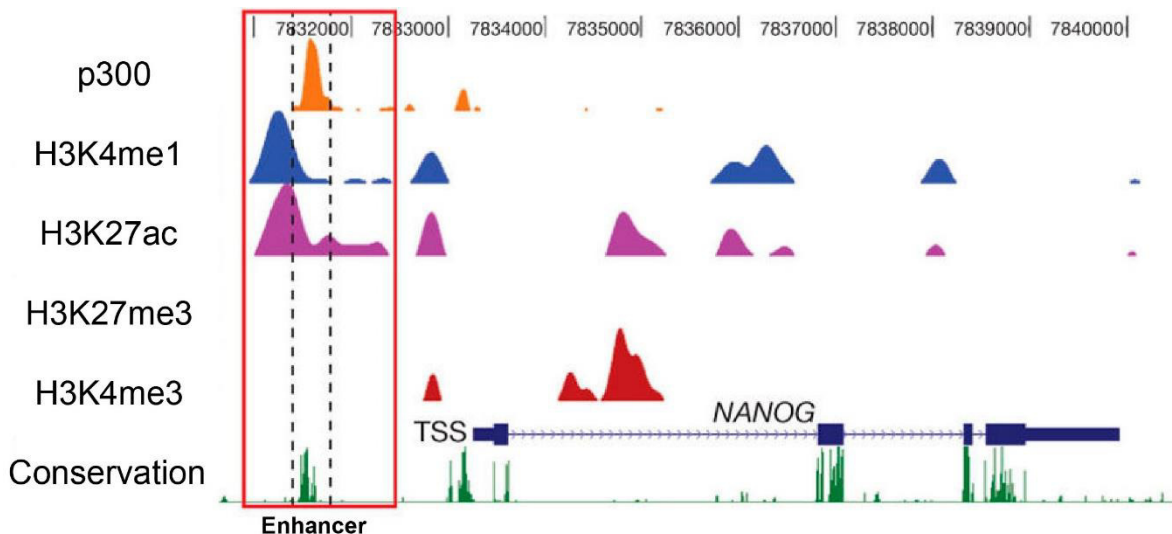


Figure I9. ChIP-seq capture displaying chromatin features of active enhancers. Adapted from (Rada-Iglesias et al., 2011).

As regards to other features and even though it has not been routinely used for enhancer identification, the presence of the HAT p300 has been shown to mark active and poised enhancers (Rada-Iglesias et al., 2011; Visel et al., 2009). In addition, low nucleosomal density and clusters of TF binding sites distally located from promoters have been demonstrated to be distinctive attributes of enhancer chromatin (Song et al., 2011; Zinzen et al., 2009).

1.1.2.2.2 Transcriptional activity at enhancers

In the year 2010, it was discovered that upon stimulation, enhancers were transcribed into nuclear unspliced bidirectional RNA species called **enhancer RNAs** (eRNAs) (Figure 110). This transcription was carried out by the RNAPII and was related to the activation of the putatively regulated genes (De Santa et al., 2010; Kim et al., 2010). Studies addressing the functional role of eRNA have been published. Interestingly, there are reports claiming that the eRNA molecule *per se* is essential for gene activation and enhancer-promoter looping (Lam et al., 2013; Li et al., 2013). Additionally, their high tissue specificity and widespread presence support the idea of the eRNAs as functional entities (Li et al., 2016).

Consequently, **eRNA transcription** has emerged as a hallmark of **enhancer activation**, and in this doctoral thesis it will be used to assess the activation status of the enhancers under study.

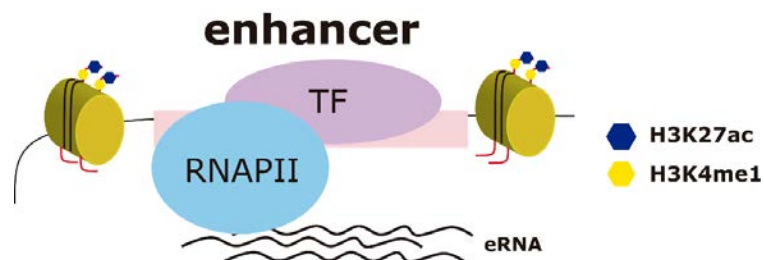


Figure 110. Schematic representation of the enhancer RNA (eRNA) transcription.

1.1.2.2.3 Superenhancers

The ability of enhancers to cooperate in order to increase gene activation in a more efficient manner has been described. **Superenhancers** are large clusters of transcriptional enhancers that drive the expression of genes that define cell identity. These large units of regulation are usually bound by lineage specific transcription factors and the Mediator complex. Furthermore, superenhancers increase gene expression in a higher level than regular enhancers (Whyte et al., 2013). Remarkably, superenhancers have been found to be regions enriched in disease-associated sequence variations and to activate essential tumorigenic genes (Hnisz et al., 2013).

Introduction

1.1.2.3 Silencers and insulators

Silencers are genetic entities that suppress gene expression. In the case of insulators, they confine it within chromatin boundaries (Kolovos et al., 2012). Curiously, silencers share many properties with enhancers (e.g. orientation- and distance-independent activity), however, they are binding sites for repressors and the chromatin state that they display avoids gene activation (Maston et al., 2006).

1.1.3 Proteins that act on the linear chromatin

In the previous sections, I have explained how the chemical modification of histones and the DNA binding motifs present in the DNA fiber configure the final gene expression output. In this section I will detailed the types of proteins that exert their roles on chromatin.

1.1.3.1 Transcription factors

The complex eukaryotic transcriptional events carried out by RNA-polymerases require a set of accessory factors that recognize promoters and enhancers to **accurately initiate transcription** (Matsui et al., 1980). In this context, TFs interpret the regulatory DNA within a genome by recognizing its sequence in a specific manner.

High throughput approaches like ChIP-Seq and structural crystallographic experiments have shed light into the *in vitro* and *in vivo* binding preferences of TFs, concluding that TFs do not bind to all their predicted sites. It is known, that other features different than just the DNA sequence affect TFs binding and recognition: chromatin state, the presence of cofactors, the necessity of TF cooperative binding and others are some examples (Slattery et al., 2014).

For this doctoral thesis, the TF SMAD3, effector of the TGF β pathway will be of special relevance. It will be detailed in the TGF β subsection 3.4.

1.1.3.1.1 Pioneer transcription factors

Nuclear DNA is not always accessible, its packaging around nucleosomes forming higher-order structures and the binding of repressor complexes sometimes make TFs binding sites unavailable for recognition. In this situation, pioneer transcription factors are proteins that have the ability of **binding compact chromatin** and by opening it up permit transcription initiation of silent regions (Zaret and Carroll, 2011). In the last years, pioneer transcription factors have acquired much attention due to their participation in reprogramming and cell lineage determination (Iwafuchi-Doi and Zaret, 2014). Along the experimental performance of this doctoral thesis, the pioneer

transcription factor ASCL1 (Achaete-scute homolog 1) has been essential and its main characteristics will be summarized in the following subsection.

1.1.3.1.1 The pioneer factor ASCL1

The TF ASCL1 belongs to the basic helix-loop-helix (bHLH) family of TFs, a family with quintessential roles in mammalian **neurogenesis**. ASCL1 is a proneural factor with roles in proliferating and differentiating neural progenitors and has been demonstrated to be both necessary and sufficient to promote the generation of neurons (Wilkinson et al., 2013). Even though it is unknown how ASCL1 exerts its pioneering activity, in the reprogramming of fibroblasts to neurons, overexpression of ASCL1 leads to its binding into cognate neural progenitor binding sites that lie on closed chromatin (Wapinski et al., 2013). Besides, and more relevant for this thesis, ASCL1 has been shown to recognize its binding sites in close chromatin in NSCs and promote this chromatin decondensation and gene activation (Raposo et al., 2015) (Figure I11).

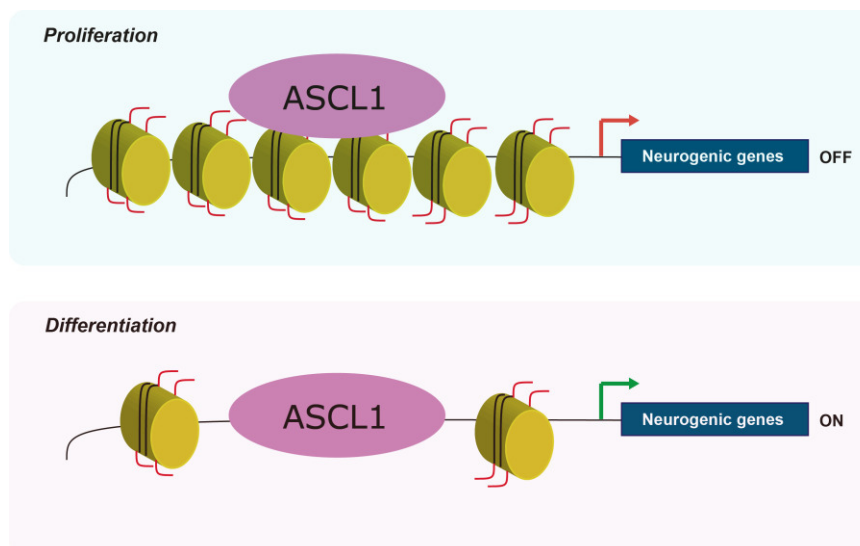


Figure I11. Schematic representation of ASCL1 pioneering activity in NSCs. Figure based in (Raposo et al., 2015).

1.1.3.2 Transcriptional coactivators

The knowledge of how activator TFs played their role led to the discovery of coactivators. Transcriptional coactivators are a diverse array of proteins that connect sequence-specific DNA binding activators to the general transcriptional machinery, thus cooperating with transcriptional activation. Accordingly, many chromatin modifying enzymes act as coactivators. This doctoral thesis addresses the role of

Introduction

the coactivator JMJD3 (explained in subsection 1.1.1.2.1.1) and of a coactivator with remodeling activity: CHD8. In the next paragraph, I will introduce the chromatin remodeler CHD8.

1.1.3.2.1 The chromatin remodeler CHD8

One of the most studied families of chromatin remodelers is the chromodomain helicase DNA-binding family of proteins (CHD). This family contains 9 members with different affinities and activities. For this thesis, I will focus on one of the members: CHD8.

The coactivator CHD8 belongs to a family of **ATP-dependent remodeling** enzymes, that are characterized for presenting a chromo domain that binds H3K4me2 and H3K4me3 histone marks and a helicase-ATPase domain that disrupts or promotes the contact between DNA and the histone octamer (Micucci et al., 2015)(Figure I12). CHD8 has been related to cell proliferation and has been found on gene promoters of cell cycle regulators interacting with the elongating form of the RNAPII (Rodriguez-Paredes et al., 2009).

Additionally, CHD8 is highly expressed in embryonic cortex and exosome sequencing experiments have found that *de novo* mutations on the *Chd8* gene are related to autism spectrum disorders (Katayama et al., 2016). *In vivo* knockdown (KD) of this protein during cortical development results in defective neural progenitor proliferation and differentiation, and ultimately generates abnormal neuronal morphology and behaviors in adult mice (Durak et al., 2016).

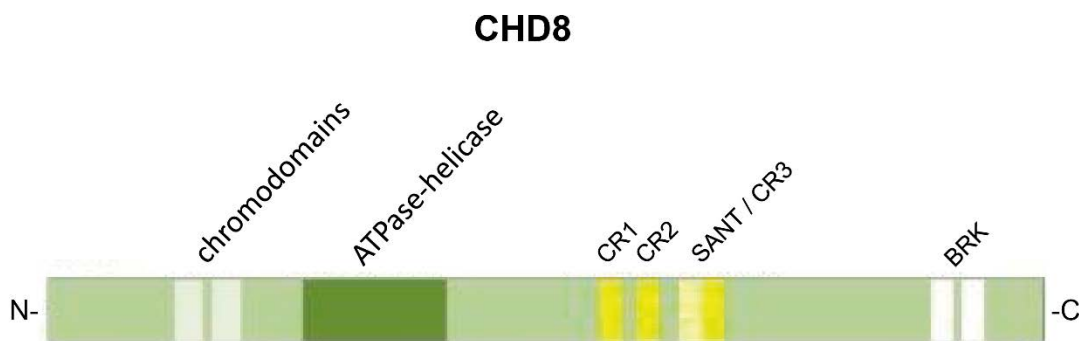


Figure I12. Schematic representation of the CHD8 chromatin remodeler domains. Adapted from (Batsukh et al., 2012).

1.1.3.3 RNA-polymerase II

All the aforementioned proteins operate on chromatin in order to support transcription, this process is carried out by RNA-polymerases that are protein complexes that also use the chromatin as substrate.

The transcription consists in the attachment of the RNA-polymerases to the DNA template in order to catalyze the **production of complementary RNA**. In eukaryotes, there exist three RNA-polymerases, but in this subsection I will focus on RNAPII, responsible of the mRNA and eRNA transcription (Sims et al., 2004).

The RNAPII is a multiprotein complex composed by 12 subunits in humans. These subunits are denominated RBPs and the one that contains the polymerase enzymatic core is RBP1. Emerging from RBP1 subunit there are tandem repeats (52 in humans) of the consensus sequence YSPTSPS that form the carboxy-terminal domain (CTD). This CTD is not necessary for the RNA-polymerase activity. However, all the steps of transcription and many co-transcriptional processes such as splicing and chromatin modification are controlled by these tandem repeats (Harlen and Churchman, 2017).

1.2 *Tridimensional components and aspects of chromatin*

In the previous paragraphs, I have detailed the linear chromatin components, and how their structure and chemical environment dictate gene transcriptional programs. In this section, I will focus on the tridimensional (3D) scope of the chromatin.

Every *Homo sapiens* cell contains approximately two meters of DNA packaged inside the nucleus. This huge amount of DNA can fit in such a small organelle due to its folding around histones and chaperones (Bednar et al., 1998).

The primary structure of the chromatin is spontaneously formed in solution under DNA native conditions. At low salt levels, a chain of nucleosomes appears like “beads on a string” or uncondensed 10 nm chromatin fibers. As the concentration of salt increases, up to 200 nucleosomes associate to form the secondary structure of chromatin, either as zigzag fibers or as solenoid fibers (30nm). The tertiary chromatin structure is shaped when the 30 nm fiber coils into sequential higher-order stages of condensation, folding and compaction giving rise to the metaphase nucleosomes. (Bascom and Schlick, 2017; Luger et al., 2012).

1.2.1 **Levels of 3D-chromatin organization**

With the development of technological advances, the understanding on how the chromatin fiber is arranged in the 3D-structure is starting to be uncovered. Due to their special relevance for this doctoral thesis, I will compile the different levels of 3D chromatin structure and their characteristics.

1.2.1.1 **Chromosome territories**

The first model accounting for how chromosomes are organized in non-dividing cells was proposed by Carl Rabl in 1885. This model claimed, that the DNA of each

Introduction

chromosome occupies a defined volume of the nucleus and only overlaps with its immediate neighbors. One century later, Thomas and Christoph Cremer empirically demonstrated this model by using DNA damage experiments, and later on, FISH (fluorescence in situ hybridization) technique confirmed the results (Cremer et al., 1982). Interestingly, the territories occupied by each chromosome are not random (Figure I14), some chromosomes are most likely located at the center of the nucleus, while others are located at the nuclear periphery. Moreover, some chromosomes are usually clustered together in a specie and cell-type specific manner. The chromosome location at the nucleus will dictate their gene expression program and many pathological processes arise from chromosome mislocalization and translocation (Lanctot et al., 2007; Parada et al., 2002).

1.2.1.2 Topological associated domains

In 2012, chromatin conformation capture techniques, and concretely, its most high throughput version, Hi-C, allowed the advancement in the understanding of 3D-chromatin structure and gene expression. Dr. Bing Ren et al. defined megabase-sized local chromatin interaction domains as topological associated domains (**TADs**) (Figure I14). These domains are conserved across species and are very stable across different cell types, indicating that their structure is intrinsic to mammalian genomes (Dixon et al., 2012). In *Drosophila*, TADs significantly correlate with epigenetic marks, and the interacting domains overlap with the previously described *Drosophila* chromatin types (Filion et al., 2010; Sexton et al., 2012). However, on mammalian genomes the relation between epigenetic marks and TADs is not so direct. TAD borders operate as insulators avoiding the spreading of H3K9me3 heterochromatic domains, but their relation with other epigenetic marks is not as clear as in the *Drosophila* genome (Dixon et al., 2012; Ea et al., 2015).

The genomic regions between TADs are named **TAD boundaries** and are regions enriched in housekeeping genes, active chromatin marks (H3K4me3, H3K36me3), short interspersed elements (SINE) retrotransposons, transfer RNAs (tRNAs) and binding motifs for insulator proteins such as CTCF (Dixon et al., 2012). At the functional level, TADs have been proposed to be the basic unit of genome folding as the genes that they contain are coregulated and share coordinated gene expression profiles across cells and tissues (Dixon et al., 2016; Nora et al., 2012). Additionally, chromatin loops between promoters and enhancers occur almost exclusively within TADs (Jin et al., 2013).

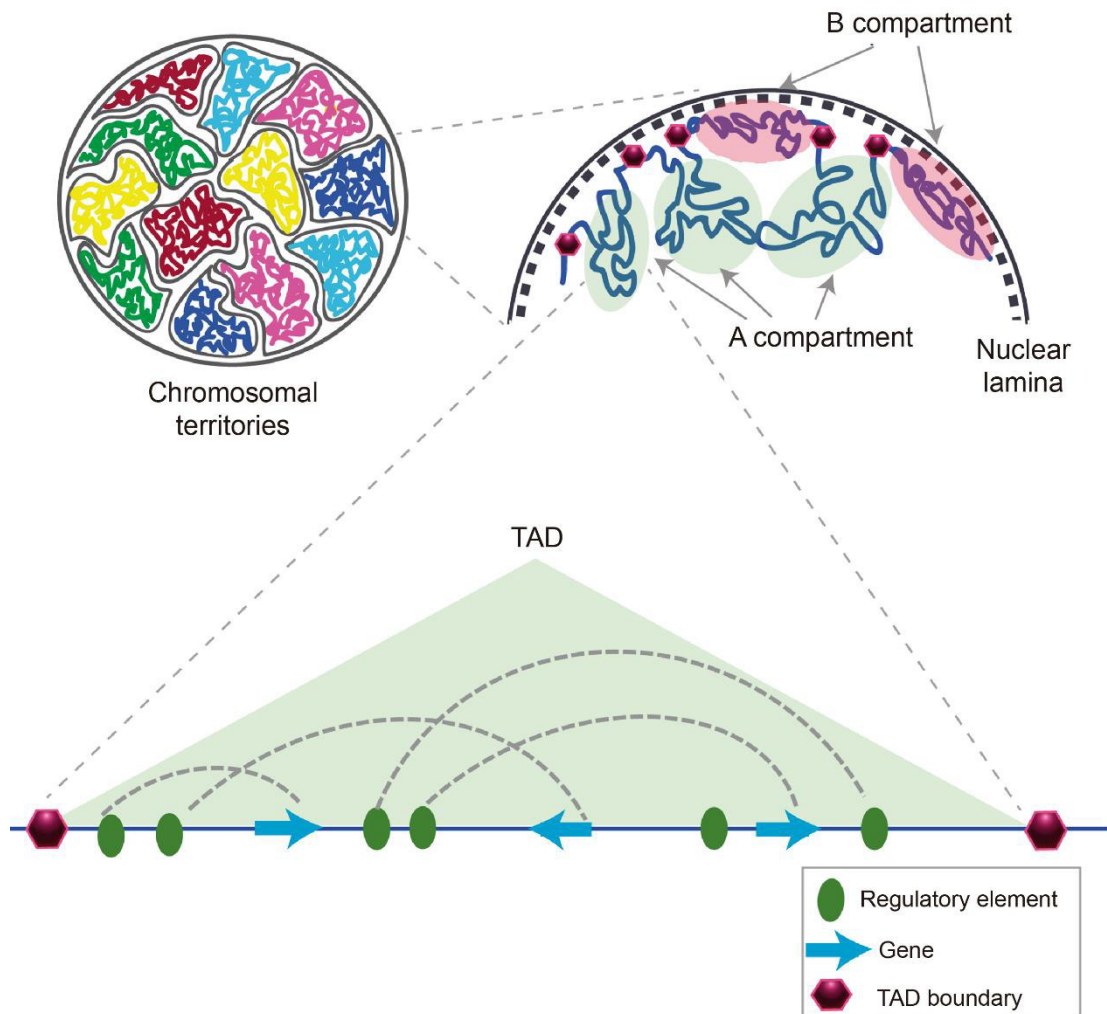


Figure I14. Schematic representation of different levels of tridimensional chromatin organization. Adapted from (Matharu and Ahituv, 2015).

1.2.1.3 Chromosomal compartments

Experiments of Hi-C have confirmed that apart from chromosomal territories and TADs, there is another organization level for chromatin into two compartments. **Compartment A**, that is transcriptionally active and **compartment B** that is transcriptionally inactive (Figure I14). In contrast to TADs, chromosomal compartments change along development and across cell types (Dixon et al., 2015; Lieberman-Aiden et al., 2009).

Introduction

1.2.1.4 Loop domains

After the characterization of TADs, a static view of the genomic contacts was briefly established. However, the improvement in the resolution of chromatin conformation capture experiments led to the delineation of genomic contacts as much more dynamic than expected. In 2014, a new version of Hi-C protocol achieved 1 Kb of resolution between genomic contacts. This great resolution showed that genomes are partitioned into contact or **loop domains** of a median length of 185 Kb conserved across species and cell types. In contrast to TADs, loop domains have characteristic epigenetic marks, and changes on these chromatin features lead to changes on the long-range contact pattern. Loop domains are generated by the anchorage of domain borders that contain CTCF proteins in a convergent fashion and Cohesin complex molecules, furthermore, they often bring together enhancers and promoters thus leading to an increase on gene expression (Rao et al., 2014) (Figure I15).

Single-cell analysis and microscopy experiments have recently demonstrated that both TADs and loop domains are **highly dynamic** (Hansen et al., 2018; Hansen et al., 2017). During mitosis, TADs and loop domains are dismantled and later, on G1 phase they are again organized (Nagano et al., 2017). How loop domains are established is still under study, but some of the machinery participating in this event is starting to be uncovered.

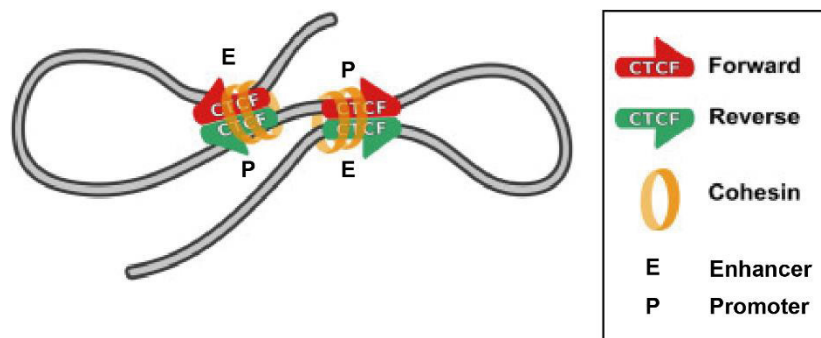


Figure I15. Picture displaying convergent CTCF sites anchoring enhancer-promoter loops. Adapted from (de Wit et al., 2015).

1.2.1.4.1 Loop maintenance complex of proteins

The previously mentioned levels of 3D-chromatin organization are maintained and formed with the participation of the loop maintenance complex of proteins (LMC), composed of **CTCF and Cohesin**.

After the definition of TADs and their boundaries, motif discovery analysis and ChIP-seq demonstrated that the insulator protein CTCF and the components of the Cohesin complex occupied TAD borders (Dixon et al., 2012; Nora et al., 2012). CTCF is an evolutionary conserved phosphoprotein that binds to DNA through the

combinatorial usage of eleven zinc fingers that are flanked by N-terminal and C-terminal domains whose role is not yet understood (Ohlsson et al., 2001). Importantly, as the CTCF binding motif is not palindromic, the orientation of CTCF binding is not random and it has been observed that chromatin loops are generated between convergent CTCF molecules (Rao et al., 2014) (Figure I15).

On the other hand, Cohesin complex is formed by SMC1, SMC3, RAD21 and SA1/2. All these components generate a ring-shaped structure inside which DNA is trapped allowing sister chromatid cohesion, homologous recombination and loop domain formation (Haarhuis et al., 2014; Ivanov and Nasmyth, 2005).

The role that these proteins exert in loop domain formation is starting to be revealed. In this time and age, the most supported model is the so-called **loop extrusion model**. In this model, Cohesin is recruited to a CTCF site and starts extruding a DNA loop until it encounters a CTCF molecule in the convergent orientation, forming a loop domain (Hansen et al., 2018; Sanborn et al., 2015).

2 The cortex development as a model of study

In this second section of the introduction of this doctoral thesis, I will cover the essential aspects of corticogenesis, in order to properly introduce my model of study, the NSCs.

2.1 Definition, function and development of the cortex

The mammalian neocortex is the region of the brain responsible of the **cognitive function**, consciousness and sensory perception. In humans it is believed to provide the uniqueness of our specie, even though the forebrain structure is conserved in all vertebrates (Wilson and Houart, 2004). At the anatomical level, the cortex is an extremely organized six-layered structure that comprises many different neuronal subtypes and glial cells (Molyneaux et al., 2007) (Figure I16).

After egg fertilization, the blastocyst divides into the three primary layers: ectoderm, mesoderm and endoderm. The **ectoderm** is the germ layer that will give rise to the nervous system (Patthey and Gunhaga, 2014). After gastrulation, the dorsal region of the ectoderm will flatten in order to generate the neural plate. Next, the neural plate will roll and will

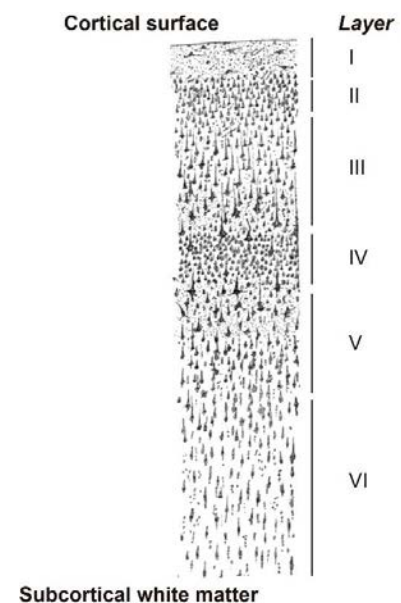


Figure I16. Drawing of cortical lamination by Santiago Ramon y Cajal. Adapted from Cerebral cortex entry in Wikipedia

Introduction

form the neural tube, whose anterior and posterior region will generate the brain and the spinal cord respectively (Ozair et al., 2013).

Cortical neurogenesis begins at gestational week 8th in humans, and at day 11 (E11) in mice embryos. This process is dependent on intrinsic and extrinsic signals that control the generation of neurons in a spatial-temporal manner. Cortical progenitors populate the proliferative zones of the dorsal telencephalon, the ventricular and the subventricular zones (VZ and SVZ respectively).

The neurogenic process begins when neuroepithelial progenitors (NE) appear, and by proliferative divisions amplify the pool of cortical progenitors. Later on, NE convert into **radial glial cells** (RG), the main subtype of cortical progenitors that will differentiate in a diverse type of neurons in order to generate the cortical structure (Tiberi et al., 2012). Neurons are generated in the deeper part of the developing brain, the VZ. Next, they migrate to the pial surface so that they generate the six-layered cortical architecture. The last step in neurogenesis is the neuron terminal translocation to the marginal zone (MZ) where it will mature developing axons and dendrites (Ohtaka-Maruyama and Okado, 2015) (Figure I17).

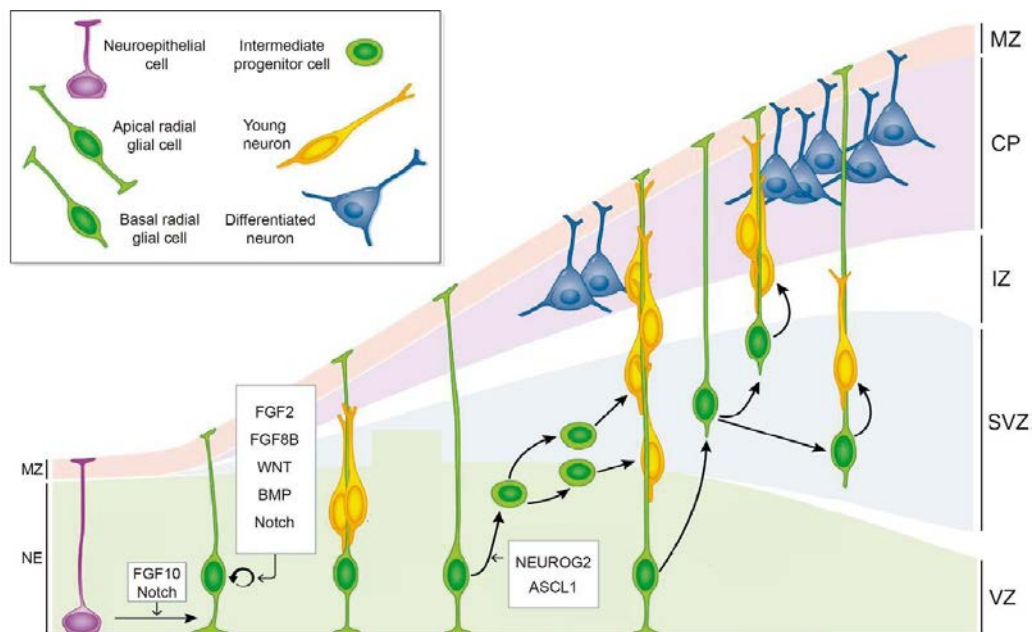


Figure I17. The drawing depicts types of progenitor cells and their lineage relationships in the developing cerebral cortex. Molecules or pathways regulating the indicated step are included in white boxes. Adapted from (Fernandez et al., 2016).

2.1.1 Developmental cues in corticogenesis

As I previously mentioned, extrinsic and intrinsic cues control neurogenesis. The decision about proliferating or differentiating is governed by signaling pathways operating in the neurogenic niche (extrinsic factors) and by proneural TF, cell polarity

and mitotic molecules functioning in a cell-autonomous manner (intrinsic factors) (Tiberi et al., 2012).

2.1.1.1 Extrinsic factors

Along neurogenesis, different signaling pathways governed by growth factors (GF) act upon cells to orchestrate the balance between proliferation and differentiation. **Fibroblast growth factor** (FGF) signaling influences the extent and level of proliferation of NSCs *in vitro* and *in vivo* (Ornitz and Itoh, 2015). **Sonic Hedhog** (Shh) pathway controls cell cycle kinetics of RG, thereby maintaining the proliferation, survival and differentiation of neurons in the neocortex (Komada, 2012). The canonical **Wnt** pathway stimulates the proliferation of the VZ neural precursors and promotes neuron fate (Chenn and Walsh, 2002; Mutch et al., 2009). From the TGF β superfamily of receptors, the bone morphogenetic proteins (**BMP**) and **TGF β** participate in forebrain development. BMPs sequentially induce neurogenesis and then gliogenesis in a NEUROG1 TF-dependent mechanism (Li et al., 1998). On the other hand, the TGF β pathway, promotes cell cycle exit of VZ progenitors by modulating the cell cycle proteins (Siegenthaler and Miller, 2005). Moreover, in cortical and hippocampal progenitors TGF β has been demonstrated to promote neuronal cell fate (Vogel et al., 2010).

2.1.1.2 Intrinsic factors

Intrinsic factors balancing proliferation versus differentiation operate mainly through a cell-autonomous manner. Some of the essential intrinsic factors are the bHLH proteins. The class II of bHLH proteins contains the **proneural TFs**, key determinants of neural cell fate. Proneural TFs are necessary and sufficient to activate neuronal differentiation programs, and in the central nervous system their expression is confined into specific cell populations (Wilkinson et al., 2013).

Members of the proneural TFs are ASCL1, the Neurogenins (NEUROG1, NEUROG2 and NEUROG3), the neurogenic differentiation proteins (NEUROD1, NEUROD2, NEUROD4 and NEUROD6) and oligodendrocyte differentiation proteins (OLIG1, OLIG2 and OLIG3) (Ross et al., 2003).

Other intrinsic factor is the **Notch pathway**, its ligand DLL1 and effector HES1 play an important role for the maintenance of the self-renewal in NSCs. HES1 expression fluctuates in NSCs and regulates ASCL1, NEUROG2 and DLL1 oscillations. Conversely, the Notch ligand DLL1 leads to maintenance of NSCs by mutual activation of Notch pathway (Shimojo et al., 2008). At the moment of cell fate choice, the oscillatory expression of a cell type specific factor becomes a sustained expression, leading to the generation of neurons if the expressed factors are

Introduction

NEUROG2 or ASCL1 and to oligodendrocytes if the factor is OLIG2 (Imayoshi et al., 2013) (Figure I18).

Other intrinsic factors affecting the corticogenesis are the **cell polarity** (Taverna et al., 2014) and the different **modes of division** of the cells (Tiberi et al., 2012), however, these will not be detailed for they are out of the scope of this doctoral thesis.

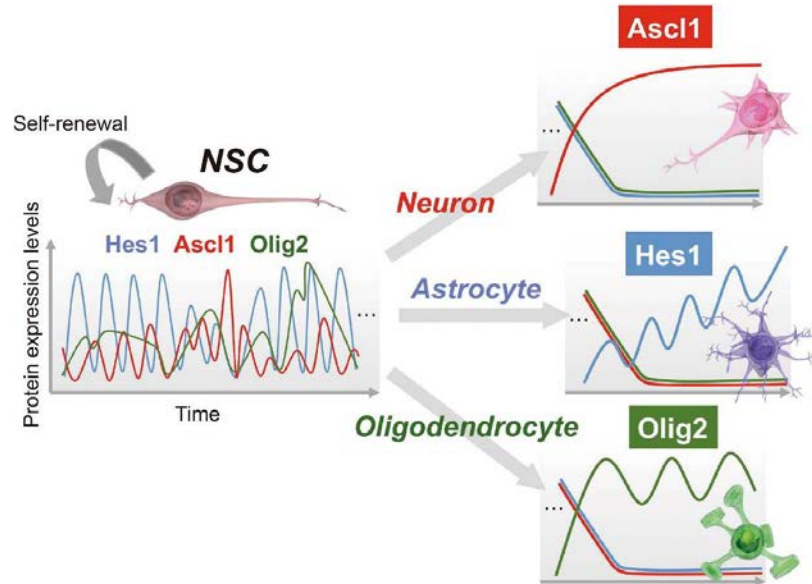


Figure I18. Schematic representation of the oscillatory expression versus the sustained expression of proneural factors in NSCs. Adapted from (Imayoshi and Kageyama, 2014)..

2.2 Neural stem cells as an *in vitro* model of study

In the experimental work of this thesis, NSCs from mice embryo cortices have been used as a model of study. As it will be explained in the Methodology section, this model is an *in vitro* system that allow us to molecularly characterize the gene expression program triggered by the TGF β pathway. It is worth noting, that vertebrate neurogenesis is an asynchronized series of events, and the study of specific mechanisms involved in transcriptional regulation is very difficult in the developing embryo. For this reason, I have used adherent cultures of NSCs, thus avoiding the higher heterogeneity of the *in vivo* systems.

2.2.1 Neural stem cells *in vitro* versus *in vivo*

NSCs are multipotent cells in the nervous system that can self-renew, proliferate in an unlimited manner and differentiate into neurons, astrocytes and oligodendrocytes. The difference between NSCs and progenitor cells is the ability of the formers of **self-renew**. Progenitors are often unipotent and their proliferative capacity is limited, thus progenitors are different from NSCs (Seaberg and van der Kooy, 2003). Notably, our *in vitro* system of NSCs displays morphology and markers similar to those of the RG in the embryo (BLBP, RC2, GLAST, PAX6 and NESTIN among others) (Pollard et al., 2006) (Table I3).

	Developmental stage	Specific markers	Rosette formation	Anteroposterior identity	Dorsoventral identity
Radial glial cells	Between E10.5 and E12.5	NESTIN, RC2, SOX2, BLBP, GLAST and PAX6	No	Forebrain	Dorsal
NSCs in this thesis	E12.5	At least NESTIN, SOX2 and PAX6	No	Forebrain	Dorsal

Table I3. This table summarizes the main characteristics of radial glial cells in comparison to the neural stem cells used in this doctoral thesis. Adapted from (Conti and Cattaneo, 2010).

Establishing and maintaining NSCs in culture is achieved by the addition of reagents to the cell plates in order to mimic their embryo origin. This artificial set up imposes some **limitations** to the NSCs as a neurogenesis model.

In vivo, NSCs occupy a microenvironment called niche that supplies the factors required for stem cell self-renewal and differentiation. On the contrary, in the adherent *in vitro* system mitogens such as FGF and EGF (epidermal growth factor) force the symmetric divisions of the NSCs retaining their multipotency, thus reducing the differentiating counterpart (Curre et al., 2007). Another important feature of NSCs maintained *in vitro*, is their loss of positional identity. *In vivo*, NSCs expressing specific bHLH TFs are characteristic of concrete regions along brain development; however, the long-term expansion of NSCs leads to the uniformity in cellular markers. Lastly, one important limitation of the NSCs culture is the limited assortment of neuronal subtypes that can be generated upon differentiation in comparison to the embryonic neurogenesis. Furthermore, the lack of functional maturation in electrophysiological terms restricts the aspects that can be studied *in vitro* (Conti and Cattaneo, 2010).

Introduction

3 The TGF β pathway

Along this last section of the introduction, I will describe the main features of the TGF β signaling cascade. Induction of the TGF β pathway has served not only as a model to study the TGF β components, but as a stimulation mechanism in order to analyze the transcriptional events that emerge from it.

3.1 The TGF β superfamily of ligands

In the living organisms, intercellular communication is achieved by the secretion and reception of messenger molecules that operate in a specific time and space. The TGF β superfamily of ligands participate in these communication events.

This superfamily is comprised of over 30 members, including Activins, Nodals, BMPs, TGF β s and Growth and Differentiation Factors. Members of this superfamily are found in metazoan species and are ubiquitously expressed in embryonic tissues and adulthood. For this reason, misregulation of any of these superfamily components leads to pathological mechanisms and disease. (Gordon and Blobel, 2008).

All the members belonging to this superfamily of ligands trigger a universal cascade that consists in the binding of the ligand to serine/threonine kinase receptors. These receptors phosphorylate SMAD proteins, that accumulate in the cell nucleus and regulate specific transcriptional programs (Weiss and Attisano, 2013) (Figure 119).

In the next subsections, I will exclusively focus on the TGF β pathway triggered by the cytokine TGF β 1 (from here on, TGF β) and its structural and functional aspects.

3.2 Definition and characteristics of the TGF β pathway

The TGF β pathway is a signaling cascade that regulates proliferation, differentiation, morphogenesis, regeneration, apoptosis and cell homeostasis. It is worth noting, that the regulation of different targets depending on the cell context is achieved due to the ability of the SMAD proteins of cooperating with different partners that endow TGF β signaling with a **versatile specificity** (Massagué, 2012).

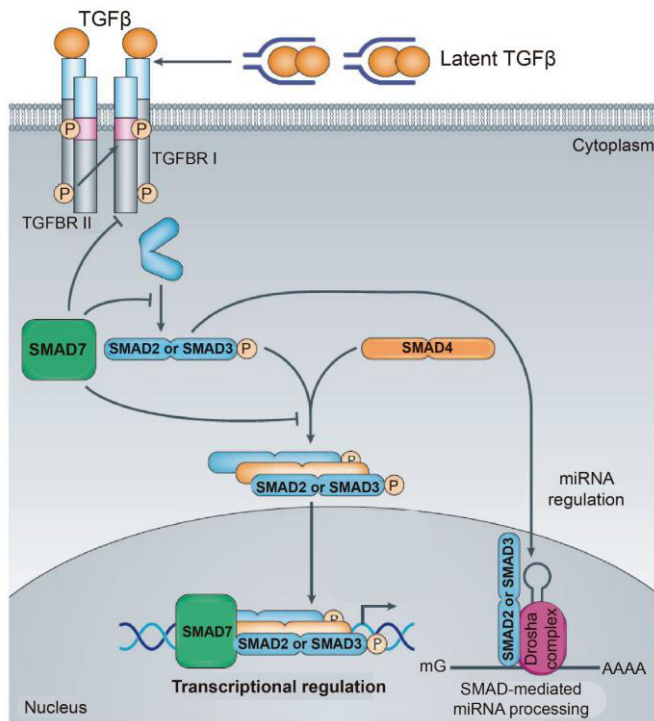


Figure I19. Schematic representation of the TGFβ signaling cascade. P: phosphate group. Adapted from (Akhurst and Hata, 2012).

3.3 TGFβ pathway signaling cascade

As I mentioned before, the cytokines belonging to the TGFβ superfamily of ligands exert their functions through a universal cellular cascade that differs between members of the superfamily in the type of membrane receptors that are targeted, and the SMAD effectors that function in every moment.

The active TGFβ cytokine is a dimer stabilized by hydrophobic interactions that are further strengthened by disulfide bridges. Each TGFβ monomer is formed by β-strands interlocked by three conserved disulfide bonds that form the so-called cysteine-knot. This dimerization of the TGFβ cytokine facilitates the interaction with **TGFβ receptors** (TGFBRs) that are also in a dimeric arrangement (Sun and Davies, 1995). Structurally, TGFBRs are proteins with an extracellular N-terminal domain that serves as ligand binding platforms, a transmembrane domain and an intracellular C-terminal domain containing serine/threonine kinase activity. There are two types of TGFBRs, type I and type II, but TGFβ has uniquely the ability of binding to type II receptors, increasing the affinity of type II for type I. Upon ligand binding, type II TGFBR that contains a kinase domain, phosphorylates type I receptor and the latter propagates the signal by phosphorylating the SMAD proteins. Concretely, in the TGFβ pathway, the cytoplasmic **SMAD2 and SMAD3** are phosphorylated by the type II TGFBR and by joining with SMAD4 enter the cell nucleus and bind along the genome using the CAGA-boxes also called SMAD binding elements (SBE) (Shi and Massague, 2003) (Figure I19).

Introduction

3.4 SMAD proteins as the effectors of the TGF β pathway

In this subsection I will deepen into the TFs that govern the TGF β pathway, the **SMADs**. TGF β stimulates the membrane receptors that directly phosphorylate SMAD proteins. Then they form transcriptional complexes to control gene expression. The versatile functionality of this signaling pathway is achieved due to the regulation of different targets depending on the cell context.

3.4.1 Types of SMAD proteins

Humans and mice encode eight SMAD proteins, SMAD 1-8. *Drosophila* species encode four and *Caenorhabditis elegans* three (Massague et al., 2005). Attending to their role in the TGF β pathway there are three types of SMADs (Figure I20). The receptor SMADs or **R-SMADs** are substrates of the TGFBRs, in the case of the TGF β signaling these are SMAD2 and SMAD3. As a second type of SMADs, there is just SMAD4. SMAD4 is the so-called **Co-SMAD** and contributes to the signaling by associating with SMAD2 and SMAD3 even though it is not phosphorylated. The last type of SMADs are the inhibitory SMADs, **I-SMADs**. To this group belong SMAD6 and SMAD7, but just SMAD7 has an inhibitory effect over the TGF β pathway. I-SMADs inhibit TGF β signaling through several mechanisms: negative feedback regulatory loops, antagonizing the phosphorylation of SMAD2/3, occupation of the SBE along the genome, and others (Massague, 1998; Yan et al., 2009).

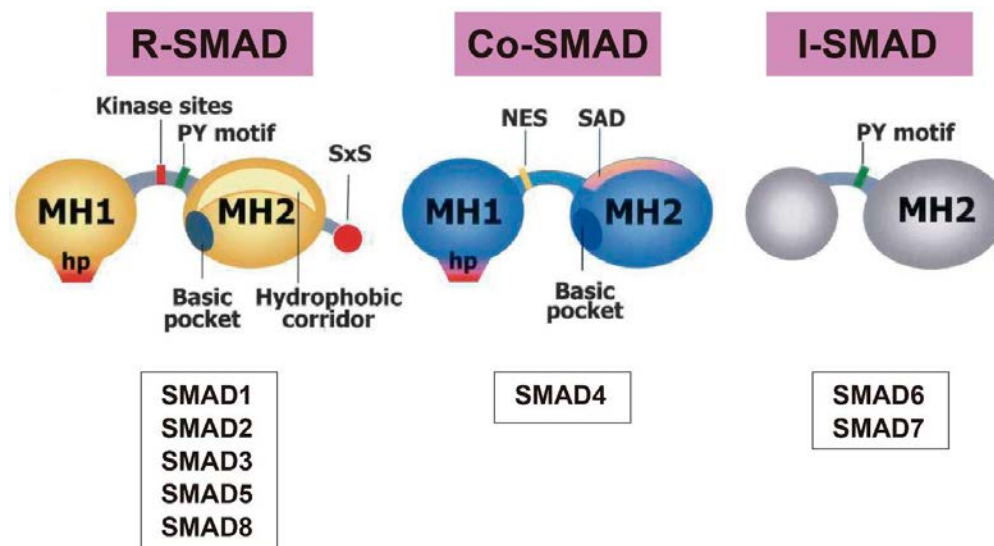


Figure I20. Cartoon model of the different types of SMADs and their structure. Adapted from (Massague et al., 2005).

3.4.2 Structural aspects of the SMAD proteins

SMADs contain around 500 amino acids. Structurally, they are formed by two globular domains and a linker region. The **MH1** globular domain is located at the N-terminal region and the **MH2** globular domain occupies the C-terminal part. The MH1 is conserved in all the SMADs excepting for the inhibitory, the **linker** region is divergent in all SMADs and the MH2 globular domain is conserved (Massague et al., 2005; Shi and Massague, 2003) (Figure I20).

At the functional level, the MH1 globular domain of the R-SMADs and SMAD4 has been demonstrated to have double strand DNA binding activity, thus binding to the SBE. On the other hand, the MH2 globular domain mediates protein-protein interactions between the R-SMADs and TGFBRs, Co-SMADs, DNA-binding cofactors, epigenetic regulators and others. The linker region functions as a substrate for PTMs and participates in the stability of SMAD proteins (Macias et al., 2015; Massague et al., 2005).

3.4.3 Functions of the SMAD proteins in the chromatin

As a starting point in this subsection, I would begin by remarking that SBEs are very degenerate, and they can often be found at the genomic sequence. This fact leads to the requirement of cofactors in order to proper target SMADs to their genomic binding regions.

It has been described that **master transcription factors** target SMAD3 to their binding regions in ESCs, myoblasts and hematopoietic lineages (Mullen et al., 2011). However, in this subsection I would like to emphasize the cooperation between SMAD3 and cofactors and/or epigenetic regulators.

One of the most studied cofactors of SMAD3 is the HAT p300/CBP. SMAD3 recruits and binds to **p300/CBP**, and this leads to transcriptional activation (Janknecht et al., 1998). Another example of SMAD3 cofactors are the members of the SWI/SNF remodeler complexes, specially **BRG1**. By recruiting BRG1, SMAD3 promotes the remodeling of the TGF β target promoters and allows gene activation (Xi et al., 2008). Finally, I would like to introduce here the cooperation between SMAD3 and the epigenetic regulator **JMJD3**. The coordinated activity between these two proteins have been observed in endodermal differentiation, in epithelial to mesenchymal transition (EMT) and most relevant for this thesis in neural development (Akizu et al., 2010; Estaras et al., 2012; Estaras et al., 2013; Kartikasari et al., 2013; Ramadoss et al., 2012). In the 3.6 section, the cooperation between SMAD3 and JMJD3 in neural development will be deeply covered.

Introduction

3.5 Functions of the TGF β pathway

As it was introduced, the TGF β pathway exerts many physiological roles along development and in adulthood. Nonetheless, over the years, many studies have addressed the role of TGF β pathway in diseases such as cancer. In this section I will describe how TGF β participates in these cellular processes.

3.5.1 Physiological functions of the TGF β pathway

Explanation of the high number of functions of the TGF β pathway in vertebrates is out of the scope of this doctoral thesis that is framed in the neural development context. Nonetheless, in this subsection I will briefly summarize the most prototypical functions of this cytokine.

3.5.1.1 Role of the TGF β pathway in proliferation

The TGF β cytokine has long been demonstrated to have cytostatic capacity. This **antiproliferative** activity has been widely studied in epithelial cells due to its relation to cancer. Nonetheless, TGF β concentration and cell environment can lead to antagonistic effects.

In normal epithelial cells, TGF β regulates the expression of genes promoting cell cycle arrest. Upon stimulation, TGF β activates two cyclin-dependent kinase inhibitors, *p21/Cip1* and *p15Ink4b*, and represses *Myc*, *Id1* and *Id2*, TFs involved in proliferation and inhibition of differentiation. Altogether, this transcriptional program has a tumor suppressor effect avoiding aberrant cell proliferation (Seoane, 2006) (Figure I21). In non-epithelial cells, TGF β has an essential role controlling T-lymphocytes growth and functions through several mechanisms. *Tgfb1* KO mice die shortly after birth due to massive inflammatory response, presumably due to the lack of TGF β inhibitory effects (Shull et al., 1992).

3.5.1.2 Role of the TGF β pathway in differentiation

The regulation of cell identity by the TGF β pathway often entails the participation of lineage specific TFs that provide specificity to the promiscuous TGF β signaling (David and Massague, 2018).

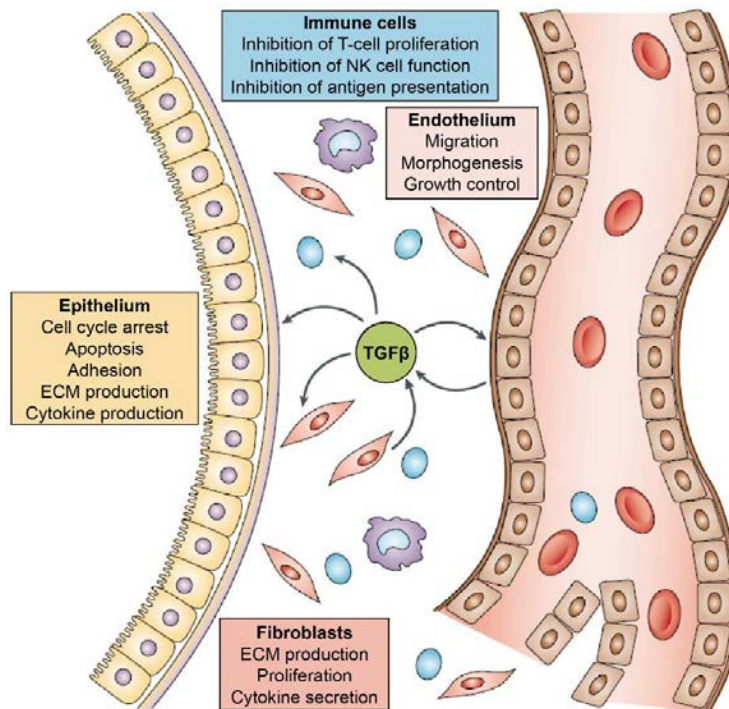


Figure I21. Physiological roles of the TGF β pathway in different cell types. Adapted from (Siegel and Massague, 2003).

In ESCs, SMAD2/3 co-occupies the genome with OCT4. Moreover, along differentiation, SMAD2/3 cooperate with the myogenic TF MYOD1 to induce the myogenic program, this is also the case for pre-B cell commitment, where SMAD2/3 colocalize with the master TF PU.1 (Mullen et al., 2011). Other contexts in which TGF β is involved in differentiation are the smooth muscle formation from neural crest cells, Schwann cells determination and osteoblast differentiation (Moses and Serra, 1996; Shah et al., 1996). Remarkably, during embryonic development, TGF β drives the EMT, a process that consists in the loss of apicobasal polarity and the acquisition of migratory mesenchymal traits (Larue and Bellacosa, 2005; Nieto et al., 2016). In morphogenesis, TGF β participates in vasculogenesis, 50% of *Tgfb1* KO mice died prenatally due to defects on yolk sac vasculature and defective hematopoiesis (Dickson et al., 1995) (Figure I21).

3.5.1.3 Role of the TGF β pathway in apoptosis and homeostasis

The role of TGF β in apoptosis is also related to cell homeostasis and to the **tumor suppressor activity**.

In epithelial cells, TGF β induces the TF TIEG1, this TF inhibits proliferation and induces apoptosis (Tachibana et al., 1997). Notably, not only the transcription program triggered by the SMAD proteins is important in apoptosis. Other mechanisms involving the potentiation of the TGF β apoptotic effect or the mediation

Introduction

between the TGF β molecule and the TGFBR type II participate in the TGF β apoptotic function (Larisch et al., 2000; Perlman et al., 2001; Siegel and Massague, 2003).

3.5.2 Pathological functions of the TGF β pathway

As it was above mentioned, TGF β pathway participates in many cellular processes. Consequently, malfunction or deregulation of this cascade is related to disease. In this subsection I will explain the implication of TGF β pathway in cancer and other diseases.

3.5.2.1 The TGF β pathway in cancer

The knowledge about TGF β in cancer is so extensive, that it will require thousands of pages to properly introduce it. Nonetheless, I would like to remark some general aspects.

The TGF β cytokine has dual roles in cancer. On one side, TGF β acts as a **tumor suppressor** in pre-malignant cells, but on the other side, TGF β drives **tumor progression** in malignant cells due to SMAD activity subversion (David and Massague, 2018). Many gastric cancers display mutations on the TGF β signaling components, concretely TGFBRs. This conduces to the loss of the tumor suppressor and apoptotic activities thus, leading to tumorigenesis (Fu et al., 2009).

In the differentiation roles of TGF β , I mentioned that TGF β drives **EMT**. In cancer, EMT promotes acquisition of mesenchymal traits by tumor cells. The migratory ability of cells increases the invasiveness and metastatic potential of the cancer (Nieto et al., 2016). In this regard, it has been demonstrated that the TGF β -promoted EMT is mediated by the KDM JMJD3. JMJD3 demethylates the promoter of the EMT regulator SNAI1 and triggers the mesenchymal program, thus contributing to metastasis (Ramadoss et al., 2012).

3.5.2.2 The TGF β pathway in other diseases

The TGF β signaling cascade is involved in many non-cancerous diseases and often the roots of these diseases are the **fibrotic processes**.

Upon inflammatory response, TGF β promotes the massive secretion of extracellular matrix components, cell migration and cell proliferation. These events form part of the chronic phase of inflammatory diseases such as arthritis, nephritis, Chron's disease, myocarditis etc. (Pohlers et al., 2009). Modulation of TGF β pathway by using agonists or inhibitory molecules is under study in order to achieve new therapeutic strategies for these inflammatory diseases (Hawinkels and Ten Dijke, 2011).

3.6 The TGF β pathway in neural development and its relation to JMJD3

In this last subsection of the thesis introduction, I will introduce the starting point of my doctoral work.

3.6.1 SMAD3 and JMJD3 coregulate the neural developmental program

Results from CNT demonstrate that the activation of the TGF β pathway leads to neuronal differentiation (Garcia-Campmany and Marti, 2007). Additionally, JMJD3 KDM has been shown to regulate neural developmental genes (Burgold et al., 2008; Jepsen et al., 2007). In this context, Dr. Marian Martínez Balbás' lab started its work aiming to elucidate whether the epigenetic regulator JMJD3 was cooperating with the TGF β pathway.

Firstly, they demonstrated that SMAD3 and JMJD3 interact in NSCs through the MH2 globular domain. Next, they performed a microarray experiment in which they observed that there was a subset of TGF β induced genes that were not so efficiently induced in a JMJD3 KD cell line (Estaras et al., 2012)

3.6.2 SMAD3 and JMJD3 colocalize on gene promoters

Another aspect covered in the work from Dr. Estarás et al., is the molecular mechanism behind the JMJD3-dependent TGF β gene activation. ChIP-seq experiments of both proteins showed an extensive genomic colocalization, on **promoters**, intragenic and intergenic regions. As their goal was to elucidate the mechanism governing transcriptional activation, they focused on gene promoters. On neural developmental gene promoters, SMAD3 recruits JMJD3 and by demethylating H3K27me₃, transcription gets activated (Estaras et al., 2012) (Figure I22).

3.6.3 TGF β promotes neurogenesis in the spinal cord in a JMJD3-dependent manner

In order to address the potential participation of JMJD3 in the neuronal differentiation *in vivo*, Dr. Estarás performed experiments in the CNT. She observed that the neuronal differentiation promoted by TGF β was dependent on JMJD3 by measuring proliferation and differentiation markers. Moreover, results suggested that the demethylation of H3K27me₃ was playing a role in this process (Estaras et al., 2012).

Introduction

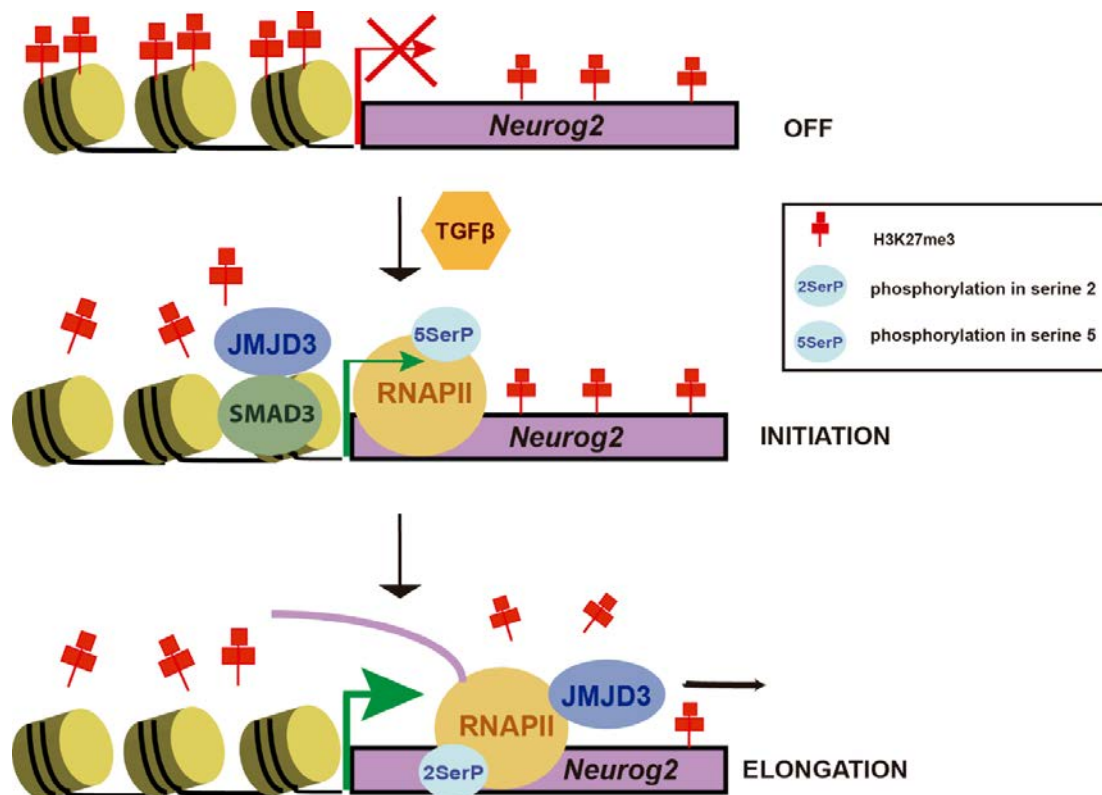


Figure I22. Cartoon model showing how the TGFβ pathway cooperates with the KDM JMJD3 in the regulation of the neurogenic program. Adapted from (Fueyo et al., 2015).

3.6.4 TGFβ induced transcription elongation is mediated by JMJD3

Another study from Dr. Marian Martínez-Balbás' lab in which I participated as part of my Master thesis, showed the role of TGFβ and JMJD3 on **transcription elongation** (Estaras et al., 2013) (Figure I22).

As I mentioned before, SMAD3 and JMJD3 co-occupy promoters, intergenic and intragenic regions along the genome. To elucidate the role of JMJD3 in intragenic regions, a ChIP-seq of the RNAPII-Ser2p upon TGFβ stimulation was performed, and it was observed, that there were intragenic regions occupied by both the RNAPII-Ser2p and JMJD3. Coimmunoprecipitation experiments demonstrated that JMJD3 interacts with the RNAPII-Ser2p and inhibition of the transcription elongation deployed JMJD3 from the intragenic regions. Conversely, in a KD cell line for JMJD3, RNAPII-Ser2p could not penetrate the intragenic region, probably, due to a lack of CDK9 recruitment on gene promoters. After TGFβ stimulation, the JMJD3 intragenic target regions suffered mild changes on H3K27me3 demethylation (Estaras et al., 2013).

Concluding the introduction, I would like to summarize, that SMAD3 and JMJD3 have repeatedly demonstrated to be cooperation partners in neural development. My doctoral thesis is framed in this context and had as a primary goal the deepening on the transcriptional role of JMJD3 after TGF β signaling stimulation.

Objectives

The PhD thesis objectives were the result of previous work from the laboratory of Dr. Marian Martínez-Balbás lab (Estaras et al., 2012; Estaras et al., 2013). In these research articles the cooperation between the TGF β pathway and the histone demethylase JMJD3 at transcription activation and elongation was analyzed.

Starting from these achievements, several goals were proposed for this PhD thesis in order to further deepen in the contribution of histone demethylase JMJD3 to the TGF β response. These are the following:

1. To analyze the role of the histone demethylase JMJD3 at enhancers in the context of the TGF β signaling activation.
2. To elucidate the molecular players that participated in the TGF β activation of enhancers.
3. To study the changes in the tridimensional structure of the chromatin triggered by the TGF β signaling cascade and the contribution of JMJD3 to these changes.

Methodology

In this section of my doctoral thesis I will provide methodological details on how the experimental work, the bioinformatic analysis and the statistics have been performed.

1. Materials

First I will item the reagents and the biological models of this doctoral thesis.

1.1. Models of study

Along this thesis I have mainly worked with *in vitro* models consisting in stem and transformed cell lines. Additionally, there is a particular experiment that has been performed in an *in vivo* model, the CNT.

1.1.1. In vitro models

As I mentioned in the introduction, my model of study has been the mouse NSCs. However, I also performed experiments in human HEK293T cells. In the following subsections I will provide details on the freezing and culture conditions of these cell lines.

1.1.1.1. Mouse neural stem cells

Mouse NSCs were dissected from cerebral cortices of C57BL/6J fetal brains (E12.5) and cultured in poly-D-lysine (Millipore # A-003-E) (5µg/mL, 2h at 37°C) and laminine (Sigma #L2020) (5 µg/mL, 4h at 37°C) precoated dishes following the previous published procedures (Currle et al., 2007) (Figure M1).

NSCs were grown in a medium prepared with equal parts of DMEM F12 (without Phenol Red, Gibco #31331093) and

Neurobasal Media (Gibco #12348-017) containing 1% of Penicillin/Streptomycin (Gibco #15140-122), 1% Glutamax (Gibco #35050061), N2 and B27 supplements (Gibco #17502-048 and #17504-044 respectively), 0.1mM non-essential amino acids (Gibco #11140035), 1mM sodium pyruvate (Gibco #11360039), 5mM HEPES (Gibco #15630056), 2mg/L of heparin (Sigma #H-4784), 25mg/L of bovine serum albumin (Sigma #A7906) and 0.01mM of β-mercaptoethanol (Gibco #31350-010) as previously described (Currle et al., 2007; Estaras et al., 2012). Fresh recombinant

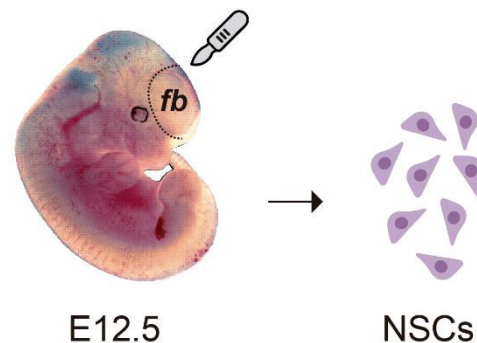


Figure M1. Schematic representation of the NSCs origin. fb: forebrain

Methodology

human EGF (R&D systems #236-EG) and FGF (Invitrogen #PHG0021) to 20 ng/mL and 10 ng/mL final concentrations respectively were added to the growing medium. Medium, supplements, EGF and FGF form the so-called expansion medium. Under these conditions, NSCs maintain the ability to self-renew and to generate a wide range of differentiated neural cell types (Currele et al., 2007; Pollard et al., 2006; Theus et al., 2012).

1.1.1.2. Human HEK293T cells

In order to perform cell transfections to produce lentivirus or to overexpress proteins, we used HEK293T cells taking advantage of their high degree of transfectability. These cells derive from human embryonic kidney transformed with the large T antigen of the SV40 virus (Graham et al., 1977). HEK293T cells were cultured in DMEM (Gibco #41965-062) supplemented with 10% of fetal bovine serum (FBS) (Gibco #10270106) and 1% of Penicillin/Streptomycin (Blanco-Garcia et al., 2009).

1.1.2. *In vivo* model

In order to address the enhancer activity *in vivo*, I collaborated with Stella Pappa, a member of the lab with expertise on CNT manipulation.

1.1.2.1. Chicken neural tube

The CNT is an *in vivo* model that has been used in this doctoral thesis in order to perform *in ovo* electroporation experiments of luciferase constructs that will be further explained in upcoming subsections. The CNT is an extraordinary system to study the effect of genetic modifications along neuronal differentiation.

Eggs from White-Leghorn chickens were incubated at 38.5°C and 70% humidity. Embryos were staged following Hamburger and Hamilton (HH) (Hamburger and Hamilton, 1992).

1.2. Reagents

Most of the reagents will be explained in the specific procedures where they are required. Nonetheless, I would like to mention here the human recombinant TGFβ1 cytokine that I have been using for the performance of experiments. TGFβ1 was purchased from Millipore (#GF111) and has been used at a concentration of 5ng/mL during the indicated times. Another reagent used for the indirect immunofluorescence experiments is DAPI (4',6-Diamidino-2-Phenylindole, Dihydrochloride), this has been purchased in Thermofisher (#D1306) and it is used at a concentration of 0.1ng/μL.

1.3. Plasmids

In the following table (Table M1), there is a summary of the plasmids used during this doctoral thesis, their origin and the experiment in which they were involved.

Plasmid	Origin / Provider
pCMV-VSVG	Dr. Timothy Thomson
pCMV-GAL-POL	Dr. Timothy Thomson
pLKO.1-shASCL1 (CCACGGTCTTTGCTTCTGTTT)	Sigma
pLKO.1-shSMAD3 (CCTTACCACTATCAGAGAGTA)	Sigma
pLKO.1-shJMJD3 (CCTCTGTTCTTGAGGGACAAA)	Sigma
pLKO.1-shCHD8 (TGCCTGGAAGAAATTGGAG)	Sigma
pCIG-HA-ASCL1	Dr. Elisa Martí
pCIG-FLAG-SMAD3	Dr. Elisa Martí
pCDNA3-MYC-JMJD3	Dr. Naiara Akizu
pCIG-FLAG-SMAD3 S/D	Dr. Elisa Martí
pCIG-FLAG-SMAD3 S/A	Dr. Elisa Martí
pGL3-promoter	Promega
pRL-TK Renilla	Promega
pGL3- <i>Ctgf</i> enh	Cloned by Raquel Fueyo
pGL3- <i>Nrip3</i> enh	Cloned by Raquel Fueyo
pGL3- <i>Neurog2</i> enh	Cloned by Raquel Fueyo
pX330-U6-Chimeric_BB-CBh-hSpCas9	Addgene #42230
pX330-U6-Chimeric_BB-CBh-hSpCas9- <i>Neurog2</i> enh left	Cloned by Raquel Fueyo
pX330-U6-Chimeric_BB-CBh-hSpCas9- <i>Neurog2</i> enh right	Cloned by Raquel Fueyo
pX330-U6-Chimeric_BB-CBh-hSpCas9- <i>Neurog2</i> pro left	Cloned by Raquel Fueyo
pX330-U6-Chimeric_BB-CBh-hSpCas9- <i>Neurog2</i> pro right	Cloned by Raquel Fueyo

Table M1. List of the plasmids used in this thesis and their origin. Green boxes correspond to virus production vectors, blue to overexpression vectors, orange to luciferase vectors and purple to CRISPR-Cas9 vectors. In the pLKO.1 plasmids brackets indicate the target sequence.

1.4. Antibodies

Along the experimental work of this thesis I have performed assays with different antibodies. Their targets, host and dilution for the different experiments can be consulted in Table M2.

Methodology

	Antibody target	Provider and reference	Dilution for the indicated application
Primary	Unspecific IgGs	Diagenode #C15410206	ChIP, Same dilution than the specific IgG
	Unspecific IgGs	Santa Cruz #sc-2025	ChIP, Same dilution than the specific IgG
	SMAD3	Abcam #28379	Western Blot 1:2000
	SMAD2/3	BD Biosciences #610842	Immunoprecipitation 1µg/0.5mL
	Phospho-SMAD3	Cell Signaling #mAb9520	Western Blot 1:1000
	ASCL1	BD Pharmingen #556604	Western Blot 1:1000
	H3K27me3	Millipore #07449	ChIP 3µg/mL
	H3K4me1	Millipore #07436	ChIP 2µg/mL
	H3K4me2	Millipore #07030	ChIP 2µg/mL
	H3K27me2	Cell Signaling #9728	ChIP 2µg/mL
	H3K27ac	Abcam #ab4729	ChIP 3µg/mL
	H3K4me3	Abcam #ab8580	ChIP 3µg/mL
	JMJD3 (serum)	Our lab	Western Blot 1:250; ChIP 12µL/mL
	CHD8 (serum)	Dr. Reyes lab	Western Blot 1:1000; ChIP 12µL/mL
	EZH2	Active motif #39875	ChIP 3µg/mL
	β-TUBULIN	Millipore #MAB3408	Western Blot 1:20000
	VINCULIN	Sigma #V9131	Western Blot 1:10000
	β-TUBULIN III (TUJ1)	Covance #MMS-435P	Immunochemistry 1:500
	NESTIN	Abcam #ab5968	Immunochemistry 1:500
	GFAP	Dako #z0334	Immunochemistry 1:500
HuC/D	Invitrogen #A21271	Immunochemistry 1:500	
FLAG tag	Sigma #F3165	Western Blot 1:5000	
HA tag	Abcam #ab20084	Western Blot 1:2000	
MYC tag	Abcam #ab9132	Western Blot 1:1000	
Secondary	Anti-Rabbit IgG HRP	Amersham #GENA934	Western Blot 1:10000
	Anti-Mouse IgG HRP	Amersham #NA9310	Western Blot 1:10000
	Anti-Rabbit IgG IRDye	LI-COR #926-32210	Western Blot 1:10000
	Anti-Mouse IgG IRDye	LI-COR #926-32210	Western Blot 1:10000
	Anti-Goat IgG IRDye	LI-COR #925-32214	Western Blot 1:10000
	Anti-Rabbit Fluor 568	Alexa #A10042	Immunochemistry 1:1000
	Anti-Mouse Fluor 488	Alexa #A21202	Immunochemistry 1:1000

Table M2. List of primary and secondary antibodies used in thesis. Providers, references, and usage concentrations are indicated.

1.5. Primers

In my doctoral thesis there are techniques that require the usage of DNA amplification in order to analyze the results. Among these procedures are ChIPs, RT-qPCR (retrotranscription followed by quantitative polymerase chain reaction) and conventional PCR (polymerase chain reaction). Additionally, we have also utilized primers for the cloning of luciferase reporters and CRISPR-Cas9 plasmids. The sequences and experimental purposes of every set of primers can be found in the Table M3.

	GENE/REGION	FW	RV	
cDNA	<i>Ctgf</i> mRNA	ACCCAACTATGATGCGAG	CAGGTCTTAGAACAGGGCG	
	<i>Nrip3</i> mRNA	CTAGCTTCCAGTCTAACCAAAA	TCTATTCCCTTACCACACTTCAC	
	<i>Fabp4</i> mRNA	GATGAAATCACCGCAGACGACA	ATTGTGGTTCGACTTTCCATCCC	
	<i>Neurog2</i> mRNA	CACAACCTAAAGGCCCGC	TCTTCGTGAGCTTGGCAT	
	<i>Ccne3</i> mRNA	TGGTCCATGGCAGTTGCGG	GGCAGAGACAGGGCGGTGCAG	
	<i>Utx</i> mRNA	CTGACCAGCAGCACAGA	CTCGGACCTTTGTGAAGC	
	<i>Rps23</i> mRNA	CGTCAGGGTGCACTCATT	GGCACGAACGCTGTGATCTT	
	<i>Chic2</i> mRNA	ACGTCACCGTATTTGGACTG	GAGGCAAATAACTGGCCACA	
	<i>Tle3</i> mRNA	TGGTGAGCTTTGGAGCTGTT	CGTTTTCCCTCCAGGAAT	
	<i>Ascl1</i> mRNA	CTGTCGCCACCATC	GTCCGAGGAGTAGGACG	
	<i>Gfap</i> mRNA	GACATCGAGATCGCCACCTA	CTTCTTTGGTGCTTTGGCCCC	
	<i>Nes</i> mRNA	GTCAGATCGCTCAGATCTCTGG	GTGTCTGCAAGCGAGAGTTC	
	<i>Chst8</i> mRNA	CTTTCCCGAGGTTCAAGGAC	GGTCCGAGAAGGTTTGGAG	
	<i>Peptd</i> mRNA	CGAGGGCATTAGCAAGTTCA	AGCGCAAACCTCTAGTTCC	
	<i>Ctgf</i> eRNA	TCACGCTGCTCCCTT	CTCTTGCTTTCCTCTGTATG	
	<i>Nrip3</i> eRNA	GGAACACCCAGGAAGC	AGCCACACTTGATACGG	
	<i>Fabp4</i> eRNA	TGAGTCCCCCACTTGCTTTA	CACCCTGTAAGGCTGGTGAT	
	<i>Neurog2</i> eRNA	TCCTTTCTCTCTATCTCCTAACAC	GACCACCTCTGTAGAAATGC	
	ChIP	<i>Neurog2</i> (-6) enhancer	TCCTGGTTCAGCGATTAGT	CTTCTCAAGCCTGCGT
		<i>Chic2</i> (-26) enhancer	AAGTGGTGGATTCCGAC	CCGCTGGAGAATGAGATA
<i>Tle3</i> (+158) enhancer		CAAATAAGAGGGACACTGCT	GGCTGGCAGGTAGGA	
<i>Ctgf</i> (-102) enhancer		TCACGCTGCTCCCTT	CTCTTGCTTTCCTCTGTATG	
<i>Nrip3</i> (-3,5) enhancer		CTCTGGCACACTTTACTAGC	GTGAACAACCAGGGCG	
<i>Neurog2</i> promoter		TAACTGGAGTGCCCTGG	CTCGTGTGTTGTCGTT	
<i>Neurog2</i> TSS		TAAGGAGAGTCGTGTGCC	GCACTCCAGTTAAAGCCG	
<i>Stx3</i> C+		AAAGGGAGCAGGAACCA	CACTGTCTTATTGTTCGGC	
<i>G6pd2</i>		ACAGTCTATGAAGCAGTCAC	CTCCACTATGATGCGGTTA	
<i>Chic2</i> promoter		TGATGCTTTCGTCTCTAGT	GGTCCAAGGTGATCCATGA	
<i>Tle3</i> promoter		CAGTTCAGACCTCGTGC	CCTCGGACCTGCGGATA	
<i>Ctgf</i> promoter		TTACTTCTGGTGTGCT	TCCACATTCCTCCGCTG	
<i>Hbb1</i> promoter		ATCACGCTTTGGACATTGCTC	CCTTCTTGCTCCCAAA	
Intergenic region		GCCACAGTTAGCCATCTGCA	GCCCTGACCTCTGGACCCTA	
<i>Neurog2</i> (-6) left cut		CACCGCTTATGTGCCACTCGTG	AAACCACGAGTGGCACATAAGC	
<i>Neurog2</i> (-6) right cut		CACCGGCTTGGCTGAACGGAT	AAACATCCGTTCCAGCAAGGCC	
CRISPR-Cas9		<i>Neurog2</i> (-6) test deltion A+B	GCTCTTATGTGCCACTCGTG	CTTAGCATCTTCCCCTGGT
	<i>Neurog2</i> (-6) test deltion C+D	GTGAGTCCAGGACAGC	GATTCAGTCTCACCATCA	
	<i>Neurog2</i> promoter left cut	CACCGACAGACGTCCTTAGAAG	AAACCTTCTAGGGGACGTCCTGTG	
	<i>Neurog2</i> promoter right cut	CACCGCCAGCCGAGGCCCCCGG	AAACCGGGGGGCTCTCGGCTGGC	
	<i>Neurog2</i> promoter test deltion E+F	CTCAGGAAAAATGGGTGTGT	ACTTACCTACGGGCTCCAG	
	<i>Neurog2</i> promoter test deltion E+G	CTCAGGAAAAATGGGTGTGT	GCCGCTGTCTATTGTGATTG	
	<i>6gpd2</i> control deletion	ACAGTCTATGAAGCAGTCAC	CTCCACTATGATGCGGTTA	
	Cloning for luciferase	<i>Ctgf</i> (-102) enhancer	ATAACGCGTAAGCGCCAAAGGACTCAA	GCTAGATCTTAGGCCACGAAAGCCATG
<i>Nrip3</i> (-3,5) enhancer		TCTTGTGGAACGACCCAC	TGACTGGAGACTTGAGGCAG	
<i>Neurog2</i> (-6) enhancer		ATAAGATCTTATCCATTTCTGAG	GCTACGCGTCTCACCAGGCTGGTAAG	
4C-seq	<i>Chst8</i> (-38) viewpoint	GGTTCCCCGTCATTATG	CCAGCTGGTTAACACAGGA	

Table M3. List of primers used in this doctoral thesis

1.6. Gene expression omnibus accessions

In this doctoral thesis we have used previously published NGS datasets to test our hypothesis. These types of genome-wide data are usually deposited in the GEO (gene expression omnibus) database (<https://www.ncbi.nlm.nih.gov/geo>). Accession numbers can be found at Table M4.

2. Experimental methods

In this section I will outline the protocols of the different technical approaches that we have been performing for this doctoral thesis.

Methodology

Data	Accession
p300 ChIP-seq	GEO GSM1052708
H3K27ac ChIP-seq	GEO GSE66961
H3K4me1 ChIP-seq	GEO GSE66961
H3K4me3 ChIP-seq	GEO GSE66961
H3K27me3 ChIP-seq	GEO GSE38269
SMAD3 0,5h TGF β ChIP-seq	GEO GSM898371
ASCL1 ChIP-seq	E-MTAB 2384
JMJD3 0,5h TGF β ChIP-seq	GEO GSM898372
JMJD3 3h TGF β ChIP-seq	GEO GSM937827
H3K27ac E17.5 ChIP-seq	GSM1264358
SMC1 ChIP-seq	GSM883646
Active enhancers in neurons Dataset	GEO GSE21161
shCHD8 RNA-seq ChIP-seq	GEO GSE72442
TGF β -regulated JMJD3-dependent (JDTA)	GSE35361

Table M4. List of the accession codes for the next generation sequencing data used in this doctoral thesis

2.1. Cell culture and genetic manipulation of cell lines

In this section I will describe methods for working with cell lines and also methods to genetically manipulate the cells.

2.1.1. Cell culture growth and maintenance

With the goal of expanding the previously mentioned cell lines to achieve the sufficient number of cells to perform the experiments, I have been working in the cell culture room following the sterile environment normative. As a rule of thumb in cell culture, cells should be frozen slowly and thaw as fast as possible in order to optimize cell survival. In this subsection, I will explain how cells are thaw and freeze and how often cells are passaged. Please, refer to subsections 1.1.1.1 and 1.1.1.2 for growing conditions.

2.1.1.1. Mouse neural stem cells

NSCs are really sensitive to mechanical and chemical manipulation. For this reason, while working with this cell line, extensive pipetting and changes of temperature must be avoided. Also, cells should not be grown over 25-30 passages in order to maintain a stable epigenome.

2.1.1.1.1. Mouse neural stem cells expansion

It is important to note that NSCs trigger differentiating signals under confluent conditions. For this fact, cells are usually split in a ratio of 1:5 every two days. To pass these cells the medium is removed and after rinsing with PBS (Phosphate Buffered Saline) the Accutase enzyme (Gibco #A11105-01) is added. This enzyme is a substitute of the classical trypsin and its effect in cell detachment is milder, thus being more adequate for NSCs. After 2 minutes at 37°C, the Accutase is diluted with PBS and cells are centrifuged at 260 rcf (relative centrifugal force) for 3 minutes. Cell pellet is gently flicked and then diluted in previously warmed expansion medium supplemented with growth factors. Then, medium containing the cells is split in the poly-D-lysine and laminine pre-coated plates.

2.1.1.1.2. Mouse neural stem cells thawing and freezing

NSCs can be stored at -80°C for weeks and in liquid N₂ for years. Freezing is performed following the same protocol as in passaging but instead of diluting the cells in regular medium they are diluted in medium containing 10% of sterile DMSO. Temperature descends in a slow manner helped by a box named Mr. Frosty (Nalgene, #5100-0001). This box contains isopropanol, an alcohol that decreases its temperature 1°C/minute, thus allowing an optimal freezing process.

Cells thawing is rapidly performed by placing the cryotube in a 37°C water bath and as soon as it reaches the liquid state the content is diluted in PBS and centrifuged at 260 rcf during 3 minutes. Cells are then diluted in the NSCs expansion medium.

2.1.1.1.3. Mouse neural stem cells differentiation experiments

For differentiation experiments NSCs were seeded in 24-well plates pre-coated with poly-D-lysine (5 µg/mL, 2 hours at 37°C) and laminine (5 µg/mL, 4 hours at 37°C) at a seeding density of 0.1×10^6 cells per well in NSCs expansion medium. After 24 hours, expansion medium was replaced by differentiating medium, consisting in the same components of the expansion medium but without EGF and FGF (Galderisi et al., 2013; Schwindt et al., 2009; Sun et al., 2017). Fresh differentiating medium was supplied every 2 days and after 3, 6 or 8 days, cells were fixed and stained for indirect immunofluorescence. Under these conditions, NSCs differentiate towards neurons, astrocytes and oligodendrocytes (Jori et al., 2007).

Methodology

2.1.1.2. HEK293T cells

Opposite to NSCs, HEK293T cells should be pipetted to avoid clumps. In the following lines I will describe how to grow, freeze and thaw these cells.

2.1.1.2.1. HEK293T cells expansion

These cells can be passed in a ratio of 1:8 every two days. To pass these cells we remove the medium and rinse with PBS, then cells are detached using the classical chemical dissociating agent trypsin (Sigma #T9935) diluted with EDTA. After one minute at 37°C, trypsin activity is stopped by adding two volumes of medium and the desired split ratio is applied.

2.1.1.2.2. HEK293T cells freezing and thawing

Like NSCs, HEK293T cells can be stored at -80°C for weeks and in liquid N₂ for years. To freeze the cells, trypsinization is stopped with PBS and cells are centrifuged at 1500 rcf for 1 minute. Then, cell pellet is diluted in FBS containing 10% of DMSO and cryotubes are introduced in the Mr. Frosty box at -80°C. HEK293T cells thawing is identical to that of NSCs.

2.1.2. Genetic manipulation of growing cells

In science, manipulation of living systems can be achieved by introducing exogenous DNA that will ultimately code for specific proteins or impede the translation of mRNAs.

In this section I will outline the procedures that we have used to deliver DNA into the cells.

2.1.2.1. Calcium phosphate transfection

This method permits the delivery of plasmids into the cells through the endocytosis of calcium phosphate precipitates that contain DNA molecules stuck on the surface. To generate these precipitates, the DNA is included into a mixture of HEBS (250mM NaCl, 9mM KCl, 1.5mM Na₂HPO₄, 10mM glucose and 50mM Hepes pH 7.12) and 0.25M CaCl₂ under vortexing conditions. After 10 minutes at room temperature, the mixture is added to the growing medium. 6 hours later, new medium is supplied to avoid cell stress due to acidity.

2.1.2.2. Nucleofection

This technique serves to deliver exogenous DNA into the cells. NSCs are not efficiently transfected with calcium phosphate and require this special method. To perform nucleofections, we used the patented Amaxa kit specific for mouse NSCs (Lonza #VPG-1004).

To nucleofect, NSCs pellet is diluted in a solution of unknown composition containing the DNA. A cuvette containing this mixture is introduced into the Amaxa Nucleofector device and after an electrical pulse of unknown parameters cells acquire the exogenous DNA. Then, NSCs are diluted in plates containing pre-heated medium. It is worth noting, that this process is harsh for the cells and normally, around 40% of NSCs die.

2.1.2.3. Lentiviral transduction

This delivery system is highly efficient in NSCs, achieving up to a 90% of genetically modified cells. In this doctoral thesis, this method has been applied for the transduction of shRNAs that will permit the knocking down of proteins in our system. The procedure consists in three steps: lentiviral production, lentiviral transduction and selection.

Lentiviral particles are produced in HEK293T cells by cotransfecting in four 10cm plates the DNA encoding the shRNA (pLKO.1-shControl, pLKO.1-shASCL1, pLKO.1-shSMAD3, pLKO.1-JMJD3 or pLKO.1-shCHD8; Table M1) together with pCMV-VSVG and pCMV-GAG-POL plasmids that encode the viral capsid and transcriptional machinery respectively. After 24 hours, supernatants containing lentiviral particles are collected and centrifuged in a sucrose bed at 57000xg during 2 hours. Then, supernatant is removed and viral particles are resuspended in NSCs medium. These particles can be stored at -80°C or be directly used for infection.

Transduction of the previously produced cells consists in the addition of the medium containing viral particle to the receptor NSCs. Approximately one production will infect 1.5×10^6 cells. One day after infection, cells are selected with the correspondent antibiotic, that for pLKO.1 plasmids is puromycin (Sigma #P8833) at a concentration of 2µg/mL. After 48 hours cells can be considered selected and the knocking down of the particular protein can be assessed by RT-qPCR or Western Blot.

2.1.2.4. CRISPR-Cas9 cell line generation

CRISPR-Cas9 technology (Ran et al., 2013) has permitted a versatile manner of genetically manipulate the cells. In this doctoral thesis I have used this technique in order to delete the enhancer and promoter corresponding to the *Neurog2* gene.

Methodology

First, we designed two pairs of gRNAs (guide RNAs) flanking *Neurog2* enhancer (genomic coordinates chr3:127326051–127334232) or promoter (genomic coordinates chr3:127335145–127336247). One primer pair generates a cut on the left side and the other on the right side, thus getting the enhancer removed. To design these gRNAs we used the online tool Genetic Perturbation Platform from the Broad Institute (<https://portals.broadinstitute.org/gpp/public/>). Selected pair of primers have an on-target score of 88 (left) and 87 (right) in the case of the *Neurog2* enhancer and 93 (left) and 86 (right) in the case of the promoter. Specificity was assessed observing that the highest off-target score for each pair of primers was theoretically low (1.4 left, 1.3 right). These gRNAs were cloned in the pX330-U6-Chimeric-BBCBh-hSpCas9 vector and then, left and right cutting plasmids were nucleofected in NSCs. After puromycin selection (2µg/mL) and detection analysis with conventional PCR, heterogeneous population carrying a majority of homozygotic deletions was used for experiments.

2.2. Chicken embryo manipulation

In this doctoral thesis the CNT has only been used for ovoid electroporation of luciferase constructs.

2.2.1. Ovoid electroporation of chicken embryonic neural tubes

Chick embryos were electroporated (EP) with purified plasmidic DNA at 1 µg/µL in H₂O with 50 ng/mL of Fast Green (Sigma # F7258). Plasmidic DNA was injected into the lumen of HH11–HH12 CNTs, electrodes were placed at both sides of the neural tube and embryos were EP by an IntracelDual Pulse (TSS-100) electroporator delivering five 50 milliseconds square pulses of 20–25V (volts) (Figure M2).

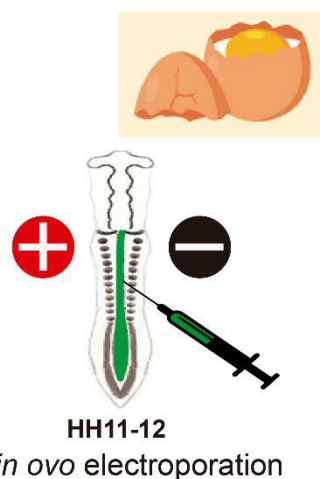


Figure M2. Schematic view of the CNT electroporation.

2.3. Molecular biology procedures

In this subsection I will detail the protocols of the molecular biology techniques performed along this thesis.

2.3.1. Nucleic acids-related

In the following lines I will describe the procedures in which either DNA or RNA are involved.

2.3.1.1. Genomic DNA extraction

Approximately 6×10^6 cells were lysed in 200 μ L of buffer (10mM Tris-HCl, 10mM EDTA, 10mM NaCl and 0.5% SDS). Then, lysates were incubated with 0.5mg/mL of proteinase K (Sigma #P2308) during 1 hour at 50°C and with 1mg/mL of RNase A (Fermentas #EN0531) during 2 hours at 50°C. Finally, the resulting mixtures of DNA and proteins were subjected to phenol-chloroform extraction to purify the DNA.

2.3.1.2. Phenol-chloroform extraction and ethanol precipitation

This procedure serves to purify DNA from complex protein-DNA mixtures. It is based in the different affinity of DNA and proteins for the organic solvent phenol.

First, 1 volumen of phenol is added to the mixture, after vortexing and centrifuging at maximum speed during 3 minutes the aqueous and the organic phases are separated, and the aqueous phase containing the DNA is moved to a clean tube. This step is repeated with chloroform and the aqueous phase obtained is subjected to ethanol precipitation.

Ethanol precipitation follows phenol-chloroform extraction in order to concentrate DNA and desalt the previously obtained aqueous phase. First, 0.1 volumes of NaAc 3M and 1 volumen of cold ethanol are added to the sample. Then, tubes are incubated at -80°C to favor precipitation and after centrifugation at maximum speed during 20 minutes at 4°C, an additional wash with ethanol 70% is performed. Finally, the pellet is dried and the DNA resuspended in a suitable buffer.

2.3.1.3. RNA extraction and DNase treatment

Trizol reagent (Invitrogen #15596018) has been used to extract RNA. Approximately 3×10^6 cells are lysed with 1mL of Trizol. Then, 200 μ L of chloroform are added and after centrifugation at maximum speed during 5 minutes the upper phase of the tube is collected and precipitated by adding 800 μ L of isopropanol. After another centrifugation at maximum speed during 10 minutes, supernatant is discarded and

Methodology

pellet is washed with 70% ethanol. After a final centrifugation of 5 minutes at maximum speed, pellet is resuspended in H₂O.

To avoid contamination of genomic DNA in the resuspended RNA, samples are treated with DNase. For this treatment, I have used the DNA-free Kit (Ambion # AM1906) that is special for DNA removal from RNA samples due to the inert beads that serve as inhibition agent. The protocol consists in the addition of 0.1 volumes of 10X buffer to the RNA sample, that later is incubated with 1 μ L of DNase during 30 minutes at 37°C. Next, 0.1 volumes of inhibition agent are added and after 2 minutes of flicking, pure RNA can be obtained by collecting the supernatant of a centrifugation at 10000 rcf during 1.5 minutes.

After this, RNA is quantified using a Nanodrop device and quality is evaluated with 260/280 and 260/230 ratios. Besides, I run an agarose gel to check RNA integrity before proceeding with other techniques.

2.3.1.4. Retrotranscription of RNA into cDNA

To quantify gene expression and enhancer transcription I performed RT-qPCR experiments. This technique allows the quantification of the RNA levels in the different conditions tested.

Reverse transcription of mRNA or eRNA was performed with 1 or 2 μ g of RNA respectively, using the High Capacity cDNA reverse transcription kit (Invitrogen #4368814). The protocol consists in the incubation of the RNA with random hexamers, DNTPs and a retrotranscriptase in a thermocycler following these parametres: 25°C 10 minutes, 37°C 120 minutes and 85°C 5 minutes. The complementary DNA (cDNA) generated can be stored at -20°C or -80°C.

2.3.1.5. Conventional PCR

Along this doctoral thesis, conventional PCR has been used for amplification of specific genomic regions for cloning or as a means for genotyping the CRISPR-Cas9 generated cell lines.

The protocol consists in the incubation of genomic DNA with deoxynucleotides (DNTPs), buffer, specific primers and Taq DNA polymerase (Fermentas # EP0401) in a thermoblock following a specific set of temperatures and incubation times depending on the target region. After the reaction, the PCR product is observed in an agarose gel.

2.3.1.6. Quantitative PCR

After retrotranscription of RNA I used qPCR to quantify cDNA. qPCR experiments are based on the quantification of the emitted fluorescence by a fluorophore that

binds DNA as PCR proceeds. A higher presence of cDNA will result in more emitted fluorescence and *vice versa*.

qPCR reactions were manually set-up in a volume of 10 μ L using SYBR Green Kit (Roche #4887352001). Reactions were carried out in 96-well plates in a LightCycler 480 (Roche) with the following cycling conditions (95°C 5 minutes, 40 cycles of 95°C 5 minutes - 60°C 10 seconds - 72°C 20 seconds, melting curve 95°C 5 seconds - 65°C 1 minute - 97°C). Specific primer pairs were designed spanning exon-exon junctions of a region conserved between splice variants (mRNA) and designed against the specific coordinates of every enhancer (eRNA). Primer pairs can be found in Table M3.

To validate qPCR results, we run non-template controls and standard curves with every new primer pair, and only primers with an efficiency of 95% or higher were kept. qPCR data were analyzed using the $2^{-\Delta\Delta CT}$ method. Outliers were defined as values that differ more than 0.5 cycles from the other two wells in the triplicates. When identified, outliers were discarded. If the non-template controls were Ct=37 or lower reaction was repeated.

In order to normalize qPCR results we usually used *Rps23* as a reference gene because after testing different housekeeping genes this has been the most constant and reliable between conditions.

2.3.1.7. Cloning of plasmidic DNA

In order to generate plasmidic vectors that were not available in the laboratory, we used different cloning strategies. In this section, I will detail the obtaining of every construct. Please note that all constructs have been Sanger sequenced after cloning in GATC Biotech (Germany).

2.3.1.7.1. pLKO.1 constructs

In the case of the protein CHD8, the pLKO.1 construct to knockdown its levels in mouse was not available to purchase in Sigma. shRNAs against CHD8 were designed in the Genetic Perturbation Platform of the Broad Institute (<https://portals.broadinstitute.org/gpp/public/>) and then, they were introduced in the pLKO.1 puro empty vector (Addgene #8453).

Firstly, the shRNA oligos were annealed by heating at 95°C in a thermoblock and then allowing them to cool down by turning it off. Then, phosphorylation of these oligos was performed with the T4 Polynucleotide kinase (PNK) (Fermentas #EK0031) in the provided buffer and supplementing the reaction with ATP (adenosine triphosphate). In parallel, 2 μ g of the destination vector was cut using the restriction enzymes EcoRI and AgeI. After purification of the digested vector from an agarose gel, the annealed phosphorylated oligos and the digested vector were

Methodology

ligated using the enzyme T4 DNA ligase in the provided buffer (Fermentas # 15224017)

2.3.1.7.2. Luciferase constructs

For the cloning of the enhancer regions in the pGL3 promoter vector (Promega # E1761), we first extracted the enhancer regions from the mouse genomic DNA using conventional PCR and the primers detailed in Table M3, it is worth noting, that these primers contain cohesive ends for the restriction enzymes that were used in the destination vector. In parallel, 2µg of pGL3 promoter destination vector was cut with MluI and BglII restriction enzymes and after purification by agarose gel, the PCR fragment and the digested vector were ligated using T4 DNA ligase.

2.3.1.7.3. CRISPR-Cas9 constructs

In order to clone the gRNAs of interest in the pX330-U6-Chimeric_BB-CBh-hSpCas9 empty vector (Addgene #42230), we used the following protocol. First, the gRNAs were designed in the Genetic Perturbation Platform of the Broad Institute (<https://portals.broadinstitute.org/gpp/public/>). Then, the ordered oligos were annealed by heating at 95°C in a thermoblock and then allowing them to cool down by turning it off. The destination vector was digested using BbsI restriction enzyme and finally, the annealed oligos and the digested vector were ligated using T4 DNA ligase.

2.3.1.8. Amplification of plasmidic DNA

Once the desired construct has been selected and obtained, DNA must be amplified in order to have enough micrograms to perform the specific techniques. For this amplification, we use *Escherichia coli* strain DH5α. Transformation is achieved by mixing 30µL of competent bacteria with DNA of concentration ranging from 20ng to 200ng. This mixture is maintained on ice during 30 minutes and then, heat at 42°C during 2 minutes. This change of temperature allows the transient opening of pores in the bacterial cell wall thus permitting the entrance of the plasmidic DNA. Afterwards, bacteria are recovered by incubating them in 1mL of LB (lysogeny broth) at 37°C and agitation and they can be either plated in agar to isolate colonies or grown in bigger volumes of LB to continue with mini- or maxipreps of DNA.

2.3.1.8.1. Mini- and maxipreparations of DNA

In molecular biology, solutions of DNA containing around 300ng/µL of DNA are known as “minipreps”, likewise, solutions of DNA containing approximately 1µg/µL

or more are known as “maxipreps”. For the obtaining of these solutions of DNA from bacteria transformation, 5mL or 500mL of LB are inoculated with either an isolated colony or 5mL of miniculture. After overnight growth, the bacterial mass is subjected to DNA purification with the alkaline lysis method following the protocol and using the buffers of the QIAprep Spin Miniprep Kit (for minipreps, QIAGEN # 27106) and of the QIAGEN Plasmid Maxi Kit (for maxipreps, QIAGEN # 12165).

The alkaline lysis method has three steps: resuspension, lysis and neutralization. The protocol consists in the sequential addition of three buffers (P1, P2 and P3) corresponding to the three mentioned steps (Buffer P1: 100µg/mL RNase A, 50mM Tris-HCl, 10mM EDTA pH 8.0; buffer P2: 200mM NaOH, 1% SDS; buffer P3: KAc 3M, pH 5.5), then the lysate is passed through a column that specifically retains DNA, and after washes, DNA is eluted and precipitated with isopropanol. Finally, the DNA is washed with 70% ethanol and after drying is resuspended in a suitable buffer.

2.3.1.9. DNA and RNA electrophoresis

This technique is used to purify and visualize DNA and RNA prior or after other applications. First, an agarose gel of the desired percentage is prepared by mixing agarose with TBE (Tris-Borate-EDTA) buffer (45 mM Tris, 45mM boric acid and 1 mM EDTA). After heating in the microwave until the agarose is dissolved, the solution is chilled and Redsafe reagent is added (Intron #21141). Redsafe is a reagent that permits the visualization of nucleic acids due to the green fluorescence that emits upon DNA and RNA binding. Then, this mixture is solidified in an electrophoretic chamber and after addition of TBE buffer, the DNA and RNA samples containing orange-glycerol are loaded. Gels are typically run at 80V during 1 hour and visualization is achieved by using an UV-transilluminator.

2.3.1.10. 4C-seq

This procedure is a mixture of classical molecular biology techniques with cutting edge technology. Circular chromatin conformation capture followed by sequencing (4C-seq) is used to determine the regions of the genome that contact a region of the genome selected by the researcher called viewpoint.

The protocol that we performed is based on the ones published in (Splinter et al., 2012) and (Stadhouders et al., 2013) (Figure M3) and it was carried out in the Centre of Molecular Medicine in Cologne under the supervision of Dr. Rada-Iglesias and with the technical help of Dr. de la Cruz-Molina.

For this protocol, I started with 12×10^6 NSCs. Cells were fixed using 1% of formaldehyde during 30 minutes at room temperature. Fixing reaction was quenched with glycine 0.125M during 10 minutes. After washes with PBS, cell pellets were

Methodology

resuspended in 5mL of cytoplasmic lysis buffer (50mM Tris-HCl pH 7.5, 150mM NaCl, 5mM EDTA, 0.5% NP-40, 1% Triton X-100 and protease inhibitors) during 10 min on ice. Lysates were centrifuged for 5 minutes at 650 rcf and 4°C.

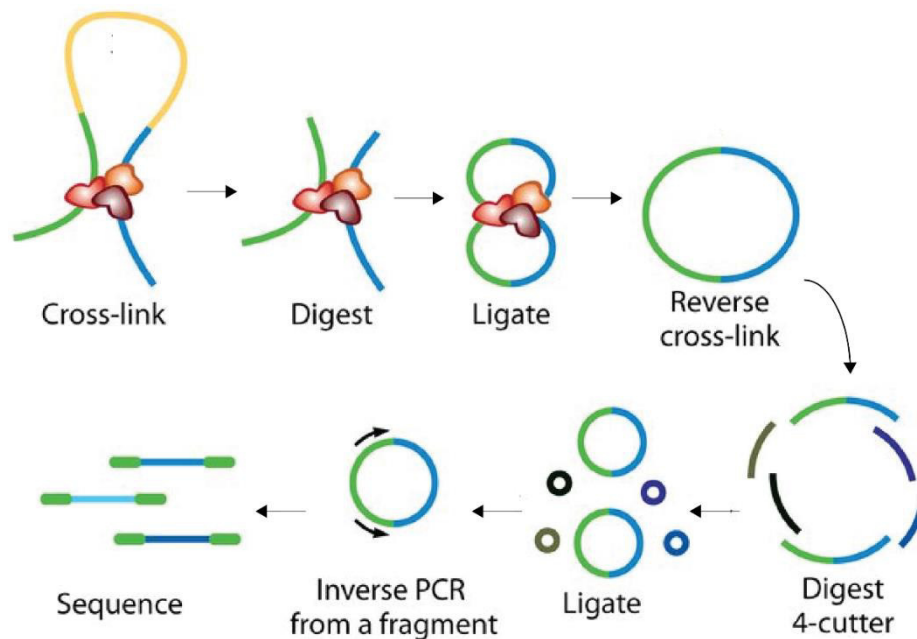


Figure M3. Cartoon describing the 4C-seq method. Adapted from (Fraser et al., 2015).

After removal of the supernatant, nuclei were resuspended in 0.5mL of NlaIII restriction enzyme buffer with 0.3% SDS and solutions were incubated at 37°C and 900 rpm for one hour. After that, Triton X-100 was added to a final concentration of 2% followed by 1 hour of incubation at 37°C and 900 rpm. These steps allow the opening of the chromatin to permit a better restriction digestion. Next, DNA is digested overnight at 37°C and 900 rpm with 400U of NlaIII (NEB, #R0125L). NlaIII is inactivated by adding SDS to a final concentration of 1.6% and incubation for 20 min at 65°C and 900 rpm. The digested chromatin is transferred to 50 mL tubes and 6.125 mL of 1.15X ligation buffer (50 mM Tris-HCl pH 7.6, 10 mM MgCl₂, 1 mM ATP, 1mM DTT) and 1% of Triton X-100 are added. This solution is incubated at 1 hour at 37°C and 1000 rpm. Next, digested chromatin is ligated with 100U of T4 DNA ligase (Invitrogen #15224041) for 8 hours at 16°C and then, this solution is treated with RNase A 1mg/mL during 45 minutes at 37°C. Decrosslinking step is performed by adding 1mg/mL of proteinase K and incubating at 65°C overnight. DNA was then purified by phenol-chloroform extraction followed by ethanol precipitation and resuspended in 100 µL of water. At this point, proper digestion and ligation are evaluated by visualizing the DNA in an agarose gel. Digested DNA generates a smear in the lane, in the ligated condition DNA must give just one band of around 10Kb (Figure M4).

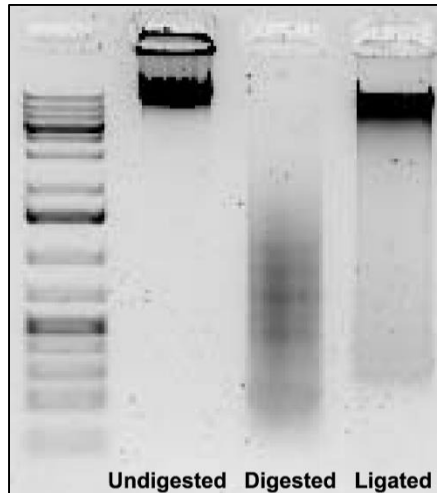


Figure M4. Agarose gel showing undigested, NlaIII digested and ligated DNA corresponding to the 4C-seq experiment.

Once the proper ligation has been tested, a second digestion with 50U of DpnII (NEB #R0543M) takes place at 37°C overnight. Enzyme and buffer are removed from the sample by phenol-chloroform extraction followed by ethanol precipitation and once purified, DNA samples are resuspended in 500 μ L of H₂O. Next, a second ligation is carried out by adding 200U of T4 DNA ligase in a final volume of 14mL of 1X ligation buffer. Mixture is incubated overnight at 16°C and after a last round of phenol-chlorophorm extraction and ethanol precipitation, the DNA is resuspended in 100 μ L of H₂O and purified with a QIAGEN PCR purification column (QIAGEN #28104). The

efficiencies of the second digestion and ligation are tested by electrophoresis, second digestion must generate a smear and second ligation too, but in this case it will end up abruptly at the size of 300 bps (Figure M5).

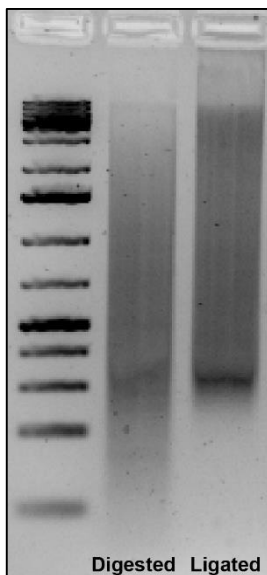


Figure M5. Agarose gel showing DpnII digested and ligated DNA corresponding to the 4C-seq experiment.

All the aforementioned protocol leads to the generation of a 4C-library that might be used to test different viewpoints. Once a viewpoint has been selected, inverse PCR reactions are performed using specific primers containing the adaptors for the Illumina sequencer (Table M3). It is worth noting, that for this PCR, a stable polymerase must be used. For this thesis it has been used the Expand long template PCR system (Roche #11681834001) with the following cycling conditions: 94°C 2 minutes, 30 cycles of 94°C 10 seconds – 60°C 1 minute – 68°C 3 minutes and 68°C 5 minutes. The product of this reaction is tested by agarose electrophoresis and if correct, it is sent for sequencing. The samples included in this doctoral thesis have been sequenced in the Erasmus Center for Biomics in Rotterdam, in an Illumina HiSeq2500 sequencer with a read depth of 100bps. Bioinformatic analysis details can be found in subsection 2.4.2.

Methodology

2.3.2. Protein-related techniques

In the following paragraphs, I will detail the protocols of the molecular biology techniques applied in this doctoral thesis whose targets are proteins.

2.3.2.1. Total protein extraction

For this doctoral thesis, we needed to extract proteins in order to detect the expression levels of a specific protein in a determined context.

To perform these protein extractions, we used a buffer called RIPA (Radioimmunoprecipitation assay buffer) that is highly astringent and breaks cytoplasmic as well as nuclear membranes (150mM NaCl, 1.0% NP-40, 0.5% sodium deoxycholate, 0.1% SDS and 50mM Tris, pH 8.0 and protease inhibitors). Approximately, 500 μ L of buffer were added to 6 \times 10⁶ cells, but the volume can vary according to concentration requirements. This suspension was incubated on ice during 20 minutes and then was centrifuged at maximum speed during 10 minutes at 4°C. The resulting supernatant contains the protein extract.

2.3.2.2. Nuclear-cytoplasmic extraction

For some applications protein extracts must be enriched in the nuclear or the cytoplasmic fraction of the cell.

To separate these compartments, we used the following protocol. Starting from 3 \times 10⁶ NSCs, pellets were resuspended in buffer A (10mM Tris-HCl, 10mM KCl, 1.5mM MgCl₂, and protease inhibitors) and kept on ice for 10 minutes. After centrifugation at 1500 rcf for 5 minutes, pellets were resuspended in buffer B (buffer A and 1% NP40) and incubated on ice for 5 minutes before centrifugation at 5000 rcf for 5 minutes. At this point, the supernatant contains the cytosolic fraction. Pellet is resuspended in buffer C (20mM Tris-HCl, 420mM NaCl, 1.5mM MgCl₂, 25% glycerol, 0.2mM EDTA, and protease inhibitors) by vortexing and incubating on ice, and the resulting suspension is centrifuged at the highest speed for 20 minutes at 4°C. The supernatant contains the nuclear fraction.

2.3.2.3. Protein quantification: Bradford

In order to measure the total amount of protein present in a sample, we performed the Bradford method.

This method is based on the reaction occurring between the proteins of the sample and the Coomassie Brilliant Blue G-250, a reagent that change its color depending on the protein concentration (Bradford, 1976).

To measure the concentration, 1 μ L of the protein extract is mixed with 1mL of Bradford solution (Bio-Rad # 5000001), after 3 minutes, the colorimetric reaction is measured in a spectrometer. Using the absorbance value of the sample and a BSA calibrate line, the concentrations of the samples are obtained.

2.3.2.4. SDS-Page electrophoresis

This classic procedure developed by Laemmli (Laemmli, 1970) serves to separate proteins in a gel according to their sizes. The SDS (sodium dodecyl sulfate) detergent provides with net negative charge to all the proteins in solution, thus ensuring that the migration will occur exclusively according to their sizes. This technique is normally followed by Western Blot that will be explained in the next subsection.

To prepare the samples, the protein extracts are mixed with Laemmli buffer (375mM Tris-HCl, 9% SDS, 50% Glycerol, 0.03% Bromophenol blue) and 5% of β -mercaptoethanol and they are heat during 5 minutes at 95°C. Then, these samples are loaded in a polyacrylamide gel that is subdivided in two gels: the stacking and the resolving. The stacking gel has a pH of 6.8 and a polyacrylamide concentration of 5%. On the other hand, the resolving gel has a pH of 8.8 and a variable polyacrylamide concentration depending on the sizes of the proteins to resolve. In Table M5 I detail the recipes for stacking and resolving gels of different percentages. After setting up the chamber with the gel, everything is covered with SDS-Page running buffer (25mM Tris-HCl, 192mM glycine, 0.1% SDS) and the power is set on at 25mA until all the sample has run through the gel.

Resolving gel	7%	10%	12%	15%	18%	Stacking gel	5%
H2O	5.2	4.8	4.3	3.6	2.8	H2O	7.25
40% Polyacrylamide	1.7	2.5	3	3.8	4.5	40% Polyacrylamide	1.25
1.5M Tris pH8.7	2.5	2.5	2.5	2.5	2.5	1.5M Tris pH6.8	1.25
10% SDS	0.1	0.1	0.1	0.1	0.1	10% SDS	0.1
10% Ammonium persulfate	0.1	0.1	0.1	0.1	0.1	10% Ammonium persulfate	0.1
TEMED	0.01	0.01	0.01	0.01	0.01	TEMED	0.01

Table M5. Recipes to polymerize the resolving gel of different polyacrylamide percentages, and recipe for the stacking gel of 5% of polyacrylamide.

2.3.2.5. Western Blot

After SDS-Page electrophoresis, we always performed the Western Blot procedure (Towbin et al., 1979) to specifically detect a protein in the sample that was resolved in the gel.

Methodology

This method starts by setting up a multilayered cassette in which from positive side to negative side are placed a sponge, Whatmann paper, a nitrocellulose membrane, the gel proceeding from the SDS-Page, Whatman paper and another sponge. This cassette is introduced in a chamber and the set is covered by Transfer buffer. It is important to note, that along this thesis we required two different transfer buffers. For regular proteins, the buffer was composed of 25mM Tris-HCl, 192mM glycine, 0.05% SDS and 10% methanol. For JMJD3 and CHD8 western blots, the transfer buffer contained 0.1% of SDS instead of 0.05%. This additional SDS increases the transferability of proteins of high molecular weight (MW) such as JMJD3 and CHD8 (240KDa and 290KDa respectively).

After chamber, cassette and buffer are set, the power source is turned on at 80V during 90 minutes for regular proteins, and during 140 minutes for JMJD3 and CHD8. At the end, proteins will be transferred to the nitrocellulose membrane.

To specifically identify a protein in the nitrocellulose membrane we first blocked the membrane with milk during 1 hour at room temperature. Then, the membrane is washed three times with PBS-Tween 0.1% and is incubated with specific primary antibodies (Table M2), usually overnight, at 4°C. Next, the membrane is washed again with PBS-Tween 0.1% and incubated with a secondary antibody (Table M2) during 1 hour at room temperature. At this point the western blot is prepared to be revealed.

During the experimental work, we have used chemoluminescent and fluorimetric methods to reveal western blots. Chemoluminescence has been used to detect proteins that are less abundant or for coimmunoprecipitations, where we expect to have low amount of protein, for this type of revealing method we used secondary antibodies bound to HRP. On the other hand, fluorescence has been used to reveal abundant proteins or proteins whose antibodies are really sensitive, for this method we used secondary antibodies bound to a fluorophore. To reveal membranes with chemoluminescence it is used the luminol-based Immobilon Western kit (Millipore #WBKLS0500) followed by revealing in an automated processor in a dark room. For fluorescence, revealing is performed in an LI-COR Odyssey scanner. At this point, visualization of the detected proteins permits the evaluation of results.

2.3.2.6. Size exclusion chromatography

In this thesis, this procedure has been used to separate a protein extract in different fractions containing protein complexes of different MW. If by western blotting different proteins are detected in an eluate, it potentially indicates that these proteins could be forming part of a complex.

To detect potential interacting proteins, we used whole NSCs protein extracts in a Superose-6 10/300 gel filtration column (GE Healthcare) on an AKTA purifier system (GE Healthcare). These extracts are performed as explained above but with a buffer

called IPH (50mM Tris-HCl pH 7.5, 150mM NaCl, 5 mM EDTA, 0.5% NP40), this buffer is less astringent and does not interfere with the gel-filtration column. Upon addition of the sample to the column, we collected different fractions and finally these fractions were used for protein detection by Western Blot.

2.3.2.7. Coimmunoprecipitation

This procedure is utilized for the detection of direct or indirect interactions between proteins. The method is based on an immunoprecipitation experiment, that consists in the binding of an antibody to a target in order to enrich the amount of target in a fraction. This experiment can be performed with overexpressed proteins or with endogenous proteins.

In this doctoral thesis we performed coimmunoprecipitations with overexpressed proteins using proteins extracted with IPH (50mM Tris-HCl pH 7.5, 150mM NaCl, 5 mM EDTA, 0.5% NP40). For endogenous coimmunoprecipitations, we used nuclear extracts performed with the protocol detailed in subsection 2.3.2.2, additionally treated with 0.3U/ μ l of Benzonase (Millipore #70746-4) during 5 hours on ice. The Benzonase is an enzyme able to degrade DNA and RNA, thus facilitating the release of proteins from the chromatin.

After generation of these protein extracts, a specific antibody is added to the extract to immunoprecipitate the target protein. Importantly, we always run a parallel reaction with IgGs of the same species in order to control the unspecific binding of the sample to the IgG *per se*. After overnight incubation at 4°C, protein A (Millipore #16-125) or protein G (Millipore #IP04) agarose beads (depending on the primary antibody specie) are included in the sample to bind to the primary antibody-protein complex. After 4 hours of incubation at 4°C, the beads are recovered and eluted by adding Laemmli buffer and shaking at 900 rpm and 55°C during 15 minutes. Finally, the Laemmli buffer containing the eluted proteins is recovered and denatured as explained in SDS-Page subsection, to proceed with electrophoresis and Western Blot.

2.3.2.8. Chromatin immunoprecipitation

This procedure permits the immunoprecipitation of proteins directly or indirectly bound to the chromatin, and the identification of the DNA bound by qPCR. This experiment has been essential during my thesis to evaluate the binding of our proteins of interest to the chromatin.

For every ChIP, 6×10^6 NSCs were fixed with 1% of formaldehyde during 10 minutes, this fixation was stopped by adding 0.125M of glycine during 5 minutes. Then, cells were lysed in 1% SDS lysis buffer (1% SDS, 10mM EDTA pH8.0, 50mM Tris-HCl pH8.1) to proceed with chromatin sonication to obtain fragments of around 300 bps

Methodology

of DNA to ensure resolution in the DNA detection. The sonication step was performed in a Bioruptor sonicator with variable parameters due to the inconsistency of the Bioruptor performance, example parameters are 30 cycles of 30 seconds on and 30 seconds off at high potency. Next, chromatin is purified by centrifugation at maximum speed during 10 minutes and the recovered supernatant is used for a sonication test in which I evaluate the correct size of the chromatin fragments.

Once chromatin is properly shredded, the immunoprecipitation step starts by diluting the chromatin tenfold with immunoprecipitation buffer (1% Triton X-100, 2mM EDTA, 150mM NaCl and 20mM Tris-HCl pH8.0) and adding the specific antibody for the protein of interest. As in coimmunoprecipitations, an identical parallel reaction is run using an unspecific IgG. After overnight incubation at 4°C the antibody-protein complexes are captured using Magna ChIP magnetic beads (Millipore #16-661) during 4 hours at 4°C. Next, antibody-protein-DNA complexes are sequentially washed with buffers TSEI (0.1% SDS, 1% Triton X-100, 2mM EDTA, 20mM Tris-HCl pH8.0 and 150mM NaCl), TSEII (0.1% SDS, 1% Triton X-100, 2mM EDTA, 20mM Tris-HCl pH8.0 and 500mM NaCl), TSEIII (0.25M LiCl, 1% NP40, 1% sodium deoxycholate, 1mM EDTA and 10mM Tris-HCl pH8.0) and TE (10mM Tris-HCl pH 8.0 and 1mM EDTA) and then eluted using elution buffer (1% SDS, 0,1M NaHCO₃) during 15 minutes.

At this point, DNA must be recovered in order to be analyzed by qPCR, to do that, samples are subjected to decrosslinking overnight at 65°C. Then, the samples are treated with 1mg/mL of RNase A during 30 minutes at 37°C and with proteinase K at 55°C during 2 hours. After this, DNA is purified by phenol-chloroform extraction followed by ethanol precipitation. Usually DNA is resuspended in 50µL of H₂O.

Lastly, ChIP DNA is analyzed by qPCR as explained in subsection 2.3.1.6, using specific primers (Table M3). Percentage of input material is used for the quantification of the immunoprecipitated DNA with respect to the total starting chromatin.

2.3.2.8.1. Chromatin immunoprecipitation of JMJD3

In this subsection I would like to detail changes on the general ChIP protocol that are required for the ChIP of JMJD3.

It starts with the fixation of 12×10^6 NSCs with two different fixing agents. First, 2mM Di(N-succinimidyl) glutarate (DSG) (Sigma #80424) during 45 minutes and then 1% of formaldehyde during 20 minutes. Quenching is performed with 0.125M glycine during 10 minutes. Cells are lysed in 1.3mL of mild buffer (0.1% SDS, 1% Triton X-100, 150mM NaCl, 1mM EDTA, 20mM Tris pH 8.0) and sonication is performed during 45 cycles in a Bioruptor. The mild nature of the buffer and the usage of two fixing agents limits the sonication capacity, so this ChIP will have less resolution. Non-diluted chromatin is used for immunoprecipitations and washes are performed

in mild wash buffer 1 (0.1% SDS, 0.1% sodium deoxycholate, 1% Triton X-100, 150mM NaCl, 1mM EDTA, 20mM HEPES), mild wash buffer 2 (0.1% SDS, 0.1% sodium deoxycholate, 1% Triton X-100, 0.5M NaCl, 1mM EDTA, 20mM HEPES), wash buffer 3 (250mM LiCl, 0.5% sodium deoxycholate, 0.5% NP-40, 1mM EDTA and 20mM HEPES) and wash buffer 4 (1mM EDTA, 20 mM HEPES). The rest of the protocol steps are performed likewise.

2.3.2.9. Indirect immunofluorescence

Along this thesis, I have collaborated with Simona Iacobucci, a member of the lab with expertise on immunofluorescence and microscopy to obtain images of the staining of specific proteins in determined conditions.

For this protocol, a coverslip with grown NSCs is fixed for 20 minutes at room temperature in 4% of paraformaldehyde and permeabilized with PBS-Triton X-100 0.1% before blocking at room temperature for 1 hour in 1% BSA (Bovine serum albumina). Then, the coverslip is incubated overnight at 4°C with primary antibodies (Table M2).

To visualize the proteins under the microscope, cells are incubated for 2 hours at room temperature with Alexa-conjugated secondary IgG antibodies (Table M2) and 0.1ng/μL DAPI (ThermoFisher #D1306). These antibodies emit fluorescence that will be detected in the microscope.

Images were captured by a Leica SP5 confocal microscope using LAS-AF software and quantification was achieved by counting cells in randomly located fields.

In the case of the neural differentiation experiments, number of neurites per cell and percentage of uni/bipolar or multipolar neurons were quantified by direct counting of 10 randomly selected fields considering as multipolar the neurons with more than three neurites.

3. Bioinformatic methods

Along my doctoral thesis I have collaborated with the team of Dr. Xavier de la Cruz for some bioinformatic analysis. Nonetheless, the bioinformatic work presented in this thesis has been performed by me using the menu-based bioinformatic platform Galaxy (Afgan et al., 2016) (excepting for the RPMs calling of 4C-seq experiment, that was performed by Dr. Rada-Iglesias).

In the following subsections I will detail which tools were used for every analysis.

3.1. ChIP-seq data acquisition and analysis

During this doctoral thesis we have used downloaded previously published ChIP-seqs for further analysis from GEO (accessions in Table M4). It is worth noting, that

Methodology

for all of the accessions excepting for ASCL1, SMAD3 and JMJD3, bed files were already available and were used for enhancer identification and Venn diagram construction.

In the case of ASCL1 ChIP-seq analysis, mapped reads were downloaded for input and ASCL1 immunoprecipitation and we called peaks using the tool MACS (Zhang et al., 2008) with the following parameters: p-value 1e-18, effective genome size 1865500000 and MFOLD 32. For SMAD3 and JMJD3 an identical pipeline was followed excepting for the p-value that was higher in this case (1e-7).

3.1.1. Enhancer identification with ChIP-seq data

As I explained in the introduction, enhancers are regions that contain specific chromatin features. In this doctoral thesis enhancers were identified by overlapping ChIP-seqs of these chromatin features.

For enhancer identification the *-intersectbed* command from BEDTools (Quinlan and Hall, 2010) with a minimum of overlapping base pairs of 50 was used.

3.1.2. Putative regulated gene assignment

Which gene is regulated by every enhancer is an intriguing question in the field. Historically, enhancers lacking further functional studies have been assigned to the nearest gene for other analysis. However, results from chromatin conformation capture experiments have shown that this is usually not the case.

In this thesis, we have used two enhancers (*Ctgf* and *Neurog2*) whose putative target gene has been discovered in functional assays as shown in Enhancer Vista Database (Visel et al., 2007). We have also studied other three enhancers (*Nrip3*, *Chic2* and *Tle3*) that have been assigned to their putative regulated genes by fetching the closest nearest gene.

3.1.3. Enhancer gene ontology analysis: GREAT

Enhancers can be assigned to specific gene ontology categories by using the software GREAT (McLean et al., 2010). This software assigns biological process categories of the nearby genes to the included enhancers.

To perform this analysis, we introduced the list of enhancers regulated by SMAD3, JMJD3 and ASCL1 and selected the genomic build mm9. For a background region we selected the whole genome option.

3.1.4. Venn diagrams construction

For Venn diagrams construction we have used R studio. A p-value statistically evaluating the overlaps accompanies some diagrams. These p-values were obtained by generating a comparable random sample with *-shuffleBed* and applying afterwards an equal proportions test (more details in section 4.5).

3.1.5. Heatmaps construction

In this doctoral thesis, we have used heatmaps to represent the colocalization of proteins on enhancers.

To generate these heatmaps, we firstly computed a matrix using the coordinates of the enhancers (as bed files) and the ChIP-seq intensities of the proteins of interest on those enhancers (as bigwig files). This matrix is generated using the *-computeMatrix* command from Deeptools (Ramirez et al., 2016). Then, the matrix is plotted in a heatmap with *-plotheatmap* tool, this tool allows the formatting and obtaining of the graphic, it also belongs to de Deeptools package (Ramirez et al., 2016).

3.1.6. ChIP-seq capture obtaining

Along the Results section I will include captions of the ChIP-seq signal of different proteins to observe specific genomic coordinates.

These captions were obtained by loading bigwig files of the different experiments in the IGV genome browser from the Broad Institute (Robinson et al., 2011) or in the UCSC browser from the University of California Santa Cruz (Karolchik et al., 2004).

3.2. 4C-seq bioinformatic analysis

Bioinformatic analysis for the 4C-seq experiment was performed by Dr. Rada-Iglesias in the Centre of Molecular Medicine in Cologne. I will briefly comment some details here.

After sequencing, Illumina adaptor sequences are removed from the obtained sequences, the first 36 bps are mapped to the *Mus musculus* genome version mm9 with Bowtie (Langmead et al., 2009). Then, mapped reads are analyzed with R3C-seq (Thongjuea et al., 2013) in order to generate bedgraphs with reads per million (RPM) for further analysis.

Methodology

3.2.1. Box-plot construction

To show the differences between signals in the conditions tested in the 4C-seq, we constructed a box-plot with R Studio. The p-values shown in the graphic are the result of a paired Wilcoxon test performed after assessing the linearity of the model by a Saphiro-Wilk test.

3.3. RNA-seq and microarray comparison

In order to compare the genes that were coregulated by SMAD3, JMJD3 and CHD8, we used previously published data by us and others (Table M4). Concretely, we crossed the genes activated by SMAD3 and JMJD3 that appeared in our previously published microarray (Estaras et al., 2012) with the downregulated genes in the RNA-seq of a CHD8 KD experiment (Durak et al., 2016). To evaluate the significance of the percentage that overlapped we performed an equal proportions test against a random set of genes.

4. Statistical analysis

In this last subsection of the Methodology I will provide details on which statistical tests have been applied to assess the reproducibility and significance of the results.

4.1. Sample size

As a general rule, experiments have been performed in triplicate. In specific cases like validation of microarray data by qPCR or immunostainings for control markers, the number of replicates have been two. Another exception is the 4C-seq experiment, this experiment just shows the results of one replicate, at the moment we are replicating it.

4.2. Standard deviation and standard error of the mean

Along the experimental work, the graphics corresponding to experiments that fit a linear model have been represented as the mean. Error bars correspond to the standard deviation (SD) in the case of indirect immunofluorescences and RT-qPCR assays, and to the standard error of the mean (SEM) in the case of ChIPs. The numeric values have been calculated with Microsoft Excel software.

4.3. Student's t-test

To assess the significance of the results that follow a linear model we have performed the Student's t-test. We have established that an experiment is statistically significant when within a 95% of confidence the result represents a true hypothesis. Asterisks represent the different p-values resulting from this test and calculation of the values was carried out with Microsoft Excel software.

* p-value < 0.05; ** p-value < 0.01; *** p-value < 0.001

4.4. Saphiro-Wilk test

This test has been performed to assess the normality of the 4C-seq signals in order to know which statistical test should be performed to evaluate the significance. Calculation was performed with R Studio using the `saphiro.test` function.

4.5. Equal proportion tests

The significance of the data that follow a non-parametric model has been assessed with the equal proportion tests. This test permits the evaluation of the differences in the proportions of a specific feature in different groups. Along the thesis I have used this test to analyze the significance of Venn Diagrams. Calculation was performed with R Studio using the `prop.test` function.

4.6. Paired Wilcoxon test

This non-parametric test is used to assess differences on the mean of matched data that are not normally distributed. We have performed this test to analyze the significance of the differences in the 4C-seq signals between the different conditions.

Concluding the methodology section, I would like to remark that I am at disposal of anyone consulting these methods with doubts or comments that I could address. In the next section I will detail the results obtained during these years.

Results

Along my doctoral thesis I have obtained results regarding the activation of enhancers upon TGF β -signaling in the linear and 3D-structure of the chromatin. Previous results from our lab elucidated the role of the TGF β -pathway and JMJD3 in the transcriptional initiation and elongation of neural genes (Figure I23) (Estaras et al., 2012; Estaras et al., 2013), I started this doctoral thesis with the goal of uncovering the participation of this signaling pathway and JMJD3 in the activation of enhancers. Some of the results have already been published in (Fueyo et al., 2018).

1. Epigenetic identification of neural enhancers

I started this work by combining different chromatin features in order to identify enhancers in mouse NSCs. Using previously published epigenomic data (Table M3) and following the well-established enhancer identification criteria (Rada-Iglesias et al., 2011), we analyzed the enhancer landscape in NSCs. For that purpose, we combined ChIP-seq of p300, H3K4me1, H3K4me3, H3K27ac and H3K27me3 generated in E12.5 cortex progenitors (Sun et al., 2015) to classify the enhancers in active enhancers: presence of p300, H3K4me1 and H3K27ac; poised enhancers: binding of p300 and presence of H3K4me1 and H3K27me3 modifications; primed enhancers: marked with H3K4me1 and p300 and heterogeneous enhancers: presence of p300, H3K4me1 and both H3K27ac and H3K27me3.

We identified 3020 putative enhancers in NSCs (defined as presence of p300, H3K4me1 and low levels of H3K4me3) (Figure R1).

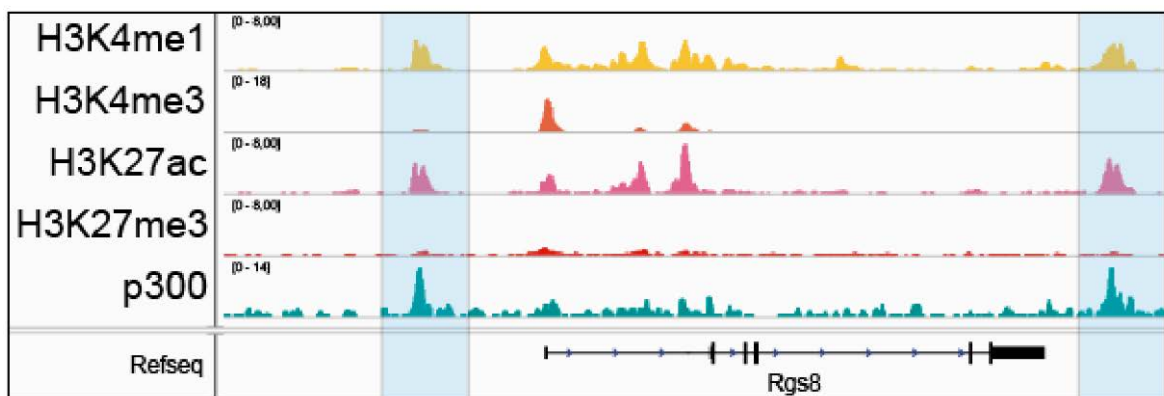


Figure R1. IGV capture showing the chromatin landscape of enhancers (highlighted in blue). Tracks display ChIP-seq in NSCs of H3K4me1, H3K4me3, H3K27ac, H3K27me3 and p300.

Among these 3020 enhancers, we found 1174 (38.9%) active enhancers, 158 (5.2%) poised enhancers, 1549 (51.3%) primed enhancers and 139 (4.6%) heterogeneous enhancers (Figure R2).

Results

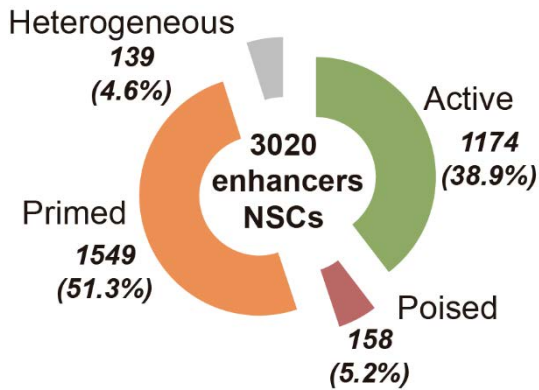


Figure R2. Number, percentage and type of identified enhancers in NSCs.

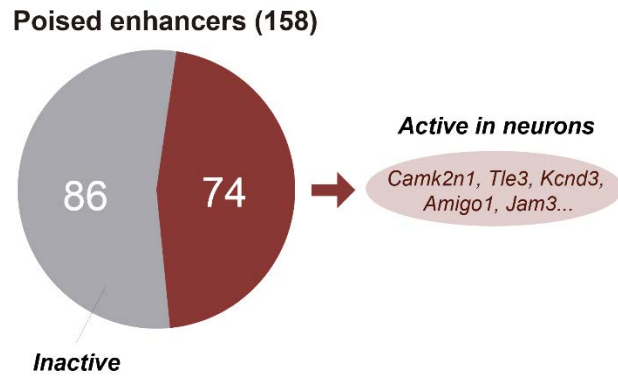


Figure R3. Number of poised NSCs enhancers that become active in neurons. Some neuronal genes associated to these enhancers are indicated.

Poised enhancers have been found in ESCs. Upon differentiation, some of them become active (Rada-Iglesias et al., 2011). In order to understand the relevance of the poised enhancers at this stage in NSCs, we analyzed how many of them become activated in terminally differentiated neurons. By comparing a previously published dataset of active enhancers in neurons from E16.5 embryos (Kim et al., 2010), we found that 74 out of the 158 poised enhancers in NSCs (46.8%) became active after the stage-transition, classical neuronal genes associated to these enhancers are included in this list (*Camk2n1, Kcnd3, Tle3, Amigo1...*) (Figure R3).

2. TGF β -regulation of neural enhancers

In this subsection I will include results on how TGF β pathway regulates neural enhancers.

2.1. SMAD3 is bound to neural enhancers upon TGF β treatment

Once identified the putative enhancers in NSCs, we continued the analysis by looking at the implication of the TGF β signaling cascade in the regulation of these enhancers. To do that, we analyzed the SMAD3 genomic distribution in NSCs upon TGF β -signaling with our previously published SMAD3 ChIP-seq (Table M4) (Estaras et al., 2012). After combining the data, we identified 1154 (38.2%) enhancers that were bound by SMAD3 after TGF β stimulation (p-value 2.2e-16), reinforcing the potential role of TGF β -signaling in NSCs (Figure R4).

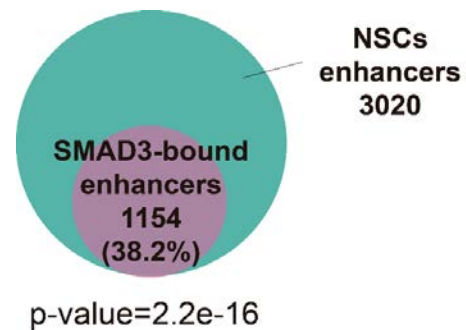


Figure R4. Venn diagram showing the number of enhancers bound by SMAD3 in NSCs treated with TGF β for 30 minutes. p-value indicates the result of an equal proportions test.

2.1.1. Interaction between the proneural factor ASCL1 and SMAD3

As it was introduced, SMAD3 genome-binding pattern shares many targets with cell-type specific TFs important for cell identity in ESCs and muscle cells (Mullen et al., 2011). The bHLH TF ASCL1 is a cell-identity molecule in NSCs that has been shown to exert a critical role regulating neural gene expression by binding and opening the chromatin structure at enhancers (Raposo et al., 2015).

In order to answer whether ASCL1 could be providing specificity to the SMAD3 binding, we first tested the physical interaction between ASCL1 and SMAD3 in NSCs. Figure R5 shows that endogenous ASCL1 and SMAD3 interact together. Next, we analyzed the colocalization of SMAD3 and ASCL1 at the genome-wide level using previously published ChIP-seq data from NSCs (Estarras et al., 2012;

Raposo et al., 2015). Doing that, we identified 762 (66%) SMAD3-bound enhancers that also contained ASCL1 (p-value 2.2e-16) (Figure R6). To see how the peaks of SMAD3 and ASCL1 colocalized, we constructed heatmaps showing the ChIP-seq intensities of the proteins in the coregulated enhancers. Figure R7 shows that SMAD3 and ASCL1 peaks significantly colocalize at the neural enhancers.

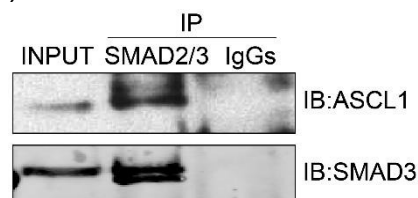


Figure R5. Immunoblots showing the coimmunoprecipitation of ASCL1 and SMAD3.

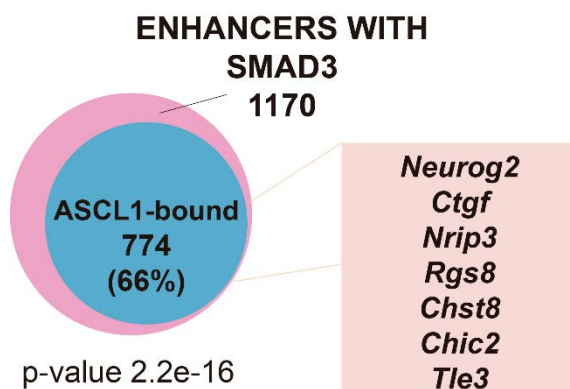


Figure R6. Venn diagram showing the number of SMAD3-bound enhancers that contain ASCL1. The names of genes putatively regulated by these enhancers are indicated.

Following these findings, we tested if TGF β and ASCL1 had any functional association. To do that, ASCL1 protein levels were transiently depleted by transduction of lentivirus containing specific ASCL1 shRNAs into the NSCs (Figure R8).

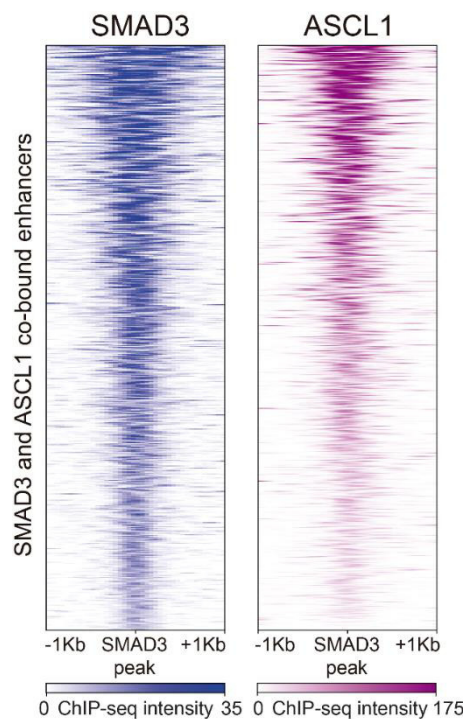


Figure R7. Heatmap representation of SMAD3 and ASCL1 binding on the SMAD3 and ASCL1 co-bound enhancers.

Results

The transient removal of ASCL1 did not affect SMAD3 levels, proliferation rate or differentiation status of the NSCs (Figures R8 and R9).

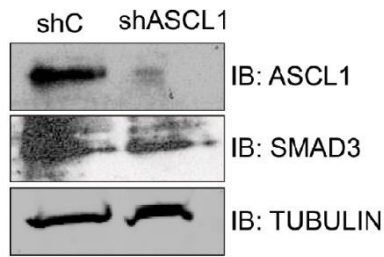


Figure R8. Immunoblots showing the levels of SMAD3 in the shC and shASCL1 cell line.

After viral transduction, the expression of some TGF β -responsive genes associated to enhancers bound by ASCL1 and SMAD3 was tested by qPCR in the shASCL1 and control cell lines. We chose *Ctgf*, *Nrip3* and *Neurog2* genes that cover the spectrum of transcriptional levels in our previously published microarray data (Estaras et al., 2012). *Ccne3* gene, which does not respond to TGF β , was used as a negative control.

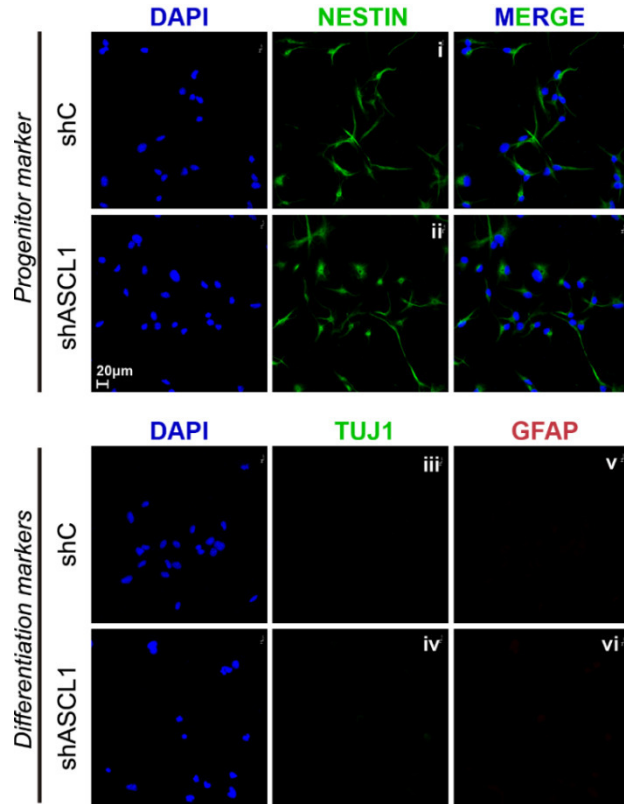


Figure R9. Immunostainings showing expression levels of neural progenitor marker NESTIN (top) and neuronal and astrocytic differentiation markers TUJ1 and GFAP (bottom) on shC and shASCL1 cell lines.

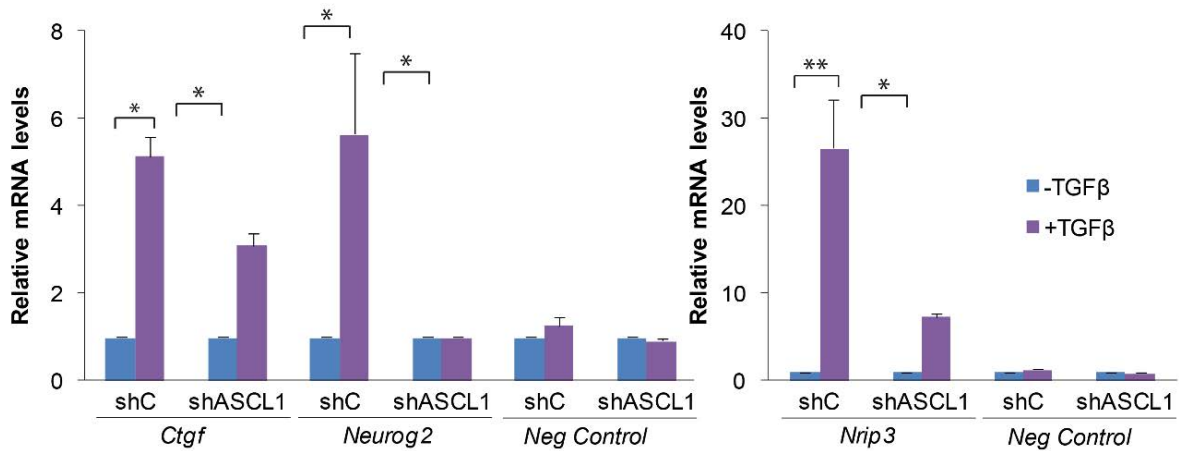


Figure R10. mRNA levels of TGF β responsive genes in the shC and shASCL1 cell lines. Data are normalized to Rps23 housekeeping gene. Error bars indicate SD. *p < 0.05; **p < 0.01 (Student's t-test).

Results in Figure R10 demonstrate that ASCL1 is essential to fully activate TGF β -targets in NSCs. All things considered, these data point to the proneural factor ASCL1 as a new partner of SMAD3 in forebrain enhancer recognition.

2.2. TGF β -dynamic activation of neural enhancers

Until now, I have described how SMAD3 is bound to enhancers pre-marked by ASCL1. In this subsection I will detail the activating role of TGF β at enhancers.

2.2.1. Activation evaluated with eRNA transcription

To analyze the activation of these cis-regulatory regions, we evaluated the transcription from the enhancers, measuring eRNAs which serve as a readout of enhancer activation (Li et al., 2016). For posterior analysis, the *Nrip3*(-3.5), *Ctgf*(-102) and *Neurog2*(-6) enhancers were selected (numbers in brackets indicate distances in Kb to the regulated promoters). These enhancers are associated to the previously analyzed genes (Figure R10) and represent the two major functional enhancer categories: active (*Nrip3* and *Ctgf*) and poised (*Neurog2*). In NSCs, eRNAs were detected as early as 30 minutes after TGF β stimulation. Interestingly, there is a strong correlation in the magnitude of the transcription between the eRNAs and their corresponding mRNAs (Figures R11).

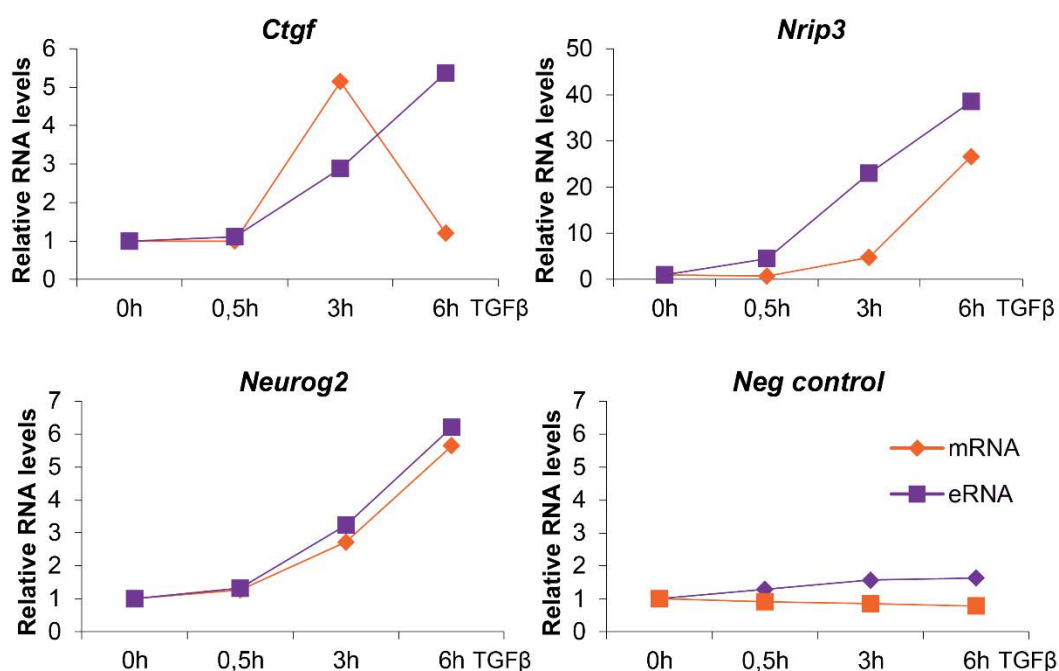


Figure R11. Kinetic curves showing the relation between eRNA and mRNA transcription upon TGF β treatment. *Fabp4* eRNA and gene were used as negative controls. Normalized with *Rps23* housekeeping gene. Error bars represent SD. *p < 0.05; **p < 0.01 (Student's t-test).

Results

In order to confirm the dependency on TGF β of the enhancer activation, we generated a shSMAD3 cell line (Figure R12) and tested the transcription of eRNAs (Figure R13). In concordance with the aforementioned results, eRNAs were hardly induced in shSMAD3 cells compared to the control cell line. Similarly, mRNA of the associated genes was not produced upon TGF β addition in the shSMAD3 NSCs (Figure R13).

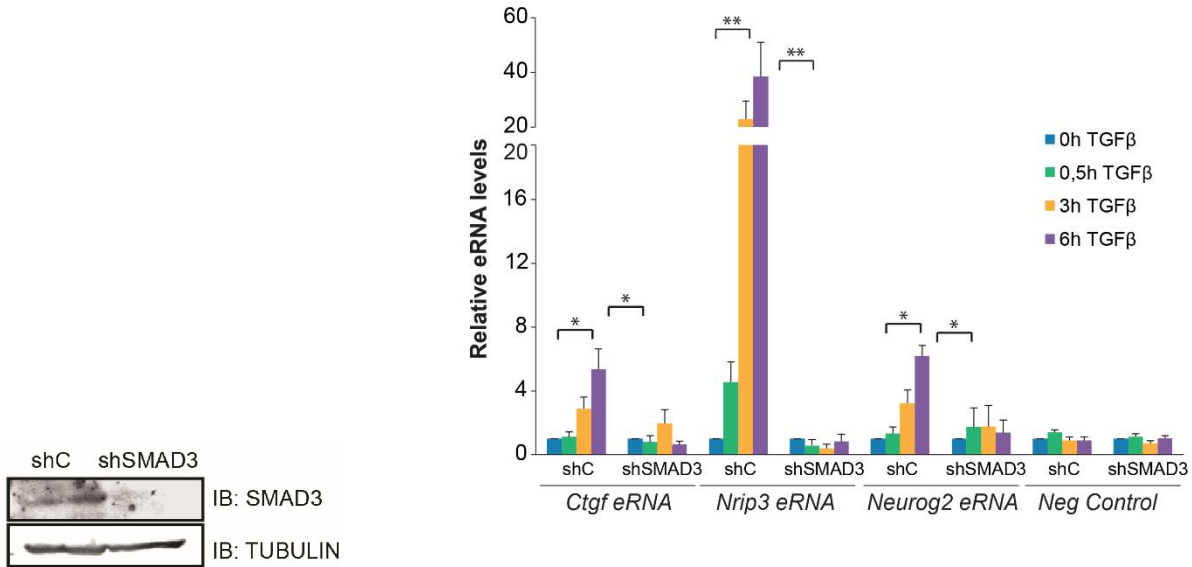


Figure R12. Immunoblot showing the levels of SMAD3 in the shSMAD3 cell line versus the shC.

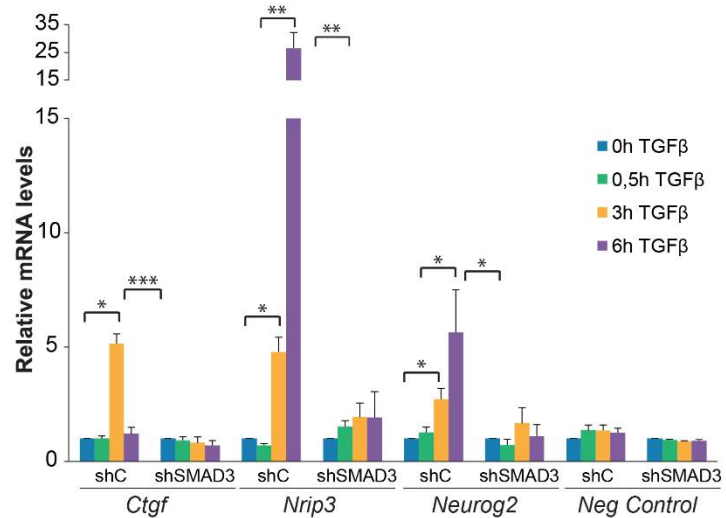


Figure R13. eRNA (top) and mRNA (bottom) transcription in shC and shSMAD3 cells treated with TGF β . Results are normalized with *Rps23* housekeeping gene. Negative control corresponds to the *Fabp4* eRNA and mRNA.

2.2.2. Activation evaluated with histone modification acquisition

Another way of measuring the enhancer activation is the analysis of the enhancer histone marks. H3K27ac is directly related to activation (Creyghton et al., 2010) and

H3K4me2 has been related to active enhancers in signaling contexts (Kaikkonen et al., 2013; Luyten et al., 2014).

In Figure R14, active enhancer histone marks upon TGF β signaling are shown, as a negative control we have used an intergenic region located at coordinates chr13:71467614-71467715 (mm9). As it is observed, TGF β increases the levels of these histone marks at the selected enhancers excepting for *Neurog2(-6)*, this result will be commented in the discussion section 2.2.2.

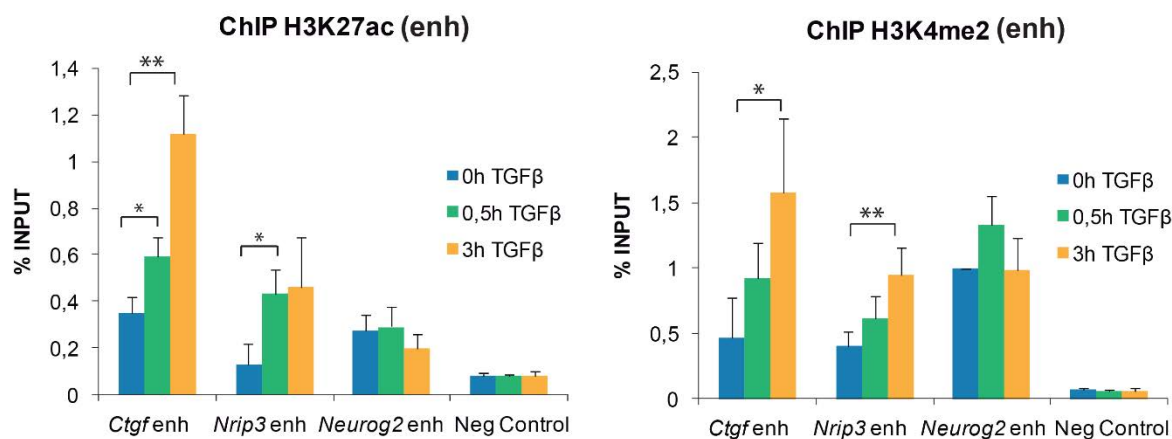


Figure R14. ChIP-qPCR showing the active enhancer histone mark levels (H3K27ac left and H3K4me2 right) upon TGF β signaling at the selected enhancers.

Enhancers are known to also contain H3K4me1 and lack H3K4me3 (Creyghton et al., 2010; Rada-Iglesias et al., 2011). For this reason, we tested these histone marks with ChIP-qPCR. Results show that at these enhancers, the H3K4 is majoritarily dimethylated and the low levels of H3K4me3 indicate that they are not promoters (Figure R15).

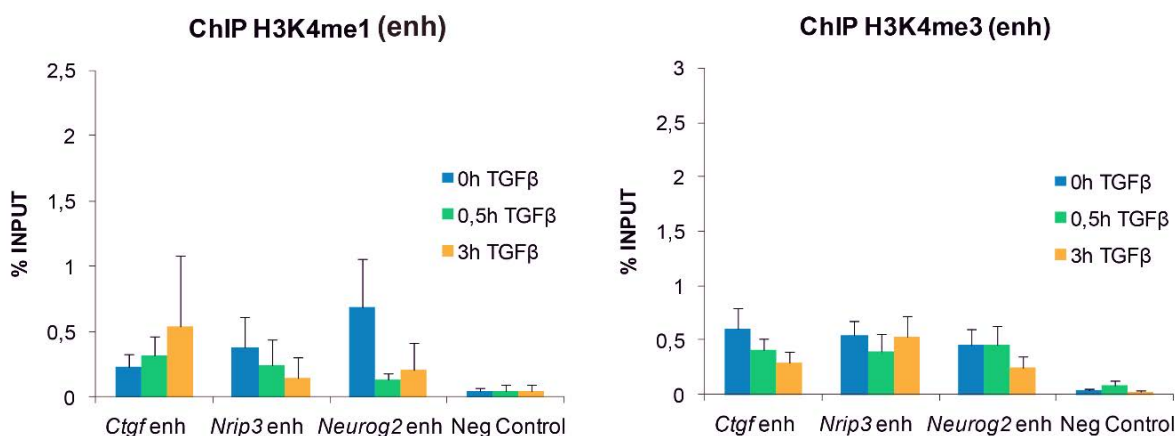


Figure R15. ChIP-qPCR showing H3K4me1 and H3K4me3 histone marks at the selected enhancers upon TGF β signaling.

Results

2.2.3. Activation evaluated with *in vivo* experiments

The above described findings support the idea of TGF β pathway activating relevant enhancers for neural development. Therefore, in collaboration with Stella Pappa, we tested whether SMAD3 regulates the analyzed enhancers in an *in vivo* model of TGF β -triggered neurogenesis, the chick embryo neural tube (Estaras et al., 2012; Garcia-Campmany and Marti, 2007).

To perform the experiments, we cloned the *Ctgf*(-102), *Nrip3*(-3.5) and *Neurog2*(-6) enhancers in a luciferase reporter vector that contains the SV40 promoter. Next, these constructions were EP *in ovo* in HH11-12 embryos. Together with the enhancers, we co-EP a constitutively active, pseudo-phosphorylated form of SMAD3, that mimicks the TGF β signaling (SMAD3-S/D), or a non-phosphorylatable form of SMAD3 that acts as a dominant negative (SMAD3 S/A). Figure R16 shows that the expression of the luciferase downstream the enhancers was increased by co-EP of the SMAD3-S/D constitutively active mutant, while the expression of the luciferase with the promoter alone (pGL3-empty) was not affected. On the contrary, co-EP of SMAD3 S/A, blocked the TGF β -induction of the analyzed enhancers. These results demonstrate that the investigated enhancers are activated *in vivo* in response to the TGF β signaling pathway (Figure R16).

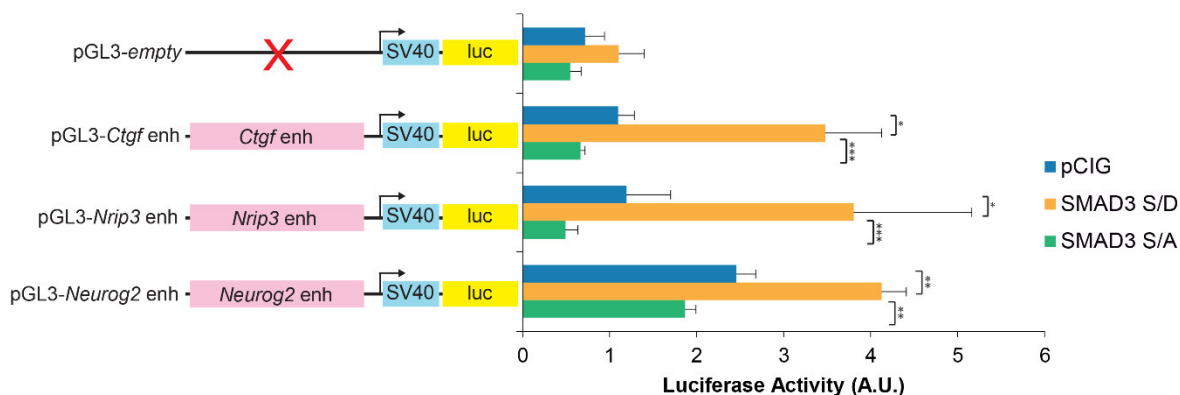


Figure R16. Luciferase experiments results showing the activation of the selected enhancers upon the overexpression of a constitutively active form of SMAD3 (SMAD3 S/D) or a dominant negative form of SMAD3 (SMAD3 S/A). Error bars indicate SD. *p < 0.05; **p < 0.01; ***p < 0.001 (Student's t-test).

3. JMJD3 cooperation in the SMAD3 and ASCL1 bound enhancers

Previous work of the laboratory has demonstrated that JMJD3 interacts and cooperates with SMAD3 at promoters to induce the TGF β neurogenic program *in vitro* and *in vivo* (Estaras et al., 2012; Estaras et al., 2013). One of the main goals of my doctoral thesis was to decipher whether JMJD3 could be cooperating with the TGF β pathway also at enhancers.

To address this hypothesis, we first identified the JMJD3-bound enhancers upon TGFβ using our previously reported ChIP-seq data (Estaras et al., 2012). Results in the left panel of Figure R17 show that JMJD3 was found in 30% of total forebrain enhancers. Interestingly, among these, 66.3% contained also ASCL1 and SMAD3 (Figure R17, right panel), showing a strong colocalization of the three proteins at cis-regulatory regions.

JMJD3 also showed a striking colocalization in terms of ChIP-seq intensity with SMAD3 and ASCL1. In Figure R18, heatmaps demonstrate that the three proteins co-bind at the enhancers.

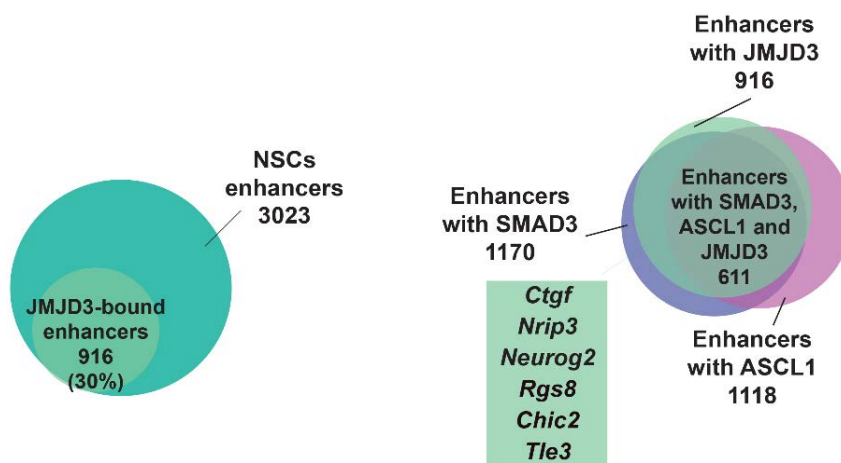


Figure R17. Venn diagrams showing the enhancers bound by JMJD3 (left) and the enhancers co-bound by SMAD3, ASCL1 and JMJD3 (right). Names in the green box correspond to the putatively regulated genes.

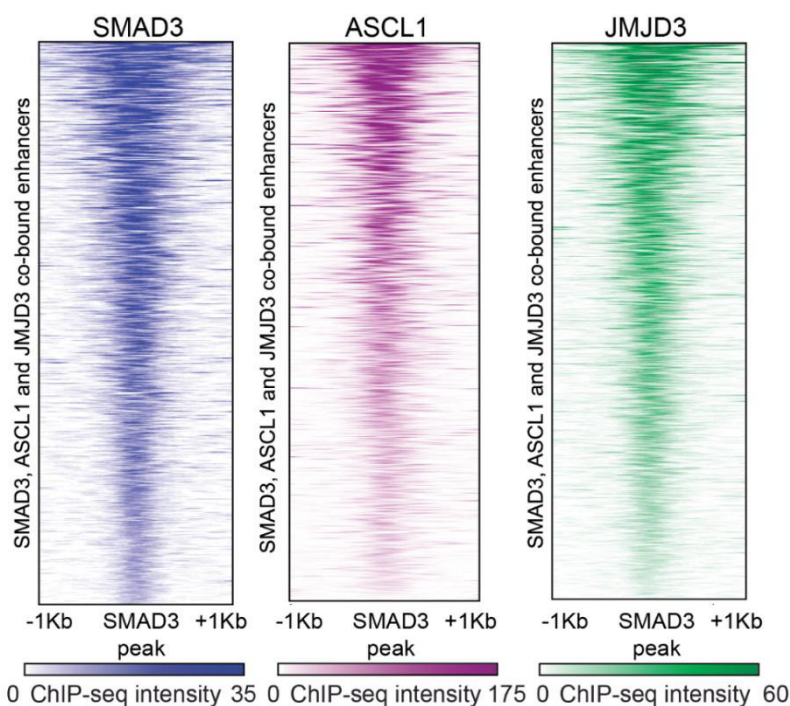


Figure R18. Heatmap representations of SMAD3, ASCL1 and JMJD3 binding on the SMAD3, ASCL1 and JMJD3 co-bound enhancers.

Results

SMAD3 and JMJD3 were found at the three types of enhancers (active, primed and poised). However, there is an increase in the percentage of active enhancers when the three partners, ASCL1/SMAD3/JMJD3, are occupying the genomic regions (p-value $1.1e-15$, equal proportions test between the ASCL1/SMAD3/JMJD3-bound enhancers and the total enhancers), suggesting that ASCL1 and JMJD3 contribute to enhancer activation in response to TGF β (Figure R19).

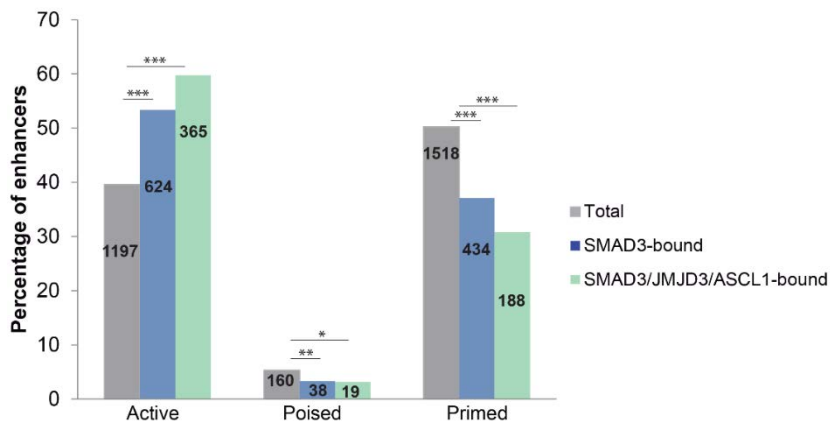


Figure R19. The percentage of active, poised and primed total enhancers, and those bound by SMAD3 or SMAD3/ASCL1/JMJD3 are depicted. Numbers inside the bars indicate absolute number of enhancers. p-values: *p < 0.05; **p < 0.01; ***p < 0.001.

Moreover, a gene ontology (GO) analysis of the enhancers bound by SMAD3/ASCL1/JMJD3 performed with GREAT (McLean et al., 2010) returned categories involved in neurogenesis, supporting the idea of these proteins regulating neuronal differentiation (Figure R20).

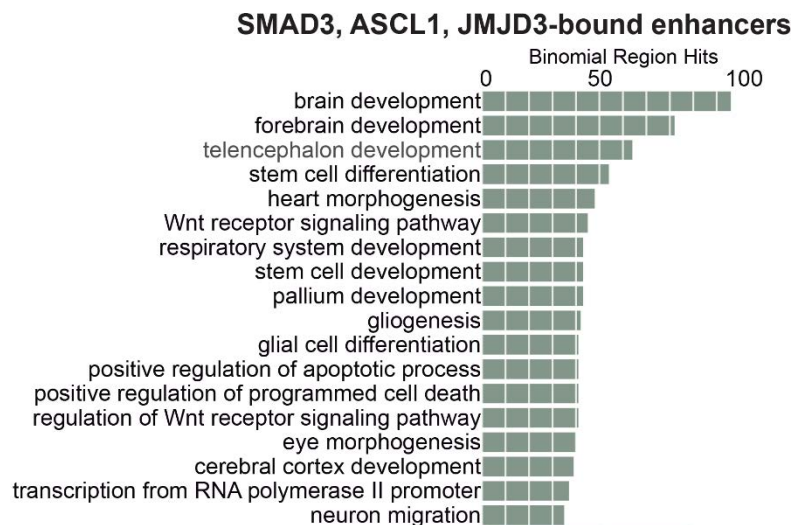


Figure R20. GREAT analysis showing GO Biological Process of the enhancers co-occupied by ASCL1/SMAD3/JMJD3.

3.1. Interaction between ASCL1 and JMJD3

As JMJD3 co-occupies enhancers together with ASCL1, we tried a coimmunoprecipitation experiment in HEK293T cells with overexpressed HA-ASCL1

and MYC-JMJD3. The immunoblots in Figure R21 show interaction between these proteins, supporting the idea that JMJD3, SMAD3 and ASCL1 are forming a functional complex at neural enhancers.

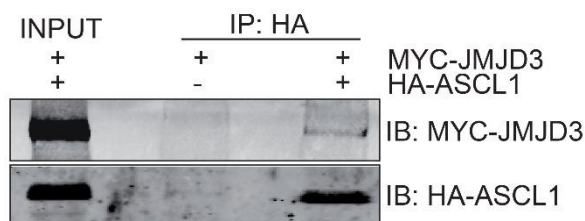


Figure R21. Immunoblots showing the coimmunoprecipitation experiment result between overexpressed JMJD3 and ASCL1.

3.2. Role of JMJD3 in the TGF β activation of enhancers

Once confirmed that JMJD3 was bound with SMAD3 and ASCL1 to neural enhancers, the contribution of this HKDM to the TGF β enhancer activation was evaluated.

First, we generated a KD cell line for JMJD3 using lentivirus containing specific shRNAs for JMJD3, immunoblot in Figure R22 displays an efficient depletion of this enzyme in NSCs. Additionally, we performed a RT-qPCR experiment in shC cells versus shJMJD3 to test the levels of the homologous HKDM, *Utx*. As it is observed in Figure R23, *Utx* does not compensate in terms of expression levels the lack of *Jmjd3*.

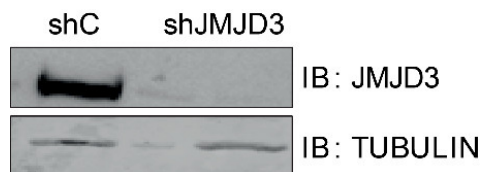


Figure R22. Immunoblot showing JMJD3 expression in the shC NSCs versus the shJMJD3 NSCs.

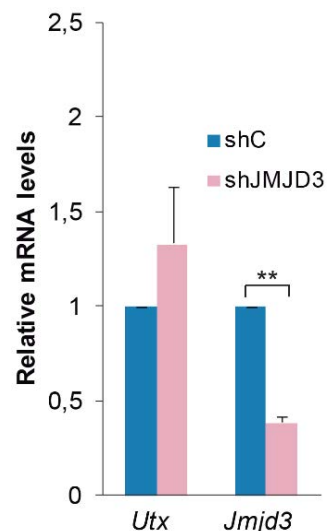


Figure R23. RT-qPCR results of the *Utx* expression, in shJMJD3 versus shC NSCs. Error bars indicate SD. **p < 0.01; (Student's t-test).

Results

3.2.1. Contribution of JMJD3 to enhancer activation evaluated by eRNA

With the shJMJD3 NSCs, we first performed RT-qPCRs to test the eRNA transcription of the *Ctgf(-102)*, *Nrip3(-3.5)* and *Neurog2(-6)* enhancers in response to TGF β . In the Figure R24 it is shown that the synthesis of eRNAs diminish in shJMJD3 NSCs upon TGF β stimulation in comparison with the shC cell line (Figure R24).

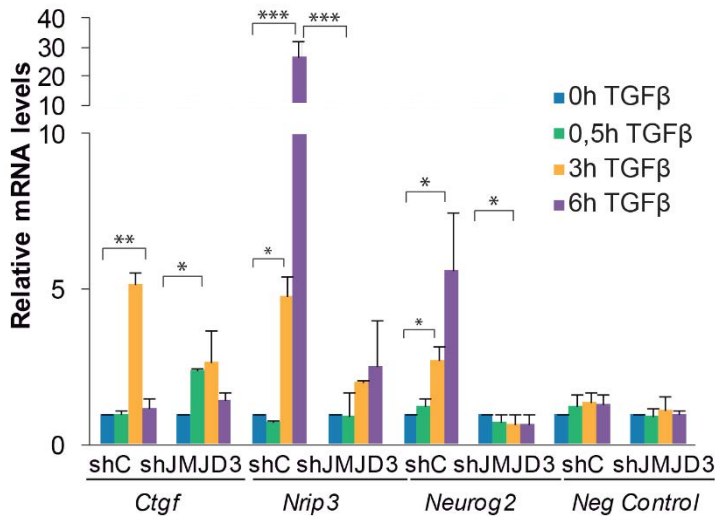
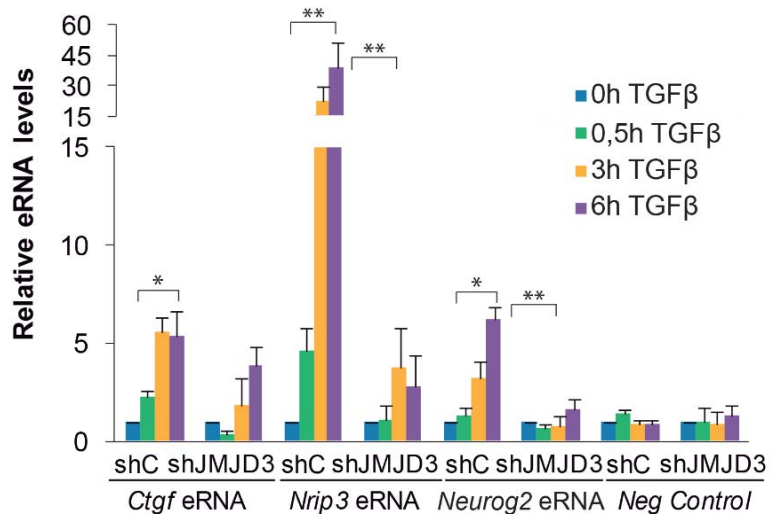


Figure R24. eRNA transcription in shC and shJMJD3 cells treated with TGF β . Results are normalized over *Rps23* housekeeping gene. Negative control corresponds to the *Fabp4* eRNA. Error bars indicate SD. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ (Student's t-test).

As in the shSMAD3 NSCs, the lack of activation of the eRNAs correlates with the lack of induction of the associated genes (Figure R25).

Figure R25. mRNA transcription in shC and shJMJD3 cells treated with TGF β . Results are normalized over *Rps23* housekeeping gene. Negative control corresponds to the *Fabp4* mRNA. Error bars indicate SD. * $p < 0.05$; ** $p < 0.01$; (Student's t-test).



3.2.2. Role of JMJD3 in the acquisition of active enhancer histone marks

After observing that JMJD3 affects eRNA transcription and mRNA expression, we tested whether this impairment was accompanied by changes on histone modifications. ChIP-qPCR experiments in Figure R26 show that the lack of JMJD3 affects the activation marks. TGF β treatment was unable to increase the H3K27ac or H3K4me2 levels in the *Ctgf*(-102) enhancer and the H3K27ac levels in the *Nrip3*(-3.5) enhancer in JMJD3 depleted cells.

These results suggest that JMJD3 affects the activation of enhancers. In the next subsection, the involvement of its catalytic activity will be studied.

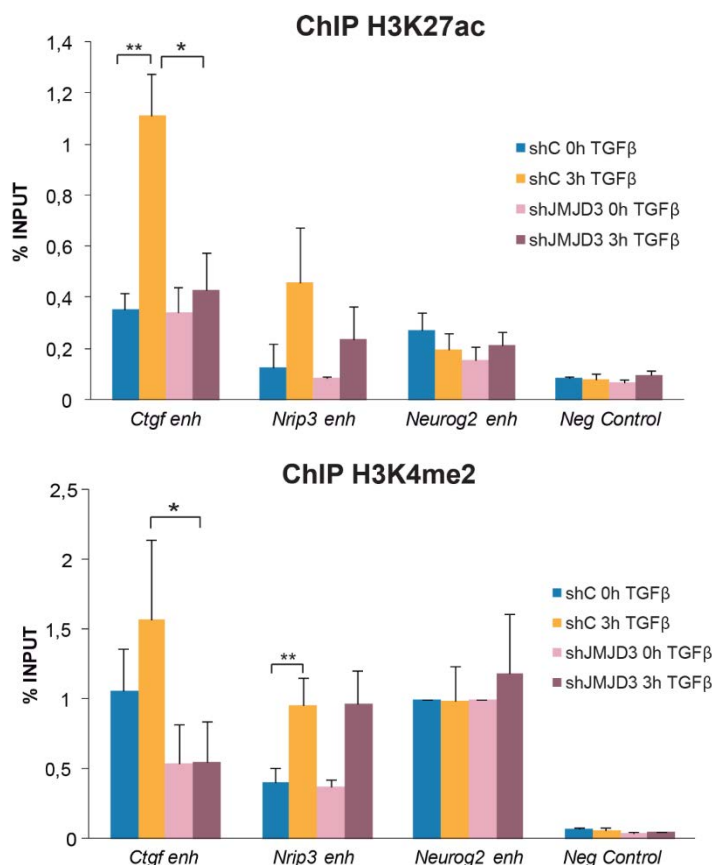


Figure R26. ChIP-qPCR results of enhancer activation histone marks at the indicated enhancers. Error bars indicate SD. * $p < 0.05$; ** $p < 0.01$.

3.2.3. Role of the H3K27me2/3 in enhancer activation

As it was introduced, JMJD3 is a HKDM with activity towards H3K27me2/3. To assess the catalytic role of JMJD3 at the enhancers, we first selected three H3K27me3-containing enhancers (enhancers in poised state) from genes that are activated upon TGF β -stimulation: *Neurog2*(-6), *Chic2*(-26), *Tle3*(-114) (Figure

Results

R27). We also added an active enhancer to the analysis, *Ctgf*(-102), this enhancer is devoid of H3K27me2/3 marks.

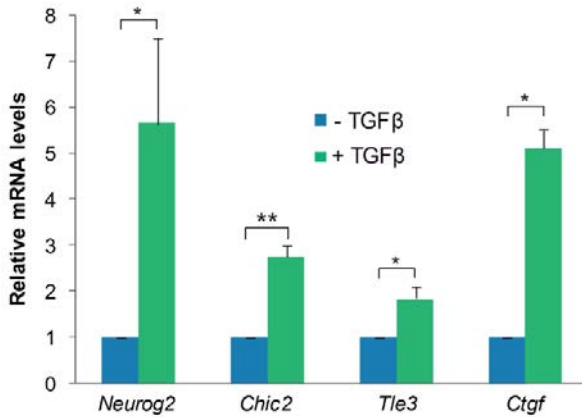


Figure R27. RT-qPCR experiment showing the activation of the poised enhancer associated genes and *Ctgf* upon TGFβ-signaling. Error bars indicate SD. * $p < 0.05$; ** $p < 0.01$ (Student's t-test).

Then, the levels of H3K27me3 in shC and shJMJD3 NSCs were determined before and upon TGFβ stimulation by ChIP-qPCR. Results in Figure R28 show that a significant change on the H3K27me3 levels between shC and shJMJD3 was only observed in the *Neurog2*(-6) enhancer upon TGFβ addition. In the case of *Chic2*(-26) and *Tle3*(-114), H3K27me3 levels decreased upon stimulation; however, these changes were independent on JMJD3. The control active enhancer *Ctgf*(-102)

did not suffer any methylation change in any of the conditions tested (Figure R28).

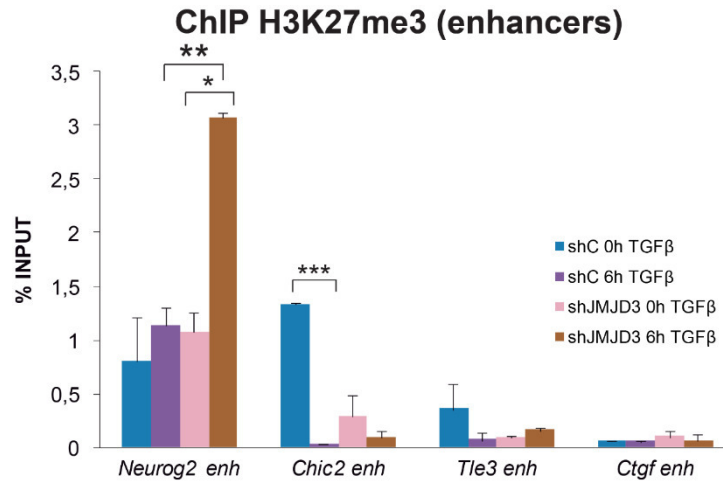


Figure R28. ChIP-qPCR showing the H3K27me3 levels at the poised enhancers and the *Ctgf* active enhancer upon TGFβ stimulation in shC versus shJMJD3 NSCs. Error bars indicate SD. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ (Student's t-test).

After observing that the demethylation of the H3K27me3 at enhancers by JMJD3 was negligible, we thought that JMJD3 could be demethylating its other substrate, the H3K27me2. This histone mark has been suggested to protect enhancers from spurious activation (Ferrari et al., 2014), suggesting that JMJD3 could be assuring activation fidelity at enhancers.

Results in Figure R29 demonstrate that the demethylation of H3K27me2 does not seem the main activity of JMJD3 at enhancers.

ChIP H3K27me2 (enhancers)

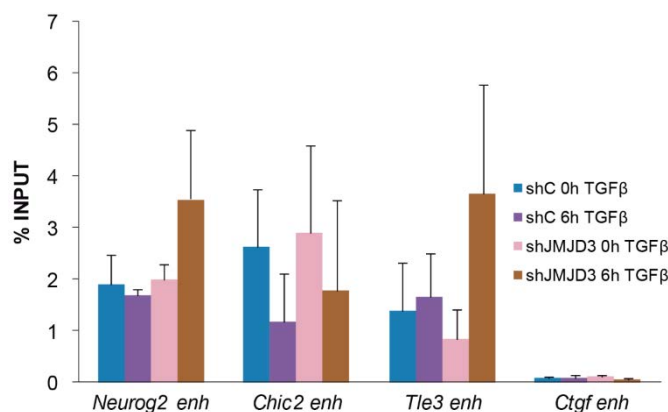


Figure R29. ChIP-qPCR showing the H3K27me2 levels at the poised enhancers and the *Ctgf* active enhancer upon TGFβ stimulation in shC versus shJMJD3 NSCs.

At TGFβ-responsive promoters, the function of JMJD3 consists in their demethylation (Estaras et al., 2012). However, we do not observe demethylation in all the poised enhancers. In order to ensure that the promoters associated to these enhancers are being demethylated upon TGFβ stimulation, we performed ChIP-qPCR experiments of H3K27me3 (Figure R30). Our results show that after stimulating with TGFβ, H3K27me3 decreases at promoters as expected.

ChIP H3K27me3 (promoters)

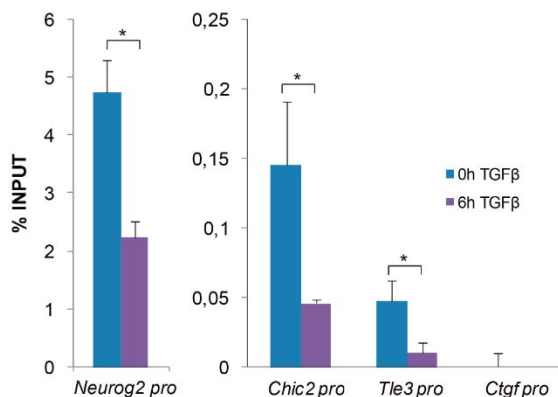


Figure R30. ChIP-qPCR showing the H3K27me3 levels at promoters corresponding to the poised enhancers and the *Ctgf* active enhancer upon TGFβ stimulation. Error bars indicate SD. *p < 0.05 (Student's t-test).

Additionally, we observe that this demethylation is accompanied by a decrease in the PRC2 complex subunit EZH2, and an increase in the active promoter histone mark H3K4me3, concluding that the role of JMJD3 at enhancers and promoters might be different (R31).

Results

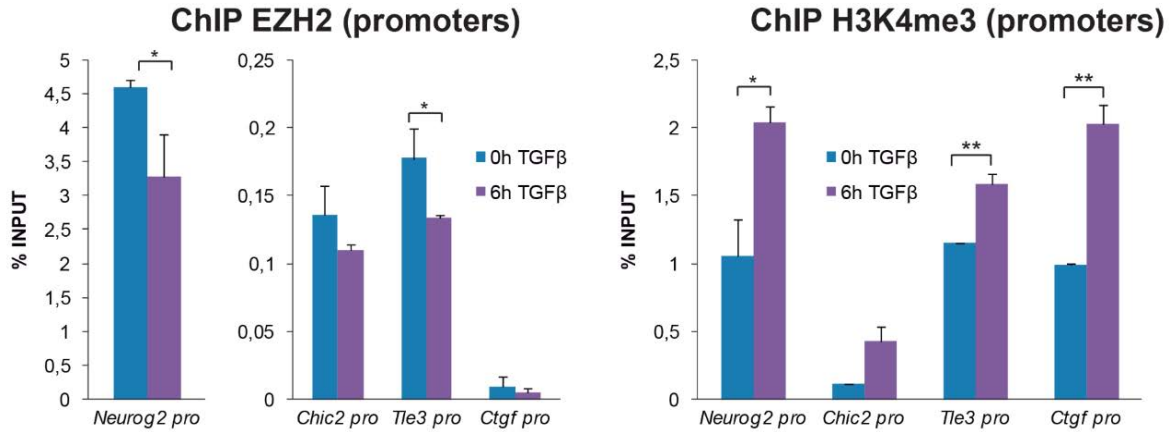


Figure R31. ChIP-qPCR showing the EZH2 levels (left) and the H3K4me3 levels (right) at promoters corresponding to the poised enhancers and the *Ctgf* enhancer upon TGFβ stimulation. Error bars indicate SD. * $p < 0.05$; ** $p < 0.01$ (Student's t-test).

4. Contribution of CHD8 to the TGFβ-responsive enhancers activation

After observing that at TGFβ-triggered enhancers, JMJD3 has catalytic dependent and independent roles, we hypothesized that JMJD3 could be recruiting other chromatin modifying proteins for the activation. Defects in eRNA transcription suggested us the possibility of a chromatin remodeler participating in this process. CHD8 belongs to a wide family of ATP-dependent remodelers that bind and open chromatin at enhancer regions (Ceballos-Chavez et al., 2015). Moreover, as JMJD3, CHD8 is highly expressed in embryonic cortex at stage E12.5 and is essential for early neurogenesis (Durak et al., 2016; Nishiyama et al., 2004; Platt et al., 2017). These facts led us to investigate whether CHD8 could be participating on the activation of the TGFβ-responsive enhancers.

To investigate this hypothesis, we first checked whether CHD8 could be physically contacting SMAD3 and JMJD3. For that purpose, we used size exclusion chromatography on whole NSCs extracts and we observed that CHD8 co-eluted with the phosphorylated form of SMAD3. Consistently, JMJD3 co-eluted in the same fractions (Figure R32). Additionally, by CoIP experiments with endogenous proteins in NSCs, we demonstrated that CHD8 and JMJD3 interact in NSCs (Figure R33).

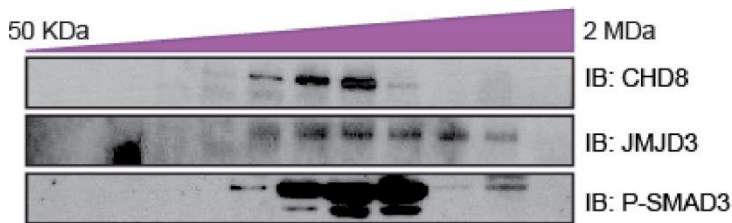


Figure R32. Size-exclusion chromatography immunoblots depicting the co-elution of CHD8, JMJD3 and phospho-SMAD3.

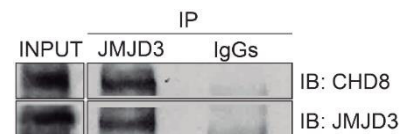


Figure R33. CoIP experiment showing the interaction between CHD8 and JMJD3 in NSCs.

As these results suggested that CHD8 could be forming a complex with JMJD3 and SMAD3, we tested whether CHD8 was bound at enhancers upon TGFβ-signaling. ChIP-qPCR results in Figure R34 demonstrate that CHD8 is already bound to these enhancers before TGFβ-stimulation (presumably, due to the basal activity of TGFβ pathway in NSCs). Nonetheless, TGFβ treatment significantly increases the levels of CHD8 at the analyzed enhancers (Figure R34).

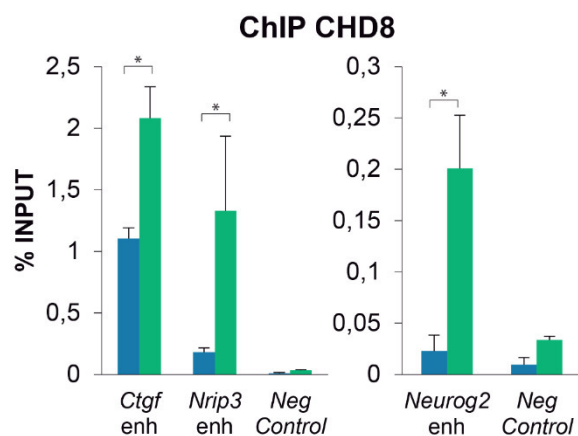


Figure R34. ChIP-qPCR experiment of the protein CHD8 at neural enhancers upon TGFβ signaling. Error bars represent SEM. *p < 0.05 (Student's t-test).

To further investigate if CHD8 was involved in eRNA transcription, we generated a KD cell line for CHD8 using specific shRNAs (Figure R35). After efficient depletion of the protein, we performed RT-qPCR to analyze the contribution of this chromatin remodeler to the eRNA transcription triggered by the TGFβ signaling cascade. The top graphic in Figure R36 displays a total inhibition of eRNA transcription in the shCHD8 NSCs, thus phenocopying the shSMAD3 and shJMJD3 effect. Consequently, the levels of the mRNAs associated to the examined enhancers did not increase upon TGFβ treatment (Figure R36, bottom).

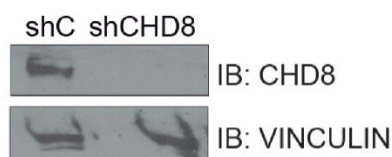


Figure R35. Immunoblot showing the levels of CHD8 in the shC cell line versus the shCHD8 cell line.

Results

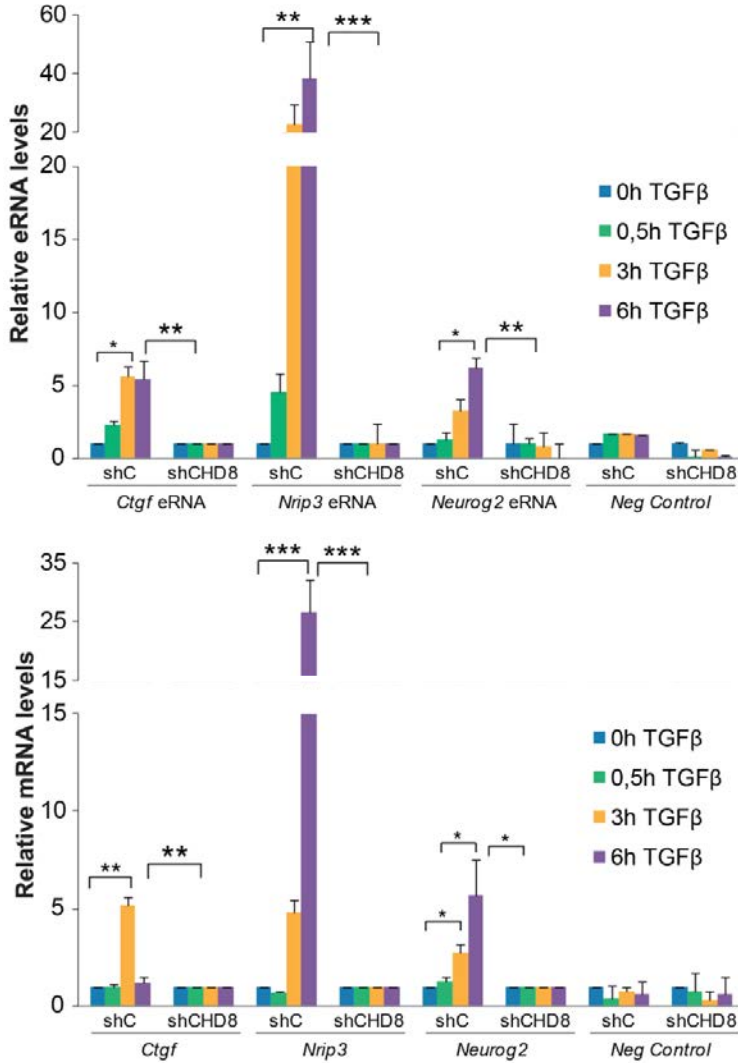


Figure R36. eRNA (top) and mRNA (bottom) transcription in shC and shCHD8 cells treated with TGFβ. Results are normalized over *Rps23* housekeeping gene. Negative control corresponds to the *Fabp4* eRNA and mRNA. Errors bars represent SD. *p < 0.05; **p < 0.01; ***p < 0.001 (Student's t-test)

Results in Figures R34 and R36 put forward to consideration the potential cooperation of CHD8 with the TGFβ pathway. Additionally, a comparison between the downregulated genes from a previously published RNA-seq performed in shCHD8 E12.5 embryo mice cortices (Durak et al., 2016) and our previously published microarray data (Estaras et al., 2012), returned that a 24% of the genes regulated by TGFβ in a JMJD3-dependent manner (JDTA genes) were downregulated in the shCHD8 RNA-seq (p-value 2.2e-16, equal proportions test against a random sample of genes) (Figure R37).

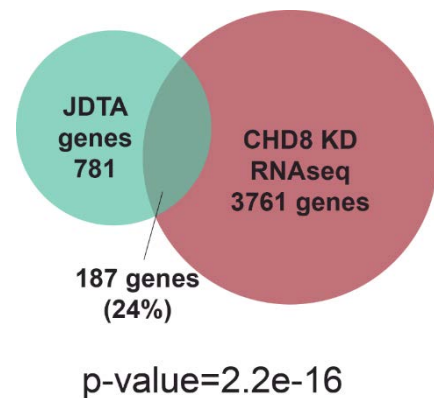


Figure R37. Venn diagram showing the number and percentage of JDTA genes that are downregulated in the shCHD8 RNA-seq.

The previous correlation suggests that CHD8 cooperates in the transcriptional regulation carried out by TGF β and JMJD3, and as CHD8 interacts with JMJD3, we thought that JMJD3 could be serving as a recruiting platform for CHD8. To test this possibility, we investigated the role of JMJD3 in the CHD8 binding to neural enhancers upon TGF β . We performed CHD8 ChIP-qPCR experiments on *Ctgf*(-102), *Nrip3*(-3.5) and *Neurog2*(-6) enhancers. The results in Figure R38 show a clear decrease on CHD8 recruitment to the analyzed enhancers upon TGF β stimulation in the shJMJD3 cells.

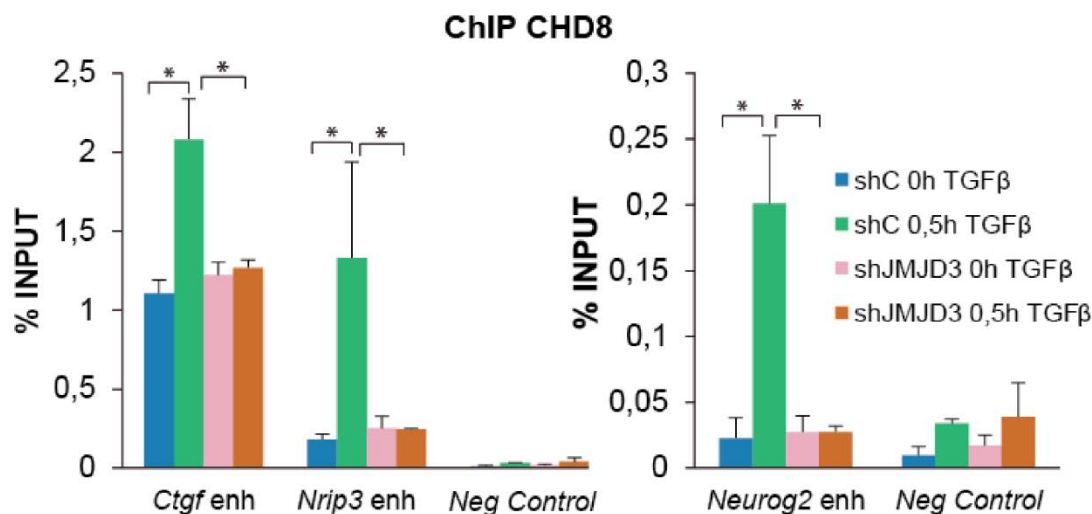


Figure R38. ChIP-qPCR experiment of CHD8 in the shC versus shJMJD3 cell line upon TGF β stimulation. Results are the mean of two biological independent experiments. Error bars represent SEM. * $p < 0.05$ (Student's t-test).

Altogether, these data demonstrate that CHD8 is a new SMAD3/JMJD3 partner at neural enhancers that is essential for full enhancer activation upon TGF β signaling.

5. TGF β re-organizes the tridimensional structure of the chromatin

In the subsection 1.2.1 of the Introduction, I detailed the different levels of 3D-chromatin folding. The most relevant level of chromatin organization for this thesis, are the loop domains (Rao et al., 2014). Contacts between enhancers and promoters have been systematically described in many models and we hypothesized that they could be playing a role in the TGF β regulation of enhancers.

To analyze our hypothesis, we performed 4C-seq experiments in collaboration with Dr. Rada-Iglesias and Dr. de la Cruz Molina in Cologne. In the next subsections I will include the rationales for the experiment and its results.

Results

5.1. Design of the 4C-seq experiment

In the methods section I detailed the protocol for the 4C-seq experiment (Figures M3, M4 and M5). This protocol allows us to analyze the contacts between a selected region (viewpoint) and the rest of the genome. Nonetheless, it has some designing criteria that might limit the enhancer-promoter contacts that could be studied due to limitations in resolution.

To design our 4C-seq experiment, we selected a TGF β and JMJD3 regulated enhancer that putatively corresponded to the promoter of a gene that responds to TGF β in a JMJD3-dependent manner, the gene *Chst8* (Figure R39).

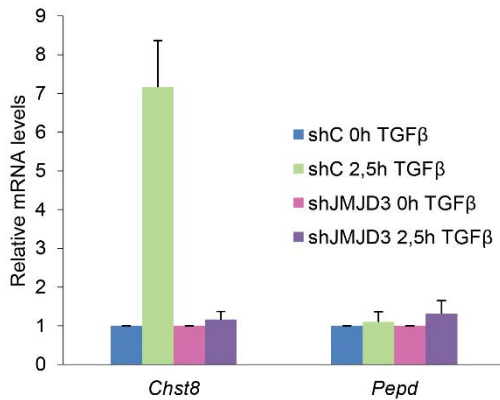


Figure R39. Transcription of the *Chst8* gene and the negative control *Pepd* upon TGF β signaling in shC and shJMJD3 cell lines.

Besides, the enhancer should be located at least 30 Kb far from its cognate promoter, in order to clearly observe whether in the tested conditions, the enhancer was contacting it. Figure R40 displays an IGV capture of the selected viewpoint and its distance to the putative regulated promoter.

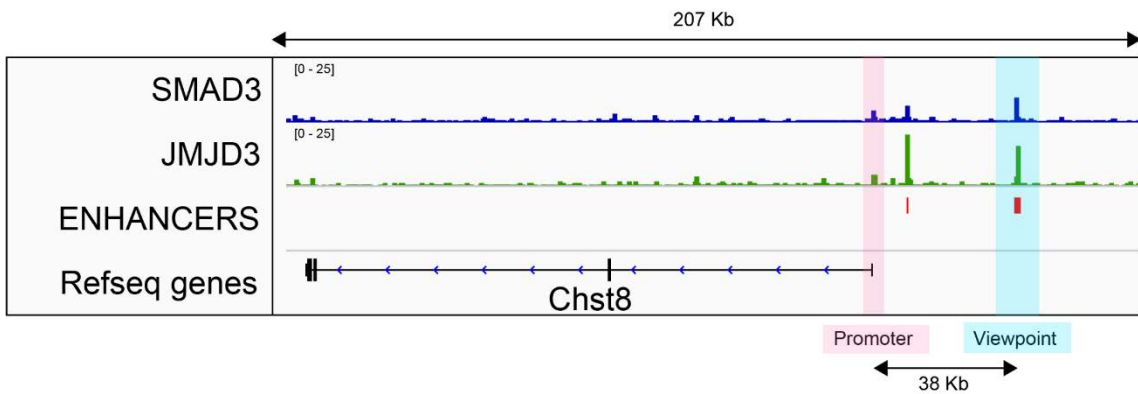


Figure R40. IGV capture showing the viewpoint selected for the 4C-seq and its cognate promoter. The location of enhancers and SMAD3 and JMJD3 peaks in CHIP-seq is displayed in the tracks.

5.2. TGF β induces enhancer-enhancer and enhancer-promoter contacts

After selecting the previously detailed region for the 4C-seq experiment, we performed the assay in shC cells treated with TGF β during three hours. This time point was selected as a time in which we have demonstrated that transcription elongation is actively occurring, thus increasing the possibilities of observing the contacts.

Capture in Figure R41 shows the obtained profiles for the untreated and TGF β treated shC NSCs. As expected, the contact between the enhancer and the promoter increased with the TGF β treatment (see region under light blue). However, we also observed contacts that decrease or disappear, concluding that TGF β is triggering a re-organization of the chromatin at this region.

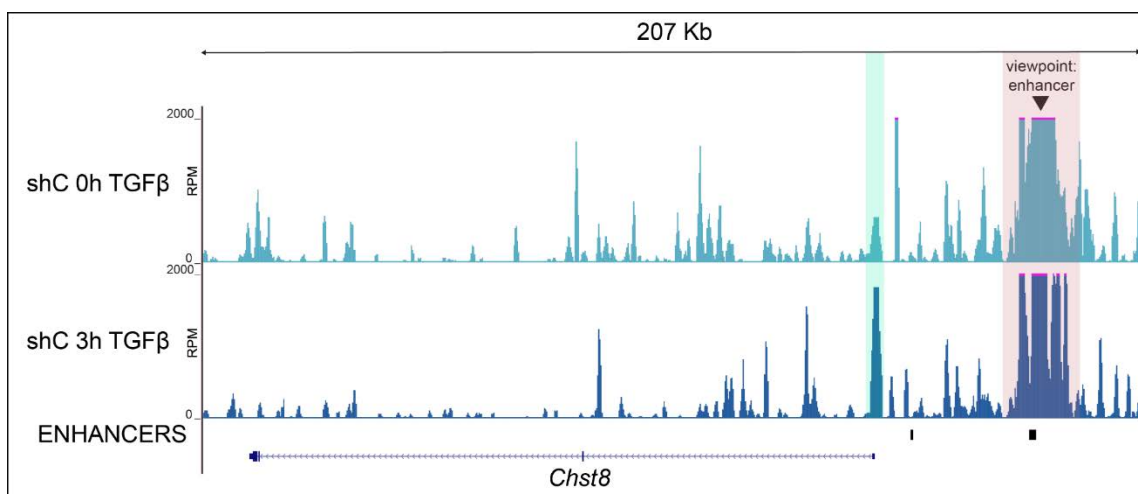


Figure R41. 4C-seq results showing the chromatin re-organization triggered by TGF β in shC NSCs.

Interestingly, zooming out the view in the browser, we could observe that many contacts were occurring between the viewpoint and inter- and intragenic enhancers located in a gene that lies at the telomeric part of the viewpoint. This gene, *Pepd*, is not regulated upon TGF β (Figure R39) and its promoter does not contain SMAD3, JMJD3 or contact the viewpoint (Figures R40 and R42). As we show in Figure R42, the number and intensity of contacts between the viewpoint and the *Pepd* gene is higher than between the viewpoint and the *Chst8* gene. Strikingly, these enhancers contain SMAD3 and JMJD3 (Figure R43), pointing to a convergence of a TGF β -regulated enhancers that could be potentially acting together to activate a TGF β -responsive gene promoter. As I mentioned in the introduction, superenhancers are clusters of enhancers that cooperate to induce cell identity genes (Hnisz et al., 2013).

Results

In this case, we observe the formation of a *functional* superenhancer that induces TGF β responsive genes.

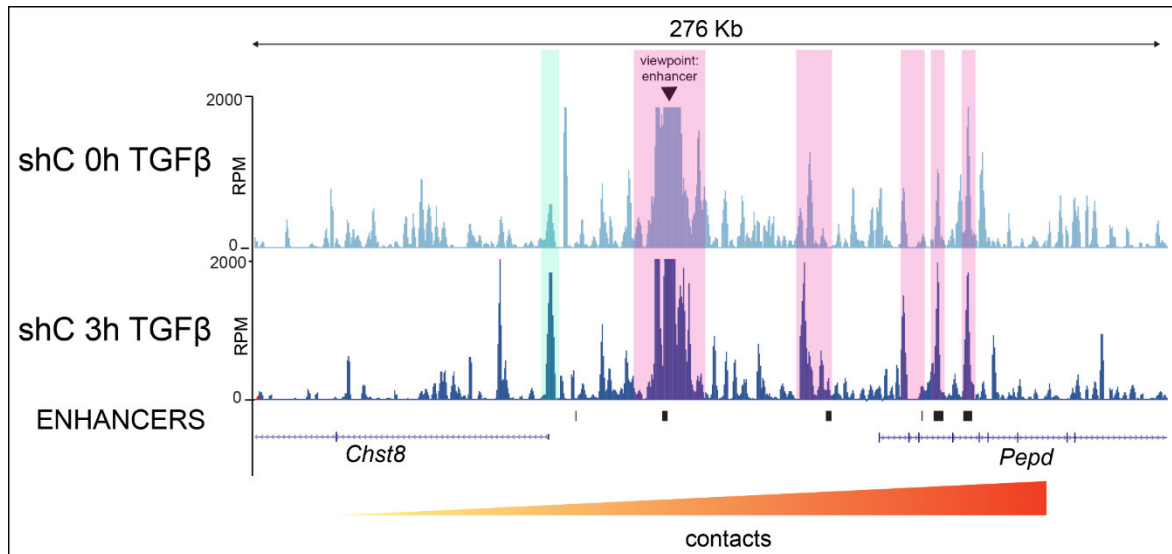


Figure R42. 4C-seq results showing the contacts formed upon TGF β -stimulation. Pink boxes indicate enhancer-enhancer contacts; light blue indicates enhancer-promoter contacts.

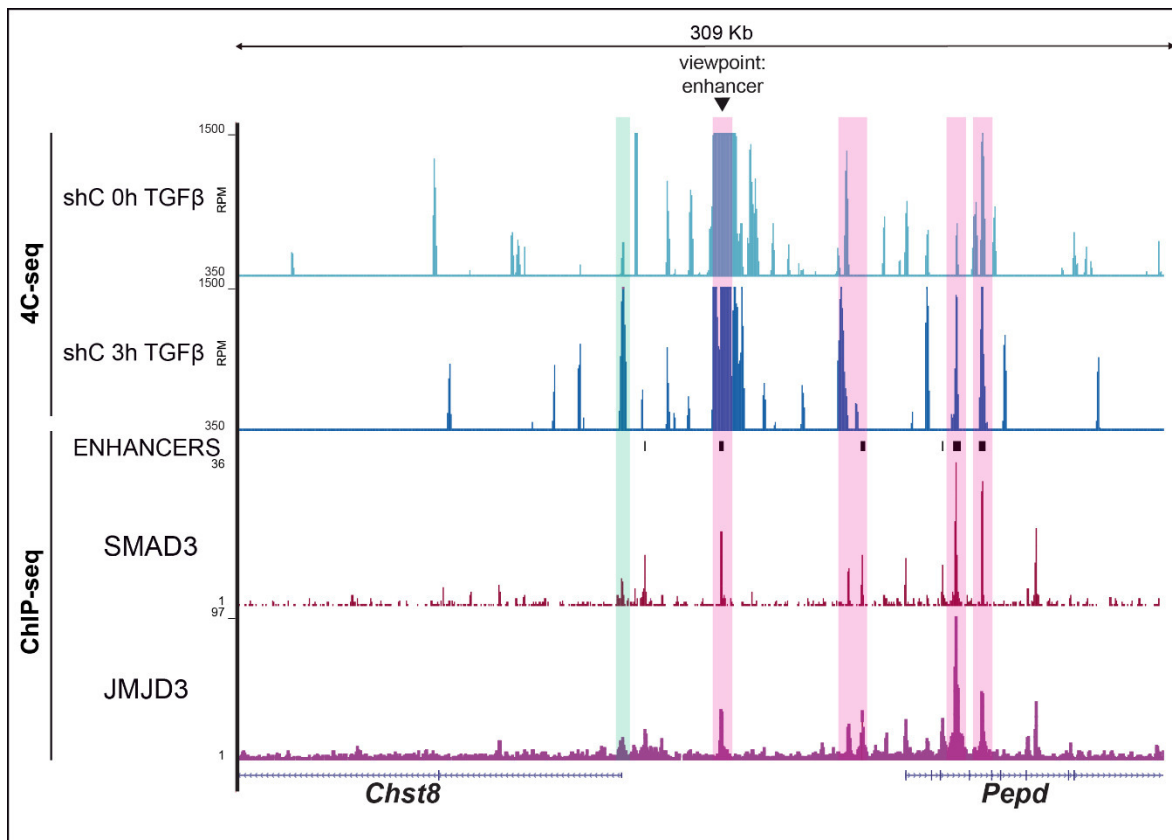


Figure R43. 4C-seq results showing the contacts formed upon TGF β -stimulation and the SMAD3 and JMJD3 ChIP-seq enrichments on these regions. Pink boxes indicate enhancer-enhancer contacts; light blue indicates enhancer-promoter contacts.

5.3. JMJD3 is necessary for TGF β -triggered contact formation

In the previous subsections, I have shown data that point to potential catalytic-independent roles of JMJD3 at enhancers. After observing that the enhancer-enhancer and enhancer-promoter contacts promoted by TGF β contained JMJD3, we wondered whether JMJD3 could be contributing to the establishment of the 3D-chromatin structure.

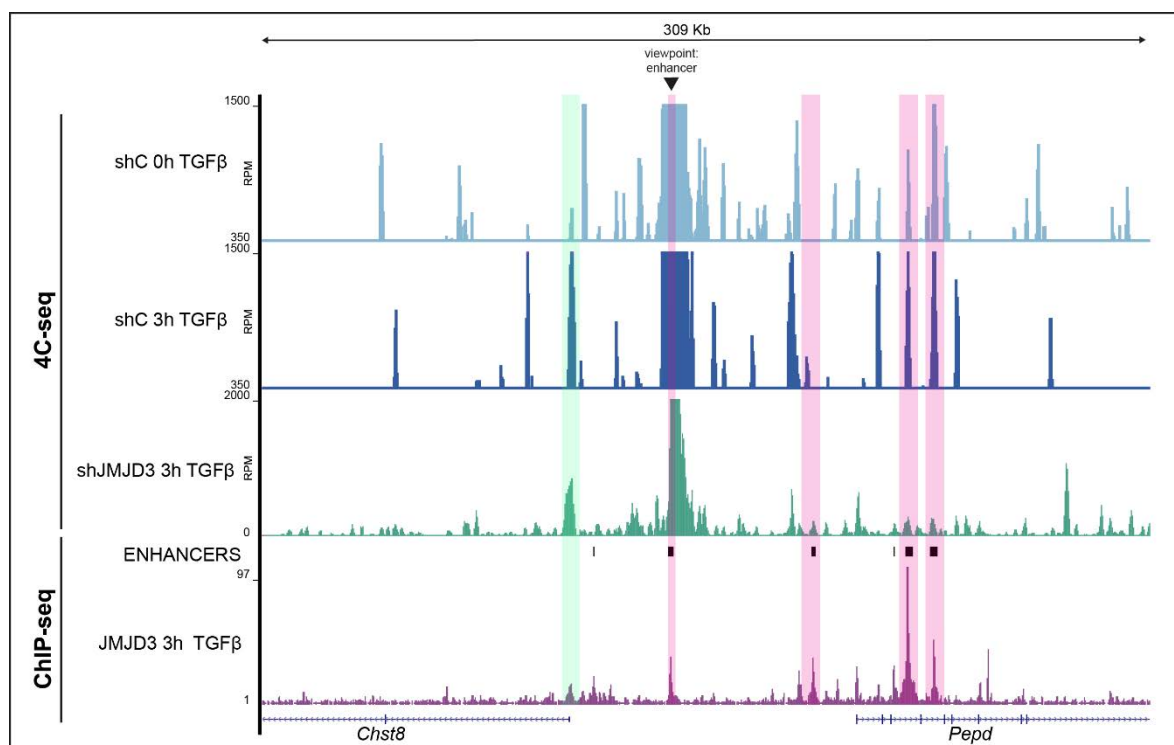


Figure R44. 4C-seq results showing the contacts formed upon TGF β -stimulation in shC and shJMJD3 cell lines. JMJD3 ChIP-seq enrichments on this region are also shown. Light pink boxes indicate enhancer-enhancer contacts; light blue indicates enhancer-promoter contacts.

For this purpose, we used our previously established shJMJD3 NSCs (Figure R22) treated during 3 hours with TGF β to perform 4C-seq experiments using the *Chst8* enhancer viewpoint. Capture in Figure R44 (and its zoom-in in R45) depicts the massive effect that the KD of JMJD3 has in the contact establishment. It is observed, that the depletion of JMJD3 leads to a loss of the contacts, thus indicating that JMJD3 is required for faithful 3D-chromatin structure acquisition upon TGF β signaling.

Altogether, these data qualitatively demonstrate that TGF β re-organizes the chromatin in a JMJD3-dependent manner.

Results

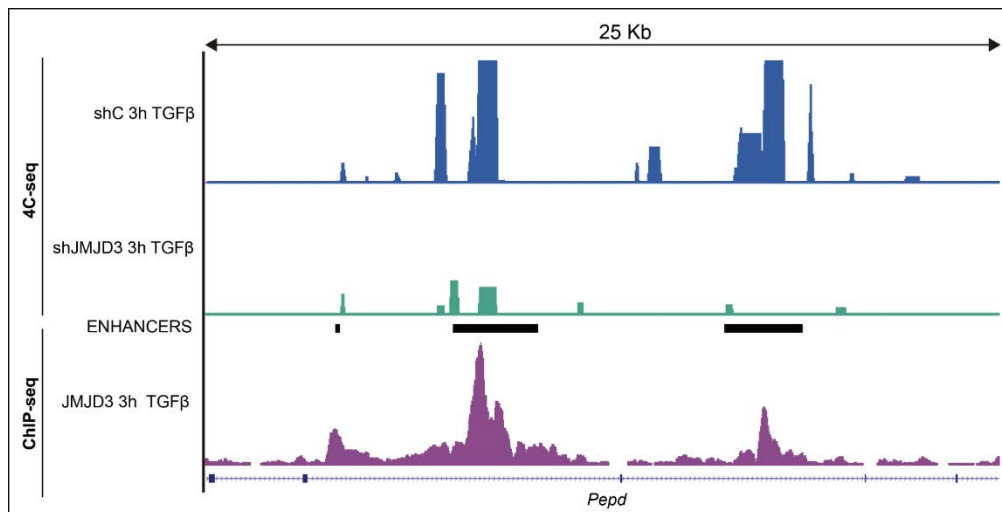


Figure R45. Capture showing a zooming-in at a region where TGF β -induced contacts are lost in shJMJD3 NSCs.

In order to show in a quantitative manner the differences between the TGF β treatment in shC and shJMJD3 cell lines, we calculated the average of the signals obtained in the restriction-based bioinformatic analysis. We averaged the values of peaks located 1 Mb around the viewpoint in every condition. The box plot in Figure R46 displays a significant increase in the average signal when shC cells are treated with TGF β (p-value $4e-4$). Between the shC and shJMJD3 NSCs treated with TGF β , it is observed a highly significant decrease in the average signal (p-value $7.2e-14$), corresponding to the loss of contacts (Figure R46).

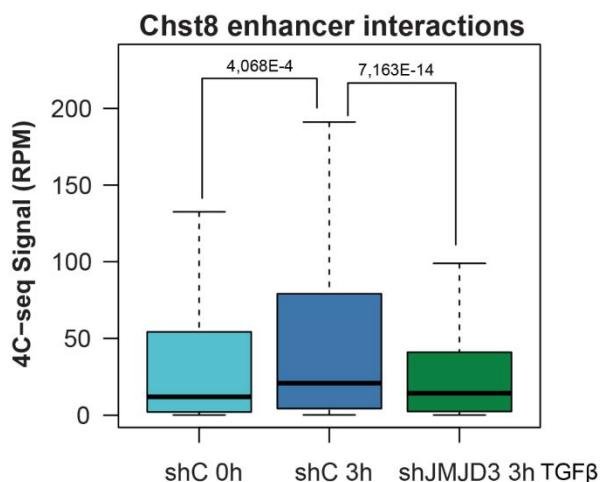


Figure R46. Box plot displaying the differences between the signals in RPM from the indicating samples. p-values are the result of a Wilcoxon-Mann-Whitney test.

5.3.1. Cohesin contribution to the 3D-chromatin structure

These results suggest that JMJD3 has a quintessential role in the establishment of contacts after TGF β stimulation. To rule out the possibility of an indirect effect due

to defects in the LMC proteins expression, we analyzed gene expression data of our previously published microarray (Estaras et al., 2012). The graphic in Figure R47 demonstrates that neither TGFβ nor JMJD3 regulate the expression of the proteins involved in this process.

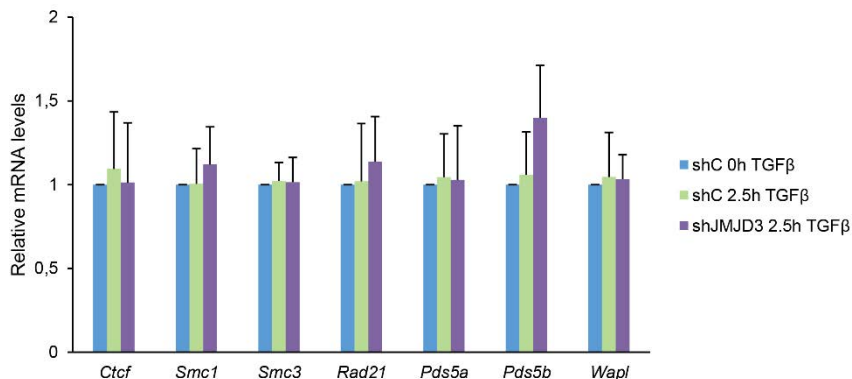


Figure R47. Microarray expression data corresponding to the protein members of the loop maintenance complex in shC and shJMJD3 cell lines treated with TGFβ.

Next, we looked at previously published ChIP-seq data of the member of the Cohesin complex SMC1, in NSCs. This protein is necessary for the establishment of dynamic contacts during gene transcription (Sanborn et al., 2015) and we hypothesized that it could be participating in the TGFβ re-organization of the chromatin cooperating with JMJD3.

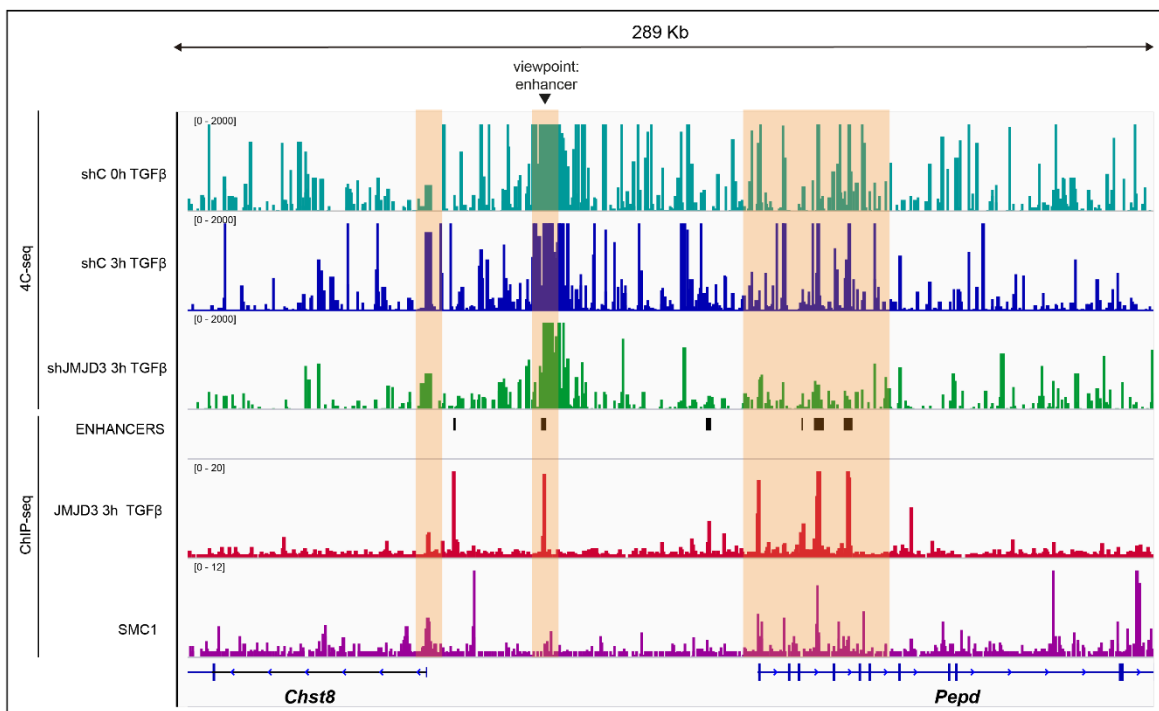


Figure R48. IGV capture showing 4C-seq results in the shC and shJMJD3 cell lines upon TGFβ treatment. ChIP-seq signals of JMJD3 and SMC1 are also shown. Regions where JMJD3 and SMC1 signals colocalize are highlighted in light orange.

Results

To test this possibility, we observed the ChIP-seq signal of SMC1 at the region surrounding the *Chst8* enhancer viewpoint. Interestingly, we noticed that some of the regions occupied by JMJD3, were also occupied by SMC1, especially those belonging to the *functional* superenhancer region. In the capture in Figure R48 the coordinates where JMJD3 and SMC1 colocalize are highlighted in light orange.

This observation prompted us to check whether JMJD3 could be co-occupying genomic regions with SMC1 in a genome-wide manner. The Venn diagram in Figure R49 shows that these proteins do not colocalize widely across the genome. Only a 19% of the peaks of JMJD3 coincide with the peaks of SMC1. On the other hand, displaying of the SMC1 ChIP-seq signal in the 19% of genomic regions bound by SMC1 and JMJD3 (heatmap, Figure R50), shows high colocalization with JMJD3. It is worth noting, that the SMC1 ChIP-seq has been performed without TGF β treatment, and this could be contributing to a smaller overlap.

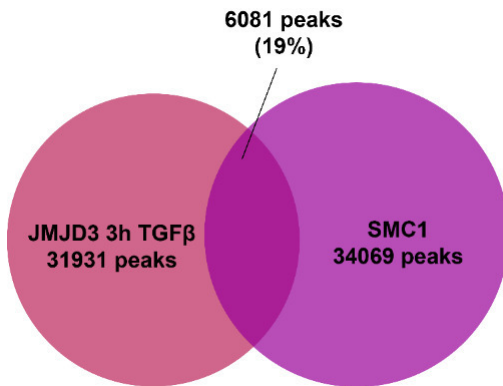
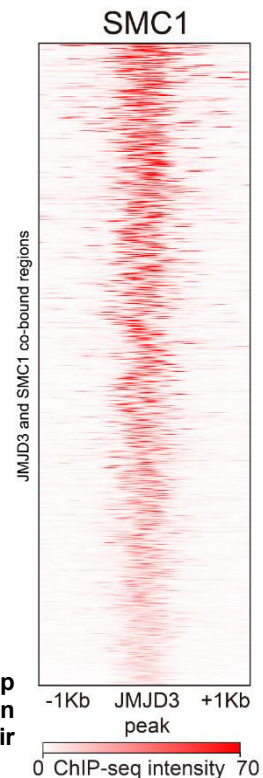


Figure R49. Venn diagram showing the colocalization of peaks corresponding to JMJD3 and SMC1 proteins

Figure R50. Heatmap showing the colocalization of SMC1 and JMJD3 at their co-bound regions



At the moment, we are performing molecular experiments to deepen into the potential cooperation between JMJD3 and SMC1. Preliminary ChIP-qPCR results do not show loss of SMC1 in the shJMJD3 NSCs treated with TGF β , thus the molecular mechanism by which JMJD3 is contributing to the contacts establishment is still unknown, and we hypothesize that is independent of LMC proteins.

6. *Neurog2(-6) enhancer and Neurog2 promoter cooperate to fine-tune transcription*

In the previous sections, I have provided data regarding the local and 3D events that contribute to the TGF β enhancer activation. In this section of my doctoral thesis, I would like to analyze the importance of enhancers and promoters for a faithful TGF β response.

The work from our lab has demonstrated how SMAD3 and JMJD3 are located at promoters and enhancers and participate in their regulation *in vivo* and *in vitro* (Estaras et al., 2012; Fueyo et al., 2018). The fact that both proteins are located at both regulatory regions makes difficult the attribution of the observed effect to enhancers or promoters. Here, I would like to show CRISPR-Cas9 results on the contribution of the different regulatory regions to the final output.

6.1. *Experimental strategy*

To address our question, we have generated cell lines with deletions in the *Neurog2(-6)* enhancer or promoter using CRISPR-Cas9 technology. Even though ideally the enhancer and the promoter deleted should have been the ones selected for the 4C-seq experiment, we decided to change our targets due to the high number of enhancers (those in the *functional* superenhancer) that potentially could be regulating *Chst8* promoter.

For our analysis we selected the *Neurog2* gene and enhancer. These regulatory regions have been molecularly characterized by us (Estaras et al., 2012; Fueyo et al., 2018) and additionally, the *Neurog2(-6)* enhancer has been functionally tested *in vivo* (Visel et al., 2007), thus ensuring that *Neurog2(-6)* is regulating *Neurog2* promoter. *Neurog2* promoter and enhancer are covered by H3K27me3 (Figure R51) in NSCs, this implies that these regions are inactive and their deletion should not have any effect in NSCs. Moreover, their activation with TGF β , permits us the manipulation of the system in an off-on manner, eliminating potential background. Another advantage of our selection is that NEUROG2 belongs to the bHLH TFs family and promotes neurogenesis. In NSCs, its expression leads to neuronal differentiation and this will allow us to analyze the phenotypic consequences of the enhancer deletion.

In Figure R51 there is a capture showing the regions deleted and their chromatin features. As it is observed, our deletions have eliminated SMAD3 and JMJD3 binding sites from the promoter and the enhancer.

Results

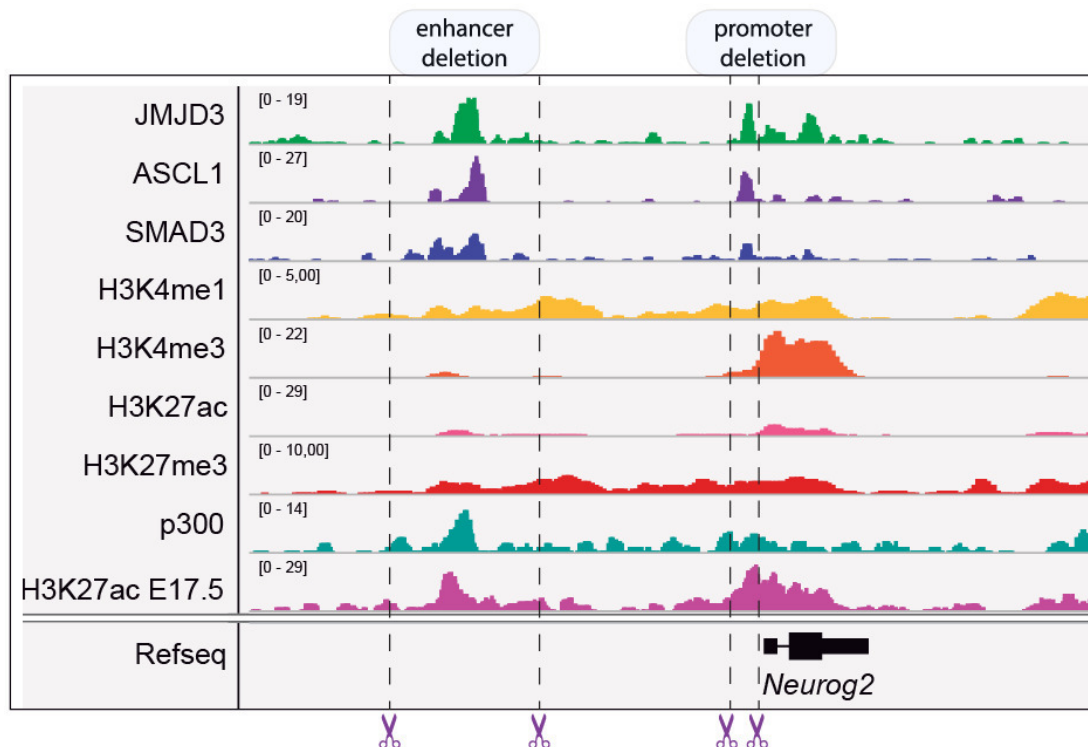


Figure R51. IGV capture showing the CRISPR-Cas9 deleted regions and their chromatin features.

In the case of the *Neurog2*(-6) enhancer we deleted 7.2 Kb that spanned all the enhancer chromatin features (Δ *Neurog2* enh cell line), in the case of the promoter, we deleted 1.1 Kb located upstream the TSS (Δ *Neurog2* pro cell line). To test that our CRISPR-Cas9 gRNAs had properly cut, we performed conventional PCRs with selected pairs of primers (Table M3). The results of these PCRs together with a scheme of the deleted regions can be found at Figure R52.

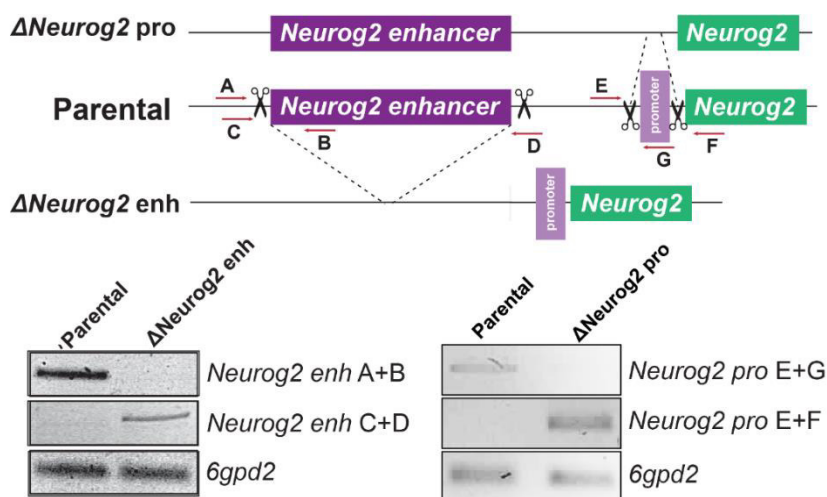


Figure R52. Scheme of the CRISPR-Cas9 strategy for the generation of Δ *Neurog2* enh or pro NSCs (top). In the bottom panel conventional PCR results for the mutant cell lines using the primers indicated with red arrows.

As we expected, both mutant cell lines grow normally and do not have any particular visible characteristic different to the Parental cell line.

6.2. Hierarchy between enhancers and promoters in the TGF β -signaling

Once the $\Delta Neurog2$ enh and $\Delta Neurog2$ pro cell lines were generated, our main goal was to dissect the contribution of every regulatory region to the final output.

6.2.1. Contribution of the *Neurog2* promoter to transcription

Promoters are regulatory regions whose main role is the control of the correct positioning of the RNAPII before starting to transcribe (Lenhard et al., 2012). Along this thesis, I have explained the importance of enhancer-promoter communication in the 3D structure of the chromatin. I will present here some data regarding the role of the promoter in the enhancer activation.

Using the Parental and the $\Delta Neurog2$ pro cell lines, we first tested whether the transcription of the eRNA from the *Neurog2*(-6) enhancer was affected by the lack of the promoter. The graphic in Figure R53 shows that in the absence of the promoter, the treatment with TGF β leads to an accumulation of the eRNA transcription. As expected, the *Neurog2* mRNA levels totally disappear, but not in the positive control *Nrip3*, that in the $\Delta Neurog2$ pro is seamlessly transcribed. This result shows that the *Neurog2* promoter is not required to activate the *Neurog2*(-6) enhancer in terms of eRNA transcription; however, this eRNA accumulation points to a misregulation of the enhancer, probably by accumulating RNAPII.

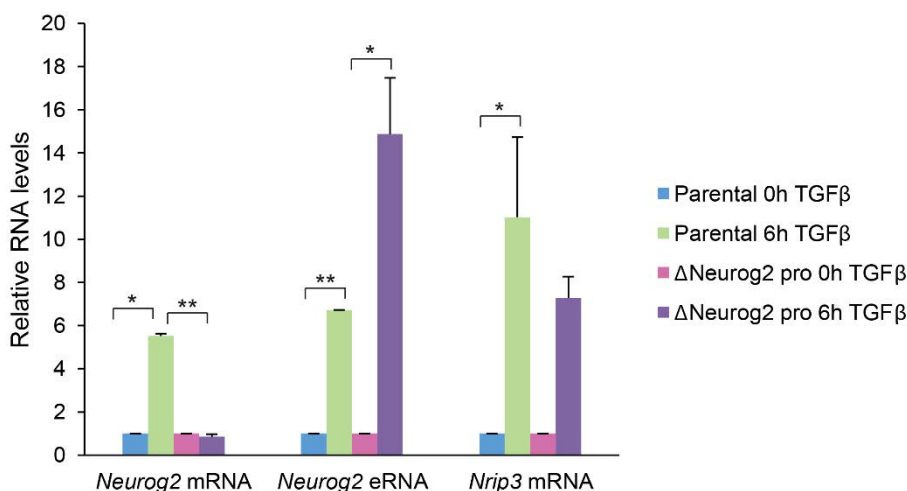


Figure R53. RT-qPCR showing the transcription levels of the *Neurog2* mRNA and eRNA. *Nrip3* is a positive control. Error bars represent SD. * $p < 0.05$; ** $p < 0.01$ (Student's t-test).

To answer whether the accumulation of the eRNAs could be related to a different recruitment of JMJD3 to the *Neurog2*(-6) enhancer in the absence of the *Neurog2*

Results

promoter, we performed ChIP-qPCR experiments in the Parental cell line versus the $\Delta Neurog2$ pro cell line. In Figure R54, it is observed that the deletion of the *Neurog2* promoter leads to a loss of JMJD3 recruitment at the *Neurog2(-6)* enhancer. In this experiment, we have used as a negative control a region where JMJD3 is not bound (*G6pd2*) and as a positive control a region where JMJD3 is recruited after TGF β independently of the deleted promoter (*Stx3*).

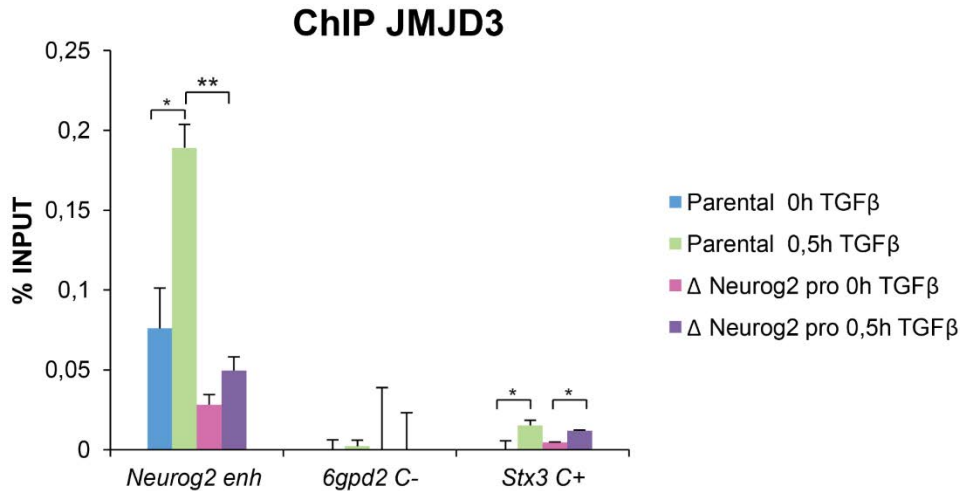


Figure R54. ChIP-qPCR experiment of JMJD3 in the Parental and $\Delta Neurog2$ pro cell lines. Error bars represent SD. *p-value < 0.05; **p-value < 0.01 (Student's t-test).

These two experiments (Figure R53 and R54) show that the promoter is required for an accurate enhancer activation. The lack of JMJD3 recruitment and the accumulation of eRNAs are not in agreement with our previous results in which we observed that shJMJD3 NSCs did not transcribe eRNAs as efficiently as the shC NSCs after TGF β stimulation, this result will be discussed in section 6.1.

6.3. Contribution of the *Neurog2(-6)* enhancer to transcription

Enhancers have historically been seen as modulators of the amount of mRNAs that are generated from coding genes (Banerji et al., 1981). Starting from this idea, we wanted to test whether the enhancer is required for proper TGF β -triggered promoter activation.

For this, we followed the same approach as in the previous subsection. First, we tested whether the lack of enhancer affects the transcription of mRNAs from the promoter. To answer this, we used the Parental and $\Delta Neurog2$ enh cell lines and we performed RT-qPCR to test the transcription of the *Neurog2* mRNA. Figure R55 presents that the lack of enhancer completely abolishes mRNA transcription. As expected, *Neurog2* eRNA was not detected and the control TGF β -responsive gene, *Nrip3*, was upregulated in both cell lines.

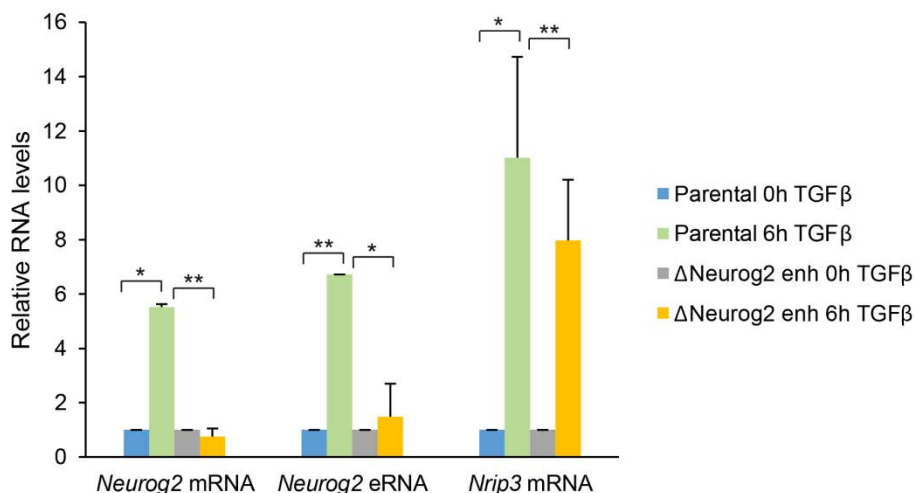


Figure R55. RT-qPCR showing the transcription levels of the *Neurog2* mRNA and eRNA. *Nrip3* is a positive control. Error bars represent SD. * $p < 0.05$; ** $p < 0.01$ (Student's t-test).

Next, we analyzed the recruitment of the cofactor JMJD3 to the *Neurog2* promoter in the Parental and the Δ *Neurog2* enh cell lines. In the Figure R56 are displayed the results showing that the enhancer is necessary for a recruitment of JMJD3 at the promoter after TGF β signaling, again, we have used as a negative control a region where JMJD3 is not bound (*G6pd2*) and as a positive control a region where JMJD3 is recruited after TGF β independently of the deleted enhancer (*Stx3*).

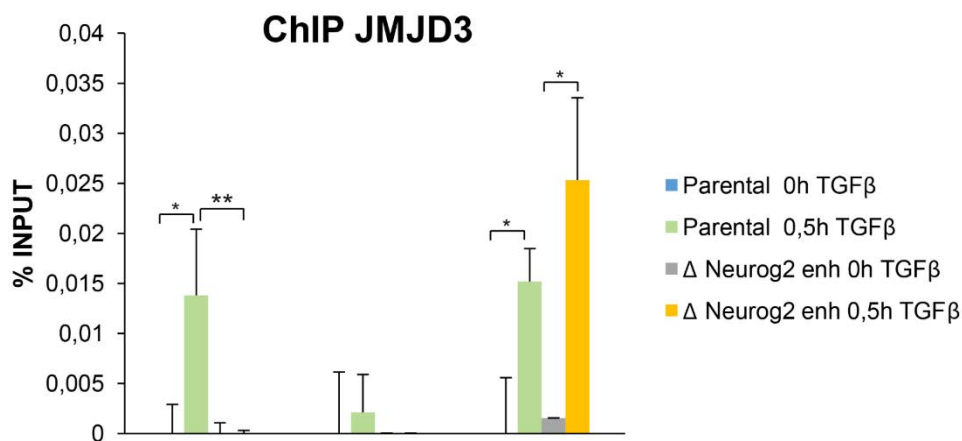


Figure R56. ChIP-qPCR experiment of JMJD3 in the Parental and Δ *Neurog2* enh cell lines. Error bars represent SD. * p -value < 0.05 ; ** p -value < 0.01 (Student's t-test).

These experiments (Figures R55 and R56) conclude that the *Neurog2(-6)* enhancer is necessary for the *Neurog2* gene activation, suggesting that enhancers are not just fine-tuners of the mRNA transcription, but essential modules for a faithful promoter activation.

Results

This doctoral thesis and previously work from the laboratory describes that JMJD3 demethylates gene promoters upon TGF β signaling (Estaras et al., 2012). Knowing that the *Neurog2* enhancer is required for JMJD3 recruitment, we hypothesized that in the Δ *Neurog2* enh NSCs, the promoter of *Neurog2* would not be properly demethylated. To test this, we performed ChIP-qPCR experiments of the histone mark H3K27me3 in the Parental and the Δ *Neurog2* enhancer cell lines (Figure R57). Accordingly with the lack of mRNA transcription and the lack of JMJD3 recruitment, in the Δ *Neurog2* enh NSCs the promoter of *Neurog2* was not efficiently demethylated, the *Ctgf*(-102) enhancer is used as a negative control for the H3K27me3.

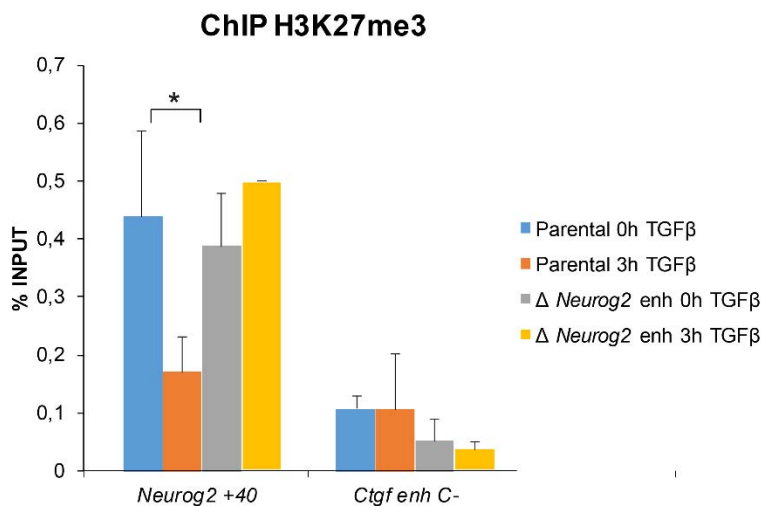


Figure R57. ChIP-qPCR of the histone mark H3K27me3 in Parental and Δ *Neurog2* enh cell lines at the TSS of the *Neurog2* gene. Error bars represent SD. * $p < 0.05$

Altogether, these experiments show that the *Neurog2*(-6) enhancer is necessary for proper gene activation due to its contribution to the recruitment of JMJD3 at the promoter.

6.3.1.1. *Neurog2*(-6) enhancer contribution to neurogenesis

As I have previously introduced, *Neurog2* is both necessary and sufficient to promote neuronal differentiation (Wilkinson et al., 2013). After establishing that the *Neurog2*(-6) enhancer is required for *Neurog2* expression, we performed neuronal differentiation experiments in order to see the phenotypical consequences of the lack of a TGF β -regulated enhancer. For these experiments, I collaborated with Simona Iacobucci, a member of the lab. She established the differentiation protocols and performed the immunofluorescence experiments.

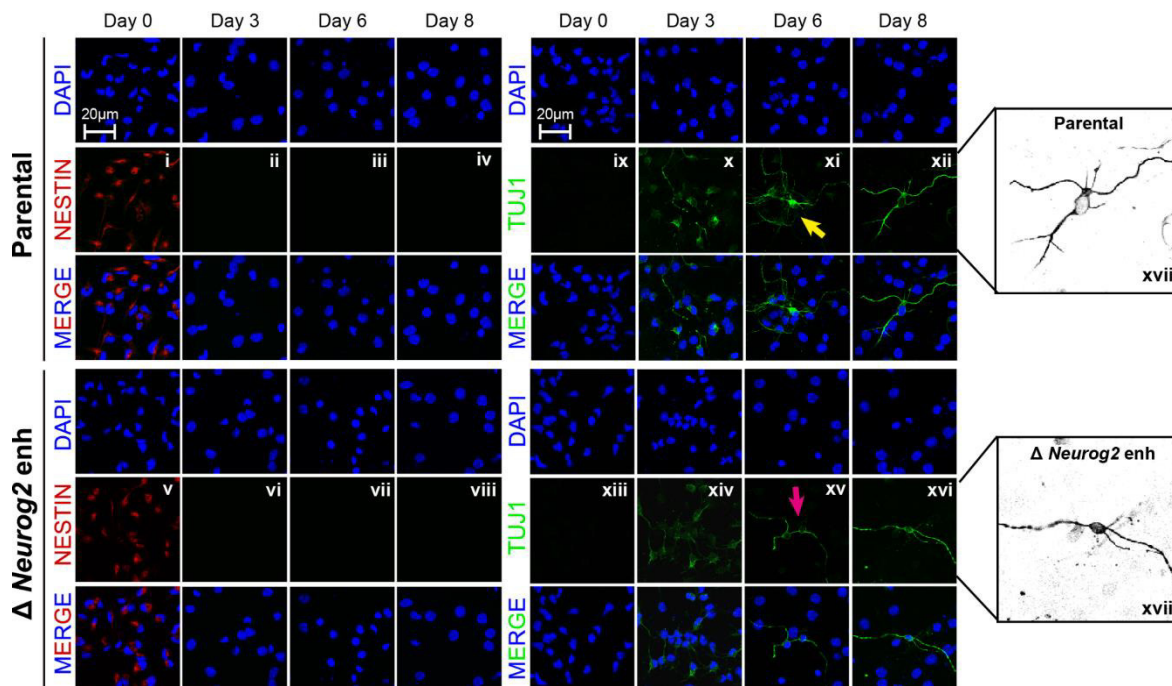


Figure R58. Parental and Δ *Neurog2* enh cell lines immunostainings of NESTIN and TUJ1.

For this procedure, Parental and Δ *Neurog2* enh cell lines were cultured under differentiating conditions (see section 2.1.1.1.3), and after 3, 6 or 8 days cells were fixed and immunostained for specific markers. In Figure R58, the indirect fluorescence representative experiment shows immunostaining for NESTIN (a progenitor marker present in dividing cells and downregulated upon differentiation) and TUJ1 (a neuronal marker). As the panel displays, after three days in differentiating medium, both cell lines stopped proliferating (Figure R58ii and vi) and the number of cells that expressed neuronal markers were similar in Parental and Δ *Neurog2* enh cells (Figure R58 x and xiv). Nevertheless, Δ *Neurog2* enh cells showed lower TUJ1 intensity than Parental NSCs (Figure R58, e.g. x versus xiv). Strikingly, Δ *Neurog2* enh NSCs failed to properly differentiate into neurons. After 6 days in differentiating medium, the number of neurites per cell was clearly reduced when compared with neurons from the Parental cell line, going from 4–5 in control to 2–3 in Δ *Neurog2* enh TUJ1+ cells (Figure R59). Moreover, in Δ *Neurog2* enh pseudoneurons, a high number of cells presented only one, two (uni/bipolar neurons) or none neurites (Figure R59).

These results demonstrate that the *Neurog2*(–6) enhancer is sufficient to modulate the acquisition of phenotypical traits during neuronal commitment, probably through its activity as a provider of transcriptional molecular machinery.

Results

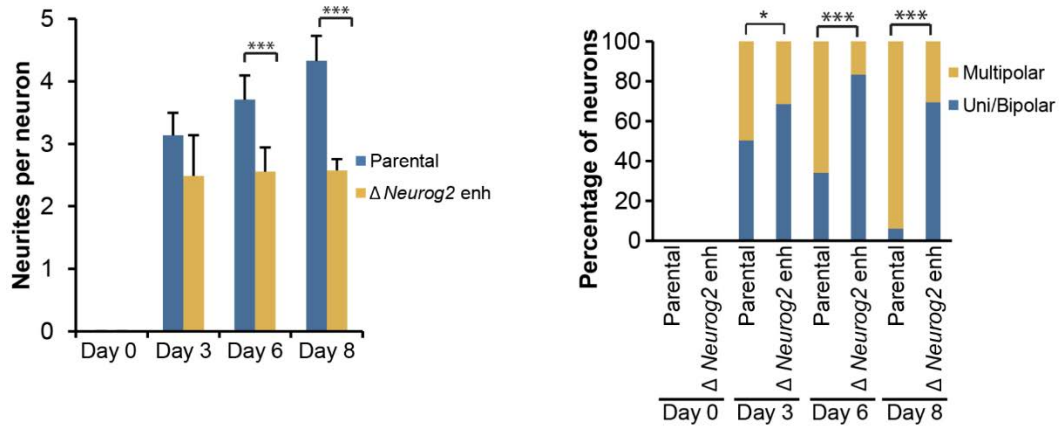


Figure R59. Counting of the number of neurites per neuron (left panel) and percentage of uni/bipolar versus multipolar neurons (right panel) in the Parental and $\Delta Neurog2$ enh cell lines.

With these data the Results section of my doctoral thesis concludes. In the next section, I will discuss and support the obtained results taking into account the actual bibliography.

Discussion

In this section I will comment on my results taking into account the state of the art of the field. The discussed topics are ordered according to their order of appearance in the Results section.

1. About the epigenetic identification of enhancers

Our enhancer identification has been performed in agreement with previously published criteria (Rada-Iglesias et al., 2011). However, the number of enhancers that we obtained (3023) is low in comparison with other analysis in which the number of enhancers range from 32693 (FANTOM) to 213260 (Ensembl) (Fishilevich et al., 2017) (Table D1).

The sequencing of the human genome in 2001 shed light into the mapping of protein coding regions and enhancers using evolutionary conservation constrains (Nobrega et al., 2003; Venter et al., 2001). Nonetheless, sequence conservation does not provide information about the spatial-temporal activity of the enhancers. In this context, research articles published by Visel et al. in 2009 and 2013 contributed to the identification of active enhancers in the developing forebrain by performing ChIP-seq of the acetyltransferase p300 *in vivo* (Visel et al., 2009; Visel et al., 2013). Our identification of enhancers starts by using the p300 ChIP-seq performed in embryonic forebrains published in (Visel et al., 2013). This ChIP-seq identified a relatively low number of p300 binding sites: 4425. In our enhancer identification we combined p300+H3K4me1 features, thus the number of enhancers is limited by the low number of p300 peaks. Other enhancer identification criteria have used the combination of H3K4me1 and DNase I hypersensitive sites. These features undoubtedly achieve higher numbers of enhancers, however, we believe that the inclusion of p300 provides astringency and reduces the number of false positives, as the binding of chromatin regulators is an essential feature of enhancer chromatin (Rada-Iglesias et al., 2011).

Source	Number of enhancers
Ensembl	213260
FANTOM	42979
VISTA	1746
ENCODE	176154
This thesis	3020

Table D1. Number of identified enhancers in the different studies and databases. Adapted from (Fishilevich et al., 2017)

1.1. About the identified poised enhancers

Poised enhancers are associated to developmental genes and are typical from stem cell states. In human ESCs there have been identified 2287 poised enhancers

Discussion

(Rada-Iglesias et al., 2011) and in mouse ESCs 1016 (Cruz-Molina et al., 2017). Our data identify for the first time poised enhancers in NSCs. Using the H3K27me3 histone mark, we mapped 158 poised enhancers and we found that indeed, these are rare features in lineage committed cell types. This result is in agreement with the poised enhancer mapping in the human neural crest cells, other committed cell type, where researchers only found 351 poised regions (Rada-Iglesias et al., 2012). Interestingly, we observe that even if the number of poised enhancers is low, 74 become active along neuronal differentiation (47%) (Figure R3), thus we believe that their developmental idiosyncrasy is conserved between cell types.

2. About the TGF β -regulation of enhancers

The TGF β pathway has long been described to be a signaling pathway that operates on gene promoters (Morikawa et al., 2013). Interestingly, the genome-wide localization of SMADs has been related to the differentiation state of the cell, being preferentially bound to promoters and exons in ESCs and to introns and intergenic regions after germ layer specification (Kim et al., 2011). Our data are in agreement with the re-localization of SMAD3 postulated in the study of Kim et al. In mouse ESCs, SMAD3 is essentially bound to promoters (Whyte et al., 2013), in contrast, in NSCs, we have identified SMAD3 binding sites at enhancers, that usually lie at intergenic and intronic regions.

2.1. About the interaction between ASCL1 and SMAD3 at enhancers

The contextual versatility of the TGF β signaling is provided by the cooperation of the SMADs with different partners in different contexts (David and Massague, 2018). Interestingly, the DNA motif recognized by SMAD3 and SMAD4 (GTCTAGAC) cannot explain the ChIP-seq genomic occupancy of SMAD3. It has been observed, that SMAD3 colocalizes genome-wide with lineage specific TFs and together they activate the promoters of essential targets for cell differentiation (Mullen et al., 2011). In this doctoral thesis, we have identified a new lineage specific TF that cooperates with SMAD3 at enhancers, ASCL1. Importantly, we do not have only detected a 66% of SMAD3-bound enhancers also occupied by ASCL1 (Figure R6), but we have demonstrated that SMAD3 and ASCL1 interact together in NSCs (Figure R5), and that the genes putatively regulated by the enhancers governed by SMAD3 are regulated by ASCL1 (Figure R10). These results are important to decipher whether there could be a genomic functional cooperation. Enhancers are TFs binding platforms and it is expected for many TFs to bind there (Long et al., 2016). By simply comparing ChIP-seq peaks, there could be obtained many false positives in terms of protein colocalization. Transcriptional experiments and coimmunoprecipitations

are essential to dissect the contribution of the proteins to the enhancer activation model.

Another aspect that is worth discussing is the potential pioneering activity of ASCL1. At the beginning of these experiments we postulated that ASCL1 would be essential for SMAD3-sites recognition and binding based on the studies of Raposo et al. and Wapinski et al. (Raposo et al., 2015; Wapinski et al., 2013). However, ChIP-qPCR experiments of SMAD3 in shASCL1 NSCs did not show a clear result (data not shown). In some enhancers, ASCL1 was required for SMAD3 binding but in others the binding was unaffected. Nevertheless, we still hypothesize that ASCL1 is targeting SMAD3 to the enhancers as OCT4 is doing in ESCs (Mullen et al., 2011), but the dynamic nature of this interaction impedes its study *in vitro*. Upon binding, ASCL1 target regions are sequentially remodeled in a process that lasts days, and afterwards, TFs bind to these regions in a dynamic manner (Wapinski et al., 2017). By treating the NSCs with a recombinant TGF β 1 cytokine, the dynamic nature of the SMAD3 recruitment to the ASCL1 binding sites along neuronal differentiation is probably bypassed. To properly address this question additional kinetic experiments would be required and preferentially imaging techniques such as single molecule tracking should be performed (Liu et al., 2015).

2.2. About the TGF β activation of enhancers

Along the experimental work we have demonstrated *in vivo* and *in vitro* that TGF β activates the *Ctgf*(-102), *Nrip3*(-3.5) and *Neurog2*(-6) enhancers using three different approaches that reached similar results with minor caveats.

2.2.1. About the enhancer activation measured as eRNA transcription

The kinetic curves in Figure R11 show a good correlation in the magnitude of the mRNA and eRNA transcription. This result suggests that the control checkpoints for RNAPII fidelity in mRNA transcription (Sydow and Cramer, 2009) could be conserved in eRNA transcription. On the other hand, in the *Ctgf*(-102) enhancer, we observe a decrease of the mRNA levels starting at 3 hours while the eRNA levels are still high (Figure R11). This result is contrary to what is published concerning the stability of the eRNAs. One of the first papers showing transcription from the enhancers tested the stability of these RNA species, concluding that they were degraded long before the degradation of their cognate mRNAs (De Santa et al., 2010). Lacking the specific degradation experiments, we observe that the *Ctgf* eRNA seems more stable than the mRNA, however, *Ctgf*(-102) enhancer could be activating other genes whose mRNA transcription would still be high.

Discussion

2.2.2. About the enhancer activation measured as histone mark acquisition

The results obtained in these experiments are in agreement with the literature (Buecker and Wysocka, 2012). Nonetheless, the preferential enrichment in H3K4me2 over H3K4me1 on these enhancers was a surprising result. After running positive controls of the H3K4me1 ChIP-qPCR experiment, we hypothesized that the specific changes on H3K4me2 that we were observing (Figure R14) could be related to signal-dependent activation of enhancers, as increases in the H3K4me2 levels had not been reported along developmental or more stable cellular changes. A paper by Kaikkonen et al. is in agreement with our data. This article shows that macrophage activation upon lipopolysaccharide stimulation comprises the activation of a subset of enhancers regulated by H3K4me2 (Kaikkonen et al., 2013). We theorize that a TGF β -specific HKMT might be operating on these enhancers. The HKMT SETD7 could be a good candidate due to its implication in the methylation of promoters upon TGF β signaling (Sun et al., 2010) (Figure D1).

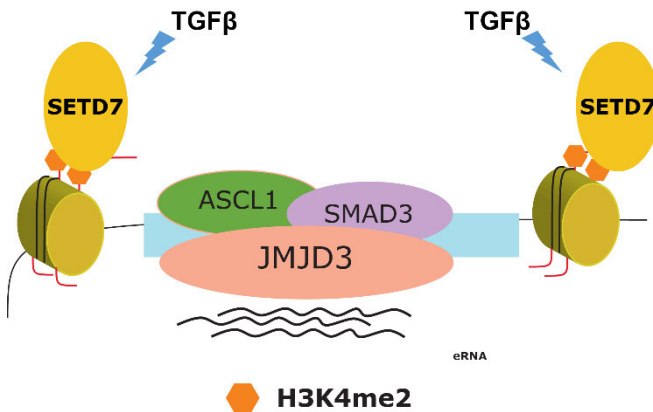


Figure D1. Cartoon of the hypothetical enhancer activation model in which SETD7 dimethylates the H3K4.

This implies that this gene and this enhancer are subjected to a tight transcriptional control that avoids its spurious activation. We postulate that the TGF β pathway alone is not sufficient to fully activate this enhancer. Our *in vitro* model allows the molecular dissection of the players participating in this event, however, we lack the *in vivo* niche where different developmental signals will converge permitting a total activation of the regulatory region. For other cell-identity regulators, we have observed that a total activation supposes an increase in the mRNA of more than 1000-fold (unpublished data), in the *Neurog2* gene case, we observe a mild activation (6-7 fold), so we would expect that having the proper conditions the *Neurog2* enhancer will properly acquire H3K27ac (Figure R51). Another potential explanation is the belonging of the *Neurog2* enhancer to a different class of enhancers described by Pradeepa et al. These enhancers display H3K64ac and H3K122ac upon activation.

I would also like to discuss the lack of H3K27ac acquisition in the *Neurog2(-6)* enhancer upon TGF β stimulation. As it will be further commented along this discussion, the *Neurog2(-6)* enhancer belongs to the poised enhancer category and, additionally, is a developmental enhancer associated to a gene that controls the neuronal cell fate (Wilkinson et al., 2013). This

These acetylations are located at the globular domain of the H3 and have been found at enhancers related to development (Pradeepa et al., 2016). Sadly, we have not tested whether *Neurog2*(-6) enhancer is enriched in these histone marks upon TGF β signaling, but as we have tested this enhancer activation with orthogonal techniques (eRNA transcription and luciferase) we defend that TGF β is activating this enhancer.

2.2.3. About the enhancer activation *in vivo*

Our luciferase results demonstrate that TGF β is activating the *Ctgf*(-102), *Neurog2*(-6) and *Nrip3*(-3.5) enhancers *in vivo* (Figure R16). *Ctgf*(-102) and *Nrip3*(-3.5) enhancers are inactive without TGF β . Interestingly, *Neurog2*(-6) enhancer is already active in basal conditions, and SMAD3 overexpression increases its reporter activity. This requirement of different stimuli to fully activate the *Neurog2*(-6) enhancer supports our hypothesis of *Neurog2*(-6) enhancer controlling cell-identity genes (section 2.2.2). The combination of mutants that we have used (SMAD3 S/D and SMAD3 S/A) strengthens our postulate of TGF β contributing to the activation of these enhancers.

Even though results favor our hypothesis, episomal reporter experiments are not the most rigorous approach to solve enhancer activation questions. As I have detailed along the Introduction and the Results sections, the 3D-chromatin conformation is the basis of the enhancer-promoter communication. Enhancers and promoters are binding platforms for TFs and thus, downstream luciferase reporters will function as a readout of the binding of transcriptional machinery. However, plasmidic DNA cannot acquire the proper 3D-conformation and thus, is not a representative context of the enhancer native location. Moreover, the promoter driving the luciferase expression is usually the SV40 promoter, a promoter with binding sites for many TFs that does not mimmick the endogenous conditions of activation (Zentner and Scacheri, 2012). These disadvantages can be overcome by introducing the classical LacZ reporter in the native chromatin using techniques such as enhancer-TRAP. This transgenic assay permits the *in vivo* functional analysis of enhancers in a high throughput manner, providing spatial and temporal data about the activity of enhancers along development (Kvon, 2015).

To test the enhancer activity, the combination of different approaches is essential to avoid misleading results. The linear and 3D aspects of the chromatin make this validation much more challenging, but at the same time more trustworthy and exciting.

Discussion

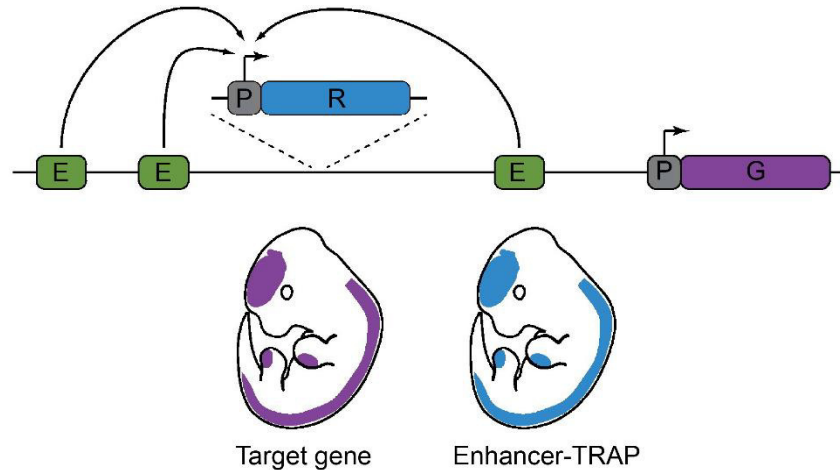


Figure D2. Schematic view of the enhancer-TRAP rationale. E: enhancer; P: promoter; R: reporter; G: gene. Adapted from (Kvon, 2015).

3. About the localization of JMJD3 at TGF β -activated enhancers

Even though we have modularly introduced the molecular players of the TGF β activation of enhancers, since the beginning of the work we hypothesized that JMJD3 was coregulating the cis-regulatory regions with SMAD3.

The fact that JMJD3 regulates enhancers has previously been reported in the DNA damage context (Williams et al., 2014), in endodermal differentiation (Kartikasari et al., 2013) and in postnatal and adult brain neurogenesis (Park et al., 2014). Among the results published in these papers, we find agreements as well as controversies in terms of the involvement of the catalytic activity of JMJD3 in the activation. Nonetheless, this will be commented in coming subsections. Here, I would like to compare our results with the results obtained in Park et al. 2014. In this report, they elegantly show the *in vivo* role of JMJD3 in postnatal neurogenesis. However, to point that JMJD3 is regulating neurogenic promoters and enhancers, they use our previously published JMJD3 ChIP-seq (Estaras et al., 2012) without mentioning that is performed after TGF β stimulation. Even though there might be common enhancers and promoters regulated by TGF β and other pathways in neurogenesis, our lab has shown that depending on the pathway, JMJD3 will activate different targets (Akizu et al., 2010). The number of enhancers that they identify (4425) and their gene ontology perfectly matches ours, thus supporting our enhancer identification.

3.1. About the interaction of JMJD3 and ASCL1

Figures R17 to R21 show that the KDM JMJD3 colocalizes and interacts with the bHLH TF ASCL1. In the section 2.1 of this discussion I have argued that ASCL1 could be participating in the TGF β enhancer activation as a pioneer TF. Here, I would

like to postulate a potential role of JMJD3 and ASCL1 as partners at the onset of the neurogenic program. Results published by Dai et al. (Dai et al., 2010) show that JMJD3 demethylates the promoter of *Ascl1* leading to a high expression of ASCL1 that promotes the beginning of the neurogenesis. Additionally, our data demonstrate that JMJD3 colocalizes and interacts with ASCL1 at enhancers, and a potential ASCL1 enhancer is regulated by JMJD3. These results point to JMJD3 as an essential epigenetic regulator during stem cell commitment (Figure D3). Accordingly, results in endoderm differentiation show that JMJD3 activates and cooperates with the master TF EOMES to trigger endodermal differentiation (Kartikasari et al., 2013). We hypothesize that JMJD3 is controlling the expression of cell-identity genes to start differentiation programs through positive feedback regulatory loops (Figure D3).

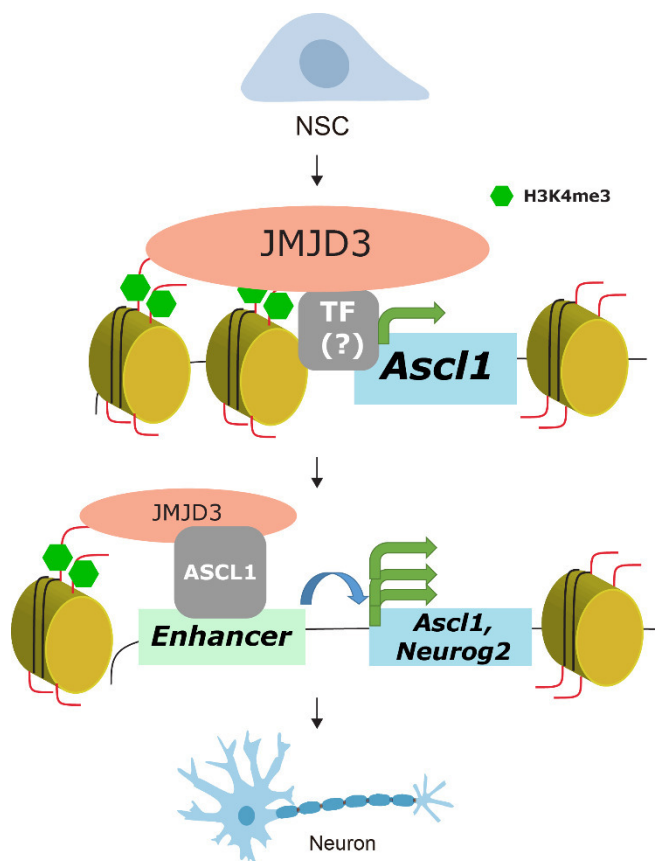


Figure D3. Cartoon model of the hypothetical cooperation between ASCL1 and JMJD3 to promote neurogenesis.

3.2. About the role of JMJD3 in the activation of enhancers

An essential point to highlight in this subsection is that the majority of enhancers do not contain H3K27me3 (Buecker and Wysocka, 2012). In relation to JMJD3, this points directly to catalytic-independent roles of this enzyme at cis-regulatory regions. In the literature, there are examples claiming that the H3K27me3 catalytic activity is necessary to activate specific enhancers (Dai et al., 2010; Kartikasari et al., 2013).

Discussion

However, we and others support that the removal of H3K27me3 is not the main function of JMJD3 at enhancers (Williams et al., 2014). Results in Figure R28 and R29 indicate that the H3K27me2/3 levels are not altered in the analyzed poised enhancers during activation in a JMJD3-dependent manner. Studying the data, two mechanisms might be envisioned to explain the role of JMJD3. In *Chic2*(-26) and *Tle3*(-114) enhancers, the H3K27me3 levels decrease upon TGF β in a JMJD3-independent manner. At these enhancers, JMJD3 function might not be related to H3K27me3, and could be linked to demethylation of other locus like the promoters (Figure R30). Interestingly, another potential role of JMJD3 at these enhancers could be the release of the RNAPII at the TSS of the related promoters. Our lab has previously published that JMJD3 is required for the RNAPII-S2p elongation. The lack of JMJD3 abolishes mRNA transcription and the presence of the RNAPII inside the genes (Estaras et al., 2013). The addition of the enhancers to this equation includes a new layer of transcriptional regulation governed by the enhancer, the idea of the enhancer as a provider of JMJD3 molecules for the functions at the promoter is an efficient manner of controlling the unspecific beginning of transcription before all the machinery is set. It has been described that another demethylase with activity towards arginines, JMJD6, participates in the RNAPII pause-release by cooperating with the elongation factor BRD4 and demethylating histone and non-histone substrates (Liu et al., 2013). We cannot rule out that JMJD3 could be playing a similar role. However, in an *in silico* screening for potential demethylation sites in non-histone substrates (PTEFb, RNAPII and others...) (data not shown) we could not find suitable chemical environments for the action of JMJD3 (Chen et al., 2006). Regarding other poised enhancer, *Neurog2*(-6), the increase observed in H3K27me3 upon TGF β in shJMJD3 cells might be related to the ability of the PRC2 complex to bind this enhancer. It has been recently shown that PRC2 might be working as an activator of neural poised enhancers, by facilitating loop formation (Cruz-Molina et al., 2017). Thus, it is possible that in the absence of JMJD3, EZH2 could be targeting more efficiently the enhancer increasing the levels of H3K27me3 in response to TGF β signal.

Along this doctoral thesis, we have tested the participation of JMJD3 in enhancer activation studying its role in eRNA transcription and histone mark acquisition. These two events have been related in the literature, for example, Dorigi et al. demonstrated with catalytic mutants of MLL3/4 (the main methyltransferases of enhancers) that the H3K4me1 was not necessary for enhancer function; however, the proper MLL3/4 proteins are. The enhancers in MLL3/4 KO cells did not gain RNAPII upon activation, and consequently did not produce eRNAs (Dorigi et al., 2017). In other research article, the authors showed that transcription of eRNAs correlated with other enhancer features (H3K4me1, H3K27ac etc.) (Hah et al., 2013), they inhibited the transcription of eRNAs with the CDK9 inhibitor flavopiridol, and they observed that all the enhancer features were properly acquired. These two

works lead to the conclusion of a dynamic series of events in which upon activation, the enhancers first gain the activating histone marks and then, the eRNAs are produced in a MLL3/4 dependent manner. This hypothetical mechanism is not directly in agreement with our results in Figures R24 and R26. In those results we observe that the removal of JMJD3 has only effect in histone marks acquisition in the *Ctgf*(-102) enhancer; however, the transcription of eRNAs is almost unaffected. On the other hand, at *Nrip3*(-3.5) and *Neurog2*(-6) enhancers, the removal of JMJD3 does not have any effect on histone marks but it does have a dramatic effect at the eRNA transcription level.

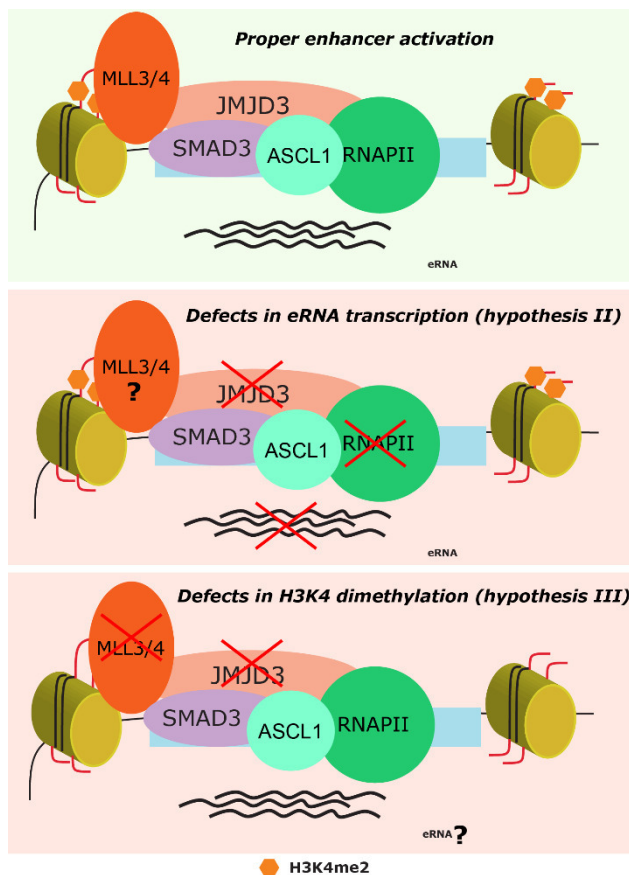


Figure D4. Cartoon model representing the hypothesis for enhancer malfunctioning in shJMJD3 NSCs.

Here, we provide some potential explanations for these disagreements (Figure D4):

- (I) In our results there are more than one source of enhancer malfunctioning, and the final output that we are obtaining is a mixture of different events.
- (II) JMJD3 interacts with the RNAPII and has been demonstrated to be required for elongation (Chen et al., 2012; Estaras et al., 2013). At the selected enhancers, the lack of JMJD3 could be impairing the recruitment of RNAPII independently of the histone modifiers.

Discussion

- (III) JMJD3 interacts with members of the MLL complex ASH2L and WDR5 (De Santa et al., 2007). We hypothesize that JMJD3 could serve as a scaffold for the binding of proteins of the MLL complex and the chromatin remodeler CHD8. In the shJMJD3 NSCs, the partial absence of JMJD3 could lead to a poor recruitment of the MLL complex of proteins and consequently to enhancer misregulation.

Our work does not uncover the role of JMJD3 at enhancers. However, the results obtained in the 4C-seq experiments (commented in subsection 5.3) point to a mixture of enzymatic and structural roles.

4. About the contribution of CHD8 to the TGF β enhancer activation

The participation of chromatin remodelers in enhancer activation has been suggested since the beginning of the investigations in this field. In fact, one accepted method to identify enhancers is the binding of BRG1, a chromatin remodeler (Rada-Iglesias et al., 2011). With our results, we claim that the CHD family remodeler, CHD8, is cooperating with JMJD3 and the TGF β pathway to activate enhancers. Which role do remodelers play at enhancers is an intriguing question. By definition, enhancers are regions depleted of nucleosomes, sensitive to DNase I treatment (Heintzman et al., 2009). We postulate that the role of these proteins at enhancers is the maintenance of open chromatin regions that permit the binding of TFs or the cooperation with pioneer TFs.

Previous results in the literature have demonstrated that JMJD3 interacts with another member of the CHD family, CHD7 (Chen et al., 2012). In our work, we have focused on CHD8 because in other models, it had been related to elongation as JMJD3 (Rodriguez-Paredes et al., 2009; Srinivasan et al., 2008). However, preliminary results and the literature, place CHD7 at many of the JMJD3-regulated enhancers (Engelen et al., 2011). CHD8 and CHD7 belong to the same subfamily of CHD remodelers, they interact together and both are highly expressed in embryonic cortex of E12.5 mice embryos (Batsukh et al., 2010). Surprisingly, these enzymes exert non-redundant roles in the living organisms, CHD7 deficiency leads to CHARGE syndrome (Vissers et al., 2004) and CHD8 deficiency is associated with autism spectrum disorders. An elegant biochemical research article by Manning et al. demonstrated that CHD8 was able of sliding nucleosomes by binding long segments of linker DNA, on the other hand, CHD7 required smaller segments. These biochemical differences could account for the different loss of function phenotypes and for the mixture of activities of CHD7 and CHD8. However, with our data we cannot rule out the possibility of CHD7 also participating in *Ctgf(-102)*, *Nrip3(-3.5)* and *Neurog2(-6)* enhancers activation.

One of the major caveats to fully understand the cooperation of SMAD3/JMJD3 and CHD8 is the lack of a CHD8 ChIP-seq. I would like to mention, that we performed this experiment and the results were not good enough to properly call peaks. Other ChIP-seqs have been published and they map CHD8 mainly at promoters, even though they identify CHD8 binding at distal regulatory regions with enhancer features (Cotney et al., 2015; Sugathan et al., 2014).

Our work states for the first time the implication of CHD8 in the TGF β pathway and its endogenous interaction with JMJD3. High throughput analysis have shown that upon removal of CHD8, expression of genes belonging to the Wnt pathway and the p53 signaling cascade are impaired (Nishiyama et al., 2009; Nishiyama et al., 2012; Thompson et al., 2008). These results reinforce our conclusions because JMJD3 has been described to be essential in the Wnt pathway (Jiang et al., 2013) and in the p53 signaling triggered by DNA damage (Williams et al., 2014), we postulate, that in these cascades where CHD8 and JMJD3 are essential, they might be cooperating as well as in NSCs.

5. About the TGF β re-organization of the chromatin

Until now, I have discussed the results concerning the linear aspects of the chromatin. In these subsections I will support the results obtained regarding the TGF β regulation of the chromatin structure.

5.1. About the design of the 4C-seq experiment

When designing the 4C-seq experiment our goals were two. First, analyzing whether TGF β triggered changes in the 3D-structure of the chromatin, and observing if JMJD3 contributed to these events; and second, studying the putative contacts between the viewpoint and the intragenic region of *Chst8* in relation to the transcriptional elongation. The second objective limited our selection, because we required a sufficiently long gene for detecting with enough resolution the contacts. For this reason, we selected the *Chst8* gene, that is 138 Kb long. It is worth noting, that we have not found contacts along the *Chst8* intragenic region that could point to sequential encounters of the enhancer with the body of the gene. However, I will provide rationales about our hypothesis.

JMJD3 colocalizes with the RNAPII-S2p inside genes, and JMJD3 is required for TGF β -promoted transcriptional elongation. In a research article by Lee et al. (Lee et al., 2015), it was proposed, that there are dynamic enhancer contacts throughout the gene bodies that track with RNAPII-S2p. The authors proposed a model in which the gene bodies changed their position relative to a stable enhancer-promoter complex. Other papers have shown that the elongation machinery has been found at

Discussion

enhancers (Lin et al., 2013; Wang et al., 2017). Altogether these data supported our hypothesis of the enhancer contacting the body of the gene during elongation and our selection of a long gene for the experiment. However, our data did not show contacts between the viewpoint and all the intragenic region of the gene. We did observe isolated contacts spanning the first 50 Kb that could suggest a relation to elongation (Figure R43); however, with the levels of expression of the *Chst8* gene after 3 hours of TGF β treatment (Fold change of 7-8, Figure R39) and being our sample the result of a population of cells, we expected contacts all along the gene. Nonetheless, replication of the experiment and analysis of other enhancers are required to completely dismiss the elongation hypothesis.

In terms of resolution, we have performed the most resolutive 4C-seq variant, where two rounds of 4-bp cutters restriction digestions have been used. This protocol led us to a resolution of around 0.5 Kb, more than the 1 Kb that can be nowadays achieved with HiC (Rao et al., 2014).

5.2. About the TGF β -induction of a functional superenhancer formation

In the literature there are many reports and hundreds of HiC experiments highlighting the communication between enhancers and promoters (Zabidi and Stark, 2016). However, the majority of these research articles do not highlight one type of genomic interaction that occurs further more often than the enhancer-promoter interaction: the enhancer-enhancer contact. The promoter-centered selection of viewpoints in 4C-seq experiments and the necessity of relating the enhancer-promoter contacts to transcriptional events have biased the conclusions of many experiments. Lately, some reports are starting to remark the existence of thousands of enhancer-enhancer contacts (Beagrie et al., 2017; Proudhon et al., 2016) that could be fine-tuning the levels of mRNA transcription, regulating the spatial-temporal activation of genes or acting as safeguards for transcriptional machinery provision.

The results of our 4C-seq experiment show that upon TGF β signaling, the *Chst8*(+38) enhancer contacts the *Chst8* promoter and other enhancers occupied by SMAD3 and JMJD3 (Figure R43). These results led us to hypothesize that TGF β is triggering a *functional* superenhancer formation. The term *functional* superenhancer has never been used in the literature, it was adopted by our lab to describe a superenhancer that instead of regulating cell-identity genes, as classical superenhancers (Whyte et al., 2013), regulates genes related with a specific cellular function, in our case, the TGF β pathway. In the literature, there has been controversy on whether superenhancers are entities distinct from the classical regulatory regions. It is known, that clustering of enhancers is neither required nor sufficient for superenhancer formation, and importantly, the cutoff for enhancer inclusion in the definition of superenhancer (in terms of ChIP-seq binding of Mediator and TFs) sounds still arbitrary (Pott and Lieb, 2014). Nonetheless, we will be using the term

functional superenhancer to name the cooperativity nature of enhancers for the potential activation of a downstream gene.

Undoubtedly, we require further data to confirm the coordinated activity of the *Chst8* *functional* superenhancer. However, we postulate that 4C-seq experiments using as viewpoints other enhancers of the cluster will show that these are also contacting the *Chst8* promoter, other enhancers or our actual viewpoint. Similar studies in V(D)J recombination have shown that the enhancers located at the *IgK* superenhancer are in contact with each other. Besides, the CRISPR-Cas9 deletion of any of these enhancers leads to the collapse of the 3D structure that ensures *IgK* expression (Proudhon et al., 2016). We hypothesize that the synergistic action of the enhancers in the *functional* superenhancer brings together high concentrations of TFs (ASCL1, SMAD3) and cofactors (JMJD3, CHD8 etc.) to the *Chst8* promoter upon TGF β signaling ensuring a quick and high response. It would be interesting to combine transcriptional data with imaging experiments to try to decipher whether there is any relation between the number of mRNA copies and the number of molecules of SMAD3/JMJD3 that are bound to the promoter or the enhancer.

5.3. About the contribution of JMJD3 to the functional superenhancer formation

In the Figures R44, R45 and R46 it is observed that JMJD3 is necessary for the establishment of contacts between the viewpoint and the *Chst8* promoter, and the viewpoint and other enhancers. We hypothesize that as JMJD3 seems to exert catalytic-independent roles in the activation of enhancers, it could be participating as a platform for protein-protein interaction mediating the enhancer-enhancer and enhancer-promoter contacts (Figura D5). With the actual data we cannot elucidate the role of JMJD3 in the contact formation; however, I will try to elaborate on potential roles of this demethylase:

- (I) As JMJD3 does not have a DNA-binding domain, it looks unlikely that it could be mediating the contacts by binding at the involved regions and bringing them together. This hypothesis will suppose that JMJD3 is sufficiently big to cover the distance between enhancers and promoter/enhancers or that JMJD3 dimerizes or oligomerizes. Additionally, the proteins that have been demonstrated to carry out this action are both zinc-finger proteins that bind DNA at both sides of the contacts (CTCF, YY1) (Beagan et al., 2017; Ohlsson et al., 2001), and JMJD3 does not have any structural similarity with them.

Discussion

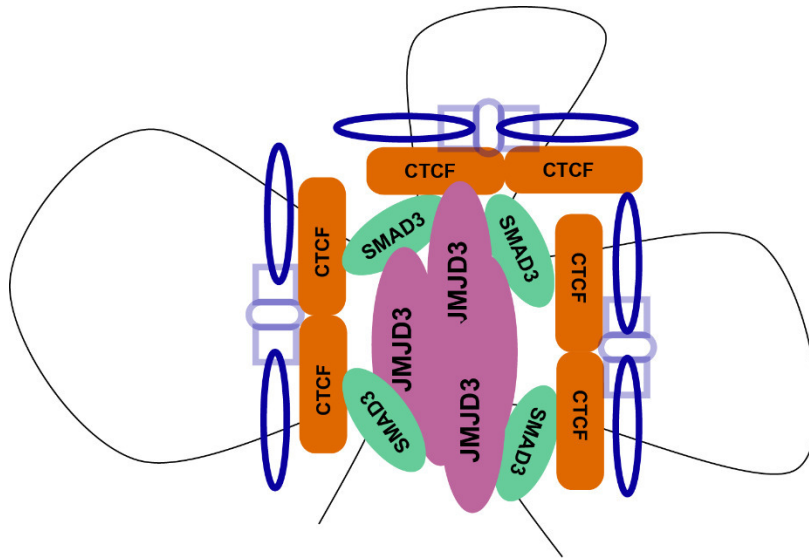


Figure D5. Cartoon model of the functional superenhancer formation. The blue modules represent the Cohesin complex

- (II) The protein JMJD3 is a big protein (1643 amino acids, 250 KDa) with just one known domain (JmJC domain, 164 amino acids). It is logical to think that JMJD3 could be serving as a platform for protein-protein interactions that will regulate genomic contacts. Strikingly, looking at the protein sequence of JMJD3 we have observed that the abundance of prolines is enormous. Prolines are highly hydrophobic amino acids whose concatenation generates sticky domains that bind rapidly and reversibly to other proteins (Williamson, 1994). The high content of prolines of JMJD3 (Figure D6) could be promoting the formation of aggregates of proteins and DNA that would generate phase transitions that would isolate and provide stability to the genomic contacts without the requirement of lipidic membranes (Figure D5). Our lab is now working to further test these hypothesis to ultimately uncover the JMJD3 function at enhancers.
- (III) Other hypothesis on how JMJD3 is mediating the contacts relies on the interaction between SMAD3 and JMJD3. At some gene promoters, JMJD3 is recruited by SMAD3 upon TGF β signaling and the lack of SMAD3 impedes the JMJD3 recruitment (Estaras et al., 2012). However, we have not tested if this is the case genome-wide and we cannot rule out the possibility of JMJD3 recruiting SMAD3 at specific regions. If this was the case, in the shJMJD3 NSCs SMAD3 will not be recruited and chromatin loops could be impaired due to the interaction of SMAD3 and CTCF. These two proteins have been demonstrated to coimmunoprecipitate, and the binding of CTCF to SMAD3 bound promoters is dependent on SMAD3

(Chen et al., 2016; Weth and Renkawitz, 2011). With impaired levels of SMAD3, CTCF will not be so efficiently recruited and the loop anchors will not be stable. This theory could answer the lack of some of the contacts, however, not all the lost contacts are occupied by CTCF in basal conditions (data not shown), so there must be other coexistent mechanisms.

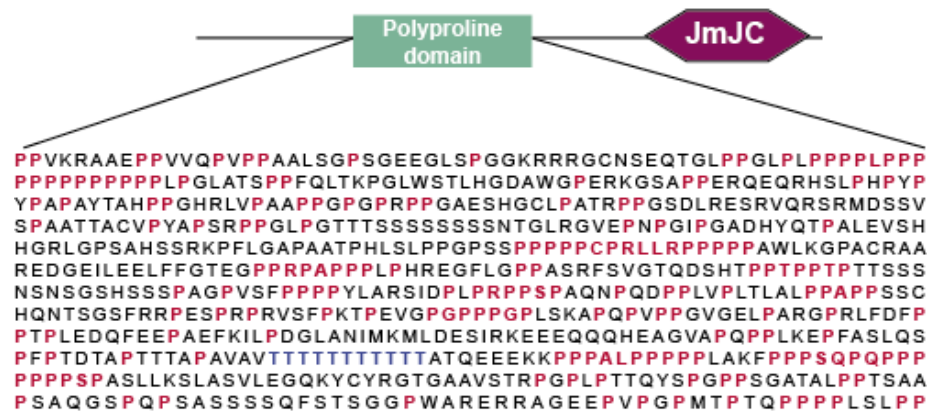


Figure D6. Scheme of the protein JMJD3. The amino acid content of the polyproline domain is indicated. Prolines are marked in pink, a track of threonines is marked in blue.

6. About the *Neurog2(-6)* enhancer and *Neurog2* promoter cooperation in the fine-tuning of transcription

Most of the enhancers that are occupied by SMAD3 and JMJD3 are putatively regulating genes whose promoters are also occupied by SMAD3 and JMJD3. This fact makes difficult the distinction between the effect that can be attributed to promoters and to enhancers. With our CRISPR-Cas9 approach we have tried to dissect the contribution of every genomic region to the final cellular output. Other studies have disrupted chromatin loops to study the contribution of CTCF or Cohesin complexes to the enhancer-promoter communication (de Wit et al., 2015; Rao et al., 2017). However, these studies did not assess the contribution of the individual elements to the final transcriptional response.

Overall, our conclusion is that both enhancers and promoters are essential to maintain a faithful transcriptional activation upon TGF β signaling. However, I will comment more on the individual participation.

Discussion

6.1. About the contribution of promoters to transcription

The RNAPII molecules start mRNA productive elongation at promoters (Kwak and Lis, 2013). By definition, the removal of the promoter would lead to the total ablation of the mRNA transcription as RNAPII and the GTFs would not have a proper transcription initiation site. However, many genes contain alternative promoters whose usage permits the increase in the number of protein isoforms that could be generated in a specific cell type (Vacik and Raska, 2017). From the beginning, we doubted that this could be the case for the *Neurog2* gene as it does not have alternative isoforms. Nonetheless, if we assume that the *Neurog2*(-6) enhancer is in close proximity to the *Neurog2* promoter, the enhancer could act as a promoter to ensure *Neurog2* expression. Results in Figure R53 show that this is not the case, removal of the promoter totally abolishes *Neurog2* transcription upon TGF β treatment. Interestingly, intragenic enhancers have been described to function as alternative promoters (Kowalczyk et al., 2012), it would be insightful to test whether the intragenic enhancers of the *functional* superenhancer could work as alternative promoters for the *Chst8* gene.

Surprisingly, the deletion of the *Neurog2* promoter generates a transcription accumulation at the enhancer. In Figure R53, it is observed that the levels of eRNAs upon TGF β signaling are almost the triple in the Δ *Neurog2* pro cell line than in the Parental cell line. This could be indicating several things:

- (I) On the one hand and taking into account that poised enhancers are in close 3D proximity to their cognate promoters before activation (Cruz-Molina et al., 2017), the lack of promoter could be generating an accumulation of RNAPII molecules at the enhancer thus increasing the transcription of eRNAs.
- (II) On the other hand, the promoter could be recruiting machinery for eRNA degradation and its absence will also end in accumulation.

Interestingly, the work of Magnuson et al. backs both hypothesis. By performing BruUV-seq, a technique that permits the enhancement of nascent RNA detection, they found that blocking the mRNA transcription highly induces the transcription of eRNAs. When RNAPII is stalled, the nuclear RNA exosomes cannot degrade RNAs as long as their 3'-ends are attached to the RNAPII (Magnuson et al., 2015; Tiedje et al., 2015). ChIP-qPCR experiments of the RNAPII would help us to test this hypothesis.

Another experiment that we performed with the Δ *Neurog2* pro cell line was a ChIP-qPCR with JMJD3. Results in Figure R54 show that the deletion of the *Neurog2* promoter impedes the recruitment of JMJD3 at the *Neurog2*(-6) enhancer. There are

many explanations for this impairment. However, the lack of reports describing promoter deletions difficults the orientation of the hypothesis. We postulate, that:

- (I) The deletion of the promoter could be affecting the binding affinity of the JMJD3 molecules that are shared between the enhancer and the promoter and thus, after TGF β signaling JMJD3 could not be efficiently recruited to the enhancers.
- (II) Another possibility is that the removal of the promoter generates a 3D-conformational change in the chromatin fiber that situates the enhancer in an environment where JMJD3 (or SMAD3) molecules cannot penetrate. When a chromatin loop is destroyed, others might be formed rewiring the enhancer-promoter contacts (Qiu et al., 2018).

One controversial result in Figures R53 and R54 is that in the Δ *Neurog2* pro cell line we observe an accumulation of eRNAs and a lack of JMJD3 at the same time. These data are opposite to what we have observed in the analysis of eRNA transcription in the shJMJD3 cell line (Figure R24), where eRNA transcription is impaired when JMJD3 is depleted. This result, as contradictory as it seems is justifiable if we point to the different recruitment of RNAPII as the responsible for the output. In the shJMJD3 NSCs, the lack of transcription of eRNAs might be explained as a secondary effect upon JMJD3 depletion. JMJD3 through the recruitment of chromatin modifiers promotes RNAPII transcription at the enhancer. On the other hand, we postulate that in the Δ *Neurog2* pro NSCs we start with an accumulation of RNAPII at the enhancer that is not dependent on JMJD3 and occurs before the treatment with TGF β , thus bypassing the absence of JMJD3.

6.2. About the contribution of enhancers to transcription

Performing experiments with the Δ *Neurog2* enh NSCs we have learned that enhancers are essential in promoter activation. Enhancers should not longer be seen as modulators of the promoter activity, but as necessary modules for transcriptional activation, that provide cell-specificity and dose control.

With the Δ *Neurog2* enh NSCs, we show that *Neurog2*(-6) poised enhancer is necessary for the induction of its target gene upon TGF β signaling activation (Figure R55). Interestingly, *Neurog2*(-6) enhancer deletion did not result in the activation of the *Neurog2* gene in NSCs, supporting the idea that the poised enhancers have not repressive activity (Cruz-Molina et al., 2017). Major developmental and cell-identity genes are frequently regulated by multiple and sometimes redundant enhancers that confer phenotypical robustness to the living organisms (Frankel et al., 2010). However, our results clearly demonstrate that the deletion of a single poised

Discussion

enhancer totally blocks the expression of *Neurog2* after TGF β compromising normal neuronal differentiation. This extends previous observations indicating that certain enhancers can control gene expression in a nonredundant manner in different cellular contexts (Cruz-Molina et al., 2017; Huang et al., 2016).

At the mechanistic level, we observe that the *Neurog2(-6)* enhancer is essential for mRNA transcription from the *Neurog2* gene (Figure R55). We postulate that the impaired recruitment of JMJD3 to the *Neurog2* promoter, ultimately leads to a lack of H3K27me3 demethylation that impedes *Neurog2* gene activation (Figures R56 and R57). Observed in isolation, these series of events imply that the molecules of JMJD3 that act over the *Neurog2* promoter are provided from the enhancer. However, the global analysis of the enhancer and promoter contribution demonstrates that enhancers and promoters are equally important for JMJD3 recruitment, thus supporting the hypothesis of JMJD3 as a molecule that serves to facilitate enhancer-enhancer and enhancer-promoter contacts (Discussion, subsection 5.3).

6.2.1. About the contribution of the *Neurog2(-6)* enhancer to neurogenesis

We have demonstrated that the *Neurog2(-6)* enhancer controls the *Neurog2* gene promoter in E12.5 NSCs in a non-redundant manner.

Our differentiation experiments in Figure R58 show that the Δ *Neurog2* enh NSCs do not differentiate into correctly polarized neurons. However, both Parental and Δ *Neurog2* enh NSCs express the neuronal marker TUJ1, indicating that neuronal differentiation is not totally abolished. *In vivo* experiments with NEUROG2 KO mice have shown that the bHLH TF ASCL1 can compensate the lack of NEUROG2 in neuronal differentiation (Hufnagel et al., 2010), thus explaining the neuronal fate acquisition. Furthermore, our results are in good agreement with the phenotype observed in experiments performed in NEUROG2 KO pyramidal neurons. In the work of Hand et al. it was demonstrated that the absence of NEUROG2 impairs the dendritic morphology and the migratory capacity of neurons. Results displayed in Figure R59 show that the number of neurites per neuron is profoundly affected in the pseudoneurons generated from the Δ *Neurog2* enh NSCs. This impaired neuritogenesis can be explained by looking at the downstream targets of NEUROG2 during neuronal differentiation: cytoskeleton genes such as *Rnd2*, *Cdc42* or *Dcx*. We tested their expression in Parental and Δ *Neurog2* enh cell lines and we observed that the latter showed a clear lack of induction of these genes upon differentiation (Fueyo et al., 2018).

7. Global analysis of the results

During neuronal lineage commitment, the transcriptional programs orchestrated by the signaling pathways, TFs and epigenetic regulators lead to phenotypical changes that provide NSCs with neuronal features (Temple, 2001).

In this context, this doctoral thesis has contributed to the identification of molecular players participating in the neuronal differentiation triggered by TGF β and has also shed light into general mechanisms governing transcription activation through enhancers.

In summary, our results highlight enhancers as TF binding platforms where different modifying enzymes coordinate their activities to induce faithful gene activation. This doctoral thesis uncovers the molecular mechanism responsible for enhancer activation in response to TGF β signaling. This activation involves the action of the cofactors JMJD3 and CHD8, that by remodeling enhancers, previously pre-marked by ASCL1, activate the neuronal commitment program. At the 3D level, TGF β reorganizes the chromatin fiber in a JMJD3-dependent manner, promoting the establishment of enhancer-enhancer and enhancer-promoter contacts that will ultimately define the transcriptional output (Figure D7).

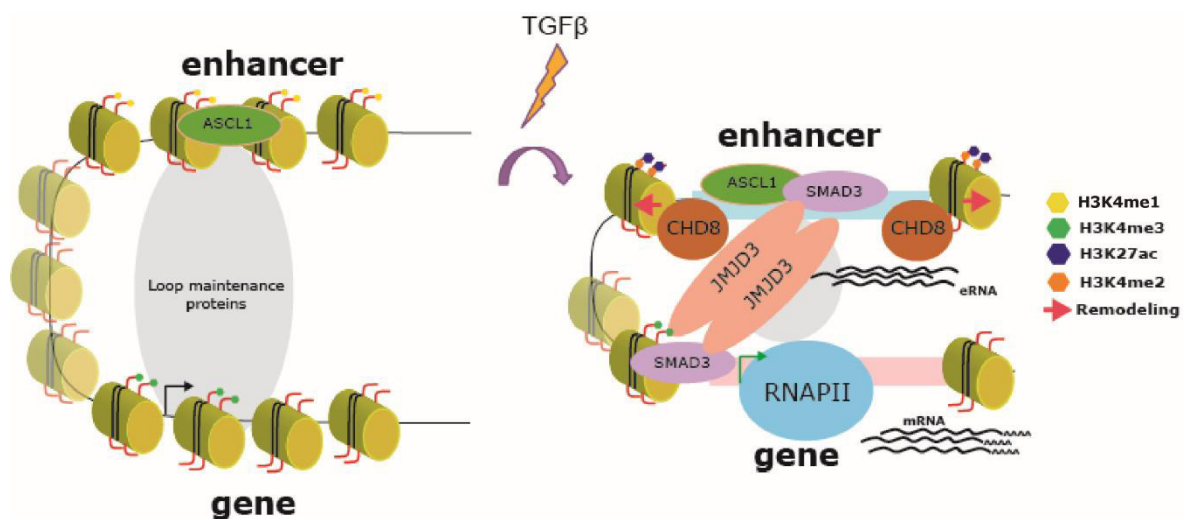


Figure D7. Cartoon model showing the molecular players that participate in the TGF β -triggered transcriptional activation in the 3D-structure of the chromatin.

Due to the broad range of TGF β functions in areas of cancer and neurodegenerative disorders, this work paves the way for investigating the SMAD3/JMJD3 contribution to transcriptional regulation in other cellular contexts and helps to move forward our understanding of the myriad of crosstalks between epigenetics and developmental program.

Conclusions

Conclusions

1. The epigenetic identification of enhancers in NSCs leads to the mapping of 3020 enhancers. Among these, 1174 are active, 1549 are primed, 158 are poised and 139 are heterogeneous.
2. The major effector of the TGF β pathway SMAD3 occupies and activates enhancers in NSCs upon TGF β stimulation.
3. TGF β activates enhancers *in vivo* in the chicken neural tube.
4. SMAD3 interacts with the proneural bHLH TF, ASCL1, in NSCs.
5. 66% of the enhancers occupied by SMAD3 contain ASCL1.
6. SMAD3 and ASCL1 coregulate transcription in NSCs
7. 603 enhancers associated to forebrain development are bound by ASCL1, SMAD3 and the histone demethylase JMJD3 in NSCs.
8. JMJD3 is essential for enhancer activation upon TGF β stimulation.
9. The demethylation of the H3K27me3 is not indispensable for enhancer activation.
10. JMJD3 interacts with the chromatin remodeler CHD8 and recruits it to the enhancers after TGF β signaling.
11. CHD8 is necessary for enhancer activation upon TGF β stimulation.
12. TGF β signaling re-organizes the genomic contacts triggering enhancer-enhancer and enhancer-promoter interactions.
13. JMJD3 is required for the establishment of enhancer-enhancer and enhancer-promoter contacts upon TGF β signaling.
14. The *Neurog2(-6)* enhancer and the *Neurog2* promoter are both required for a faithful TGF β response of the *Neurog2* gene.
15. The *Neurog2(-6)* enhancer is essential for proper neuronal differentiation.

Bibliography

- Afgan, E., Baker, D., van den Beek, M., Blankenberg, D., Bouvier, D., Cech, M., Chilton, J., Clements, D., Coraor, N., Eberhard, C., *et al.* (2016). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Res* 44, W3-W10.
- Agger, K., Cloos, P.A., Rudkjaer, L., Williams, K., Andersen, G., Christensen, J., and Helin, K. (2009). The H3K27me3 demethylase JMJD3 contributes to the activation of the INK4A-ARF locus in response to oncogene- and stress-induced senescence. *Genes Dev* 23, 1171-1176.
- Akhurst, R.J., and Hata, A. (2012). Targeting the TGFbeta signalling pathway in disease. *Nat Rev Drug Discov* 11, 790-811.
- Akizu, N., Estaras, C., Guerrero, L., Marti, E., and Martinez-Balbas, M.A. (2010). H3K27me3 regulates BMP activity in developing spinal cord. *Development* 137, 2915-2925.
- Banerji, J., Rusconi, S., and Schaffner, W. (1981). Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. *Cell* 27, 299-308.
- Bannister, A.J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. *Cell research* 21, 381-395.
- Bascom, G., and Schlick, T. (2017). Linking Chromatin Fibers to Gene Folding by Hierarchical Looping. *Biophys J* 112, 434-445.
- Batsukh, T., Pieper, L., Koszucka, A.M., von Velsen, N., Hoyer-Fender, S., Elbracht, M., Bergman, J.E., Hoefsloot, L.H., and Pauli, S. (2010). CHD8 interacts with CHD7, a protein which is mutated in CHARGE syndrome. *Hum Mol Genet* 19, 2858-2866.
- Batsukh, T., Schulz, Y., Wolf, S., Rabe, T.I., Oellerich, T., Urlaub, H., Schaefer, I.M., and Pauli, S. (2012). Identification and characterization of FAM124B as a novel component of a CHD7 and CHD8 containing complex. *PLoS One* 7, e52640.
- Beagan, J.A., Duong, M.T., Titus, K.R., Zhou, L., Cao, Z., Ma, J., Lachanski, C.V., Gillis, D.R., and Phillips-Cremens, J.E. (2017). YY1 and CTCF orchestrate a 3D chromatin looping switch during early neural lineage commitment. *Genome Res* 27, 1139-1152.
- Beagrie, R.A., Scialdone, A., Schueler, M., Kraemer, D.C., Chotalia, M., Xie, S.Q., Barbieri, M., de Santiago, I., Lavitas, L.M., Branco, M.R., *et al.* (2017). Complex multi-enhancer contacts captured by genome architecture mapping. *Nature* 543, 519-524.
- Bednar, J., Horowitz, R.A., Grigoryev, S.A., Carruthers, L.M., Hansen, J.C., Koster, A.J., and Woodcock, C.L. (1998). Nucleosomes, linker DNA, and linker histone form a unique structural motif that directs the higher-order folding and compaction of chromatin. *Proceedings of the National Academy of Sciences of the United States of America* 95, 14173-14178.
- Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., *et al.* (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 125, 315-326.
- Blanco-Garcia, N., Asensio-Juan, E., de la Cruz, X., and Martinez-Balbas, M.A. (2009). Autoacetylation regulates P/CAF nuclear localization. *J Biol Chem* 284, 1343-1352.
- Bochynska, A., Luscher-Firzlaff, J., and Luscher, B. (2018). Modes of Interaction of KMT2 Histone H3 Lysine 4 Methyltransferase/COMPASS Complexes with Chromatin. *Cells* 7.
- Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72, 248-254.
- Buecker, C., and Wysocka, J. (2012). Enhancers as information integration hubs in development: lessons from genomics. *Trends in genetics : TIG* 28, 276-284.
- Burchfield, J.S., Li, Q., Wang, H.Y., and Wang, R.F. (2015). JMJD3 as an epigenetic regulator in development and disease. *Int J Biochem Cell Biol* 67, 148-157.
- Burgold, T., Spreafico, F., De Santa, F., Totaro, M.G., Prosperini, E., Natoli, G., and Testa, G. (2008). The histone H3 lysine 27-specific demethylase Jmjd3 is required for neural commitment. *PLoS One* 3, e3034.

Bibliography

- Burgold, T., Voituron, N., Caganova, M., Tripathi, P.P., Menuet, C., Tusi, B.K., Spreafico, F., Bevengut, M., Gestreau, C., Buontempo, S., *et al.* (2012). The H3K27 demethylase JMJD3 is required for maintenance of the embryonic respiratory neuronal network, neonatal breathing, and survival. *Cell Rep* 2, 1244-1258.
- Ceballos-Chavez, M., Subtil-Rodriguez, A., Giannopoulou, E.G., Soronellas, D., Vazquez-Chavez, E., Vicent, G.P., Elemento, O., Beato, M., and Reyes, J.C. (2015). The chromatin Remodeler CHD8 is required for activation of progesterone receptor-dependent enhancers. *PLoS Genet* 11, e1005174.
- Chen, H., Xue, Y., Huang, N., Yao, X., and Sun, Z. (2006). MeMo: a web tool for prediction of protein methylation modifications. *Nucleic Acids Res* 34, W249-253.
- Chen, J., Yao, Z.X., Chen, J.S., Gi, Y.J., Munoz, N.M., Kundra, S., Herlong, H.F., Jeong, Y.S., Goltsov, A., Ohshiro, K., *et al.* (2016). TGF-beta/beta2-spectrin/CTCF-regulated tumor suppression in human stem cell disorder Beckwith-Wiedemann syndrome. *J Clin Invest* 126, 527-542.
- Chen, S., Ma, J., Wu, F., Xiong, L.J., Ma, H., Xu, W., Lv, R., Li, X., Villen, J., Gygi, S.P., *et al.* (2012). The histone H3 Lys 27 demethylase JMJD3 regulates gene expression by impacting transcriptional elongation. *Genes Dev* 26, 1364-1375.
- Chenn, A., and Walsh, C.A. (2002). Regulation of cerebral cortical size by control of cell cycle exit in neural precursors. *Science* 297, 365-369.
- Ciccone, D.N., Su, H., Hevi, S., Gay, F., Lei, H., Bajko, J., Xu, G., Li, E., and Chen, T. (2009). KDM1B is a histone H3K4 demethylase required to establish maternal genomic imprints. *Nature* 461, 415-418.
- Clissold, P.M., and Ponting, C.P. (2001). JmjC: cupin metalloenzyme-like domains in jumonji, hairless and phospholipase A2beta. *Trends Biochem Sci* 26, 7-9.
- Conti, L., and Cattaneo, E. (2010). Neural stem cell systems: physiological players or in vitro entities? *Nat Rev Neurosci* 11, 176-187.
- Cotney, J., Muhle, R.A., Sanders, S.J., Liu, L., Willsey, A.J., Niu, W., Liu, W., Klei, L., Lei, J., Yin, J., *et al.* (2015). The autism-associated chromatin modifier CHD8 regulates other autism risk genes during human neurodevelopment. *Nat Commun* 6, 6404.
- Cremer, T., Cremer, C., Baumann, H., Luedtke, E.K., Sperling, K., Teuber, V., and Zorn, C. (1982). Rabl's model of the interphase chromosome arrangement tested in Chinese hamster cells by premature chromosome condensation and laser-UV-microbeam experiments. *Hum Genet* 60, 46-56.
- Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A., *et al.* (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proceedings of the National Academy of Sciences of the United States of America* 107, 21931-21936.
- Crick, F.H., Barnett, L., Brenner, S., and Watts-Tobin, R.J. (1961). General nature of the genetic code for proteins. *Nature* 192, 1227-1232.
- Cruz-Molina, S., Respuela, P., Tebartz, C., Kolovos, P., Nikolic, M., Fueyo, R., van Ijcken, W.F.J., Grosveld, F., Frommolt, P., Bazzi, H., *et al.* (2017). PRC2 Facilitates the Regulatory Topology Required for Poised Enhancer Function during Pluripotent Stem Cell Differentiation. *Cell Stem Cell* 20, 689-705 e689.
- Currie, D.S., Hu, J.S., Kolski-Andreaco, A., and Monuki, E.S. (2007). Culture of mouse neural stem cell precursors. *J Vis Exp*, 152.
- Dai, J.P., Lu, J.Y., Zhang, Y., and Shen, Y.F. (2010). Jmjd3 activates Mash1 gene in RA-induced neuronal differentiation of P19 cells. *J Cell Biochem* 110, 1457-1463.
- Das, C., Lucia, M.S., Hansen, K.C., and Tyler, J.K. (2009). CBP/p300-mediated acetylation of histone H3 on lysine 56. *Nature* 459, 113-117.

- David, C.J., and Massague, J. (2018). Contextual determinants of TGFbeta action in development, immunity and cancer. *Nat Rev Mol Cell Biol*.
- De Santa, F., Barozzi, I., Mietton, F., Ghisletti, S., Polletti, S., Tusi, B.K., Muller, H., Ragoussis, J., Wei, C.L., and Natoli, G. (2010). A large fraction of extragenic RNA pol II transcription sites overlap enhancers. *PLoS Biol* 8, e1000384.
- De Santa, F., Totaro, M.G., Prosperini, E., Notarbartolo, S., Testa, G., and Natoli, G. (2007). The histone H3 lysine-27 demethylase Jmjd3 links inflammation to inhibition of polycomb-mediated gene silencing. *Cell* 130, 1083-1094.
- de Wit, E., Vos, E.S., Holwerda, S.J., Valdes-Quezada, C., Versteegen, M.J., Teunissen, H., Splinter, E., Wijchers, P.J., Krijger, P.H., and de Laat, W. (2015). CTCF Binding Polarity Determines Chromatin Looping. *Mol Cell* 60, 676-684.
- Dickson, M.C., Martin, J.S., Cousins, F.M., Kulkarni, A.B., Karlsson, S., and Akhurst, R.J. (1995). Defective haematopoiesis and vasculogenesis in transforming growth factor-beta 1 knock out mice. *Development* 121, 1845-1854.
- Dixon, J.R., Gorkin, D.U., and Ren, B. (2016). Chromatin Domains: The Unit of Chromosome Organization. *Mol Cell* 62, 668-680.
- Dixon, J.R., Jung, I., Selvaraj, S., Shen, Y., Antosiewicz-Bourget, J.E., Lee, A.Y., Ye, Z., Kim, A., Rajagopal, N., Xie, W., *et al.* (2015). Chromatin architecture reorganization during stem cell differentiation. *Nature* 518, 331-336.
- Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 485, 376-380.
- Dorighi, K.M., Swigut, T., Henriques, T., Bhanu, N.V., Scruggs, B.S., Nady, N., Still, C.D., 2nd, Garcia, B.A., Adelman, K., and Wysocka, J. (2017). Mll3 and Mll4 Facilitate Enhancer RNA Synthesis and Transcription from Promoters Independently of H3K4 Monomethylation. *Mol Cell* 66, 568-576 e564.
- Durak, O., Gao, F., Kaeser-Woo, Y.J., Rueda, R., Martorell, A.J., Nott, A., Liu, C.Y., Watson, L.A., and Tsai, L.H. (2016). Chd8 mediates cortical neurogenesis via transcriptional regulation of cell cycle and Wnt signaling. *Nat Neurosci* 19, 1477-1488.
- Ea, V., Baudement, M.O., Lesne, A., and Forne, T. (2015). Contribution of Topological Domains and Loop Formation to 3D Chromatin Organization. *Genes (Basel)* 6, 734-750.
- Engelen, E., Akinci, U., Bryne, J.C., Hou, J., Gontan, C., Moen, M., Szumska, D., Kockx, C., van Ijcken, W., Dekkers, D.H., *et al.* (2011). Sox2 cooperates with Chd7 to regulate genes that are mutated in human syndromes. *Nat Genet* 43, 607-611.
- Estaras, C., Akizu, N., Garcia, A., Beltran, S., de la Cruz, X., and Martinez-Balbas, M.A. (2012). Genome-wide analysis reveals that Smad3 and JMJD3 HDM co-activate the neural developmental program. *Development* 139, 2681-2691.
- Estaras, C., Fueyo, R., Akizu, N., Beltran, S., and Martinez-Balbas, M.A. (2013). RNA polymerase II progression through H3K27me3-enriched gene bodies requires JMJD3 histone demethylase. *Mol Biol Cell* 24, 351-360.
- Fernandez, V., Llinares-Benadero, C., and Borrell, V. (2016). Cerebral cortex expansion and folding: what have we learned? *EMBO J* 35, 1021-1044.
- Ferrari, K.J., Scelfo, A., Jammula, S., Cuomo, A., Barozzi, I., Stutzer, A., Fischle, W., Bonaldi, T., and Pasini, D. (2014). Polycomb-dependent H3K27me1 and H3K27me2 regulate active transcription and enhancer fidelity. *Mol Cell* 53, 49-62.
- Filion, G.J., van Bemmel, J.G., Braunschweig, U., Talhout, W., Kind, J., Ward, L.D., Brugman, W., de Castro, I.J., Kerkhoven, R.M., Bussemaker, H.J., *et al.* (2010). Systematic protein location mapping reveals five principal chromatin types in Drosophila cells. *Cell* 143, 212-224.

Bibliography

- Fishilevich, S., Nudel, R., Rappaport, N., Hadar, R., Plaschkes, I., Iny Stein, T., Rosen, N., Kohn, A., Twik, M., Safran, M., *et al.* (2017). GeneHancer: genome-wide integration of enhancers and target genes in GeneCards. Database (Oxford) 2017.
- Flemming, W. (1880). Beiträge zur Kenntniss der Zelle und Ihrer Lebenserscheinungen. Archiv für mikroskopische Anatomie 18, 151-259.
- Frankel, N., Davis, G.K., Vargas, D., Wang, S., Payre, F., and Stern, D.L. (2010). Phenotypic robustness conferred by apparently redundant transcriptional enhancers. Nature 466, 490-493.
- Fraser, J., Williamson, I., Bickmore, W.A., and Dostie, J. (2015). An Overview of Genome Organization and How We Got There: from FISH to Hi-C. Microbiol Mol Biol Rev 79, 347-372.
- Fu, H., Hu, Z., Wen, J., Wang, K., and Liu, Y. (2009). TGF-beta promotes invasion and metastasis of gastric cancer cells by increasing fascin1 expression via ERK and JNK signal pathways. Acta Biochim Biophys Sin (Shanghai) 41, 648-656.
- Fueyo, R., Garcia, M.A., and Martinez-Balbas, M.A. (2015). Jumonji family histone demethylases in neural development. Cell Tissue Res 359, 87-98.
- Fueyo, R., Iacobucci, S., Pappa, S., Estaras, C., Lois, S., Vicioso-Mantis, M., Navarro, C., Cruz-Molina, S., Reyes, J.C., Rada-Iglesias, A., *et al.* (2018). Lineage specific transcription factors and epigenetic regulators mediate TGFbeta-dependent enhancer activation. Nucleic Acids Res 46, 3351-3365.
- Galderisi, U., Peluso, G., Di Bernardo, G., Calarco, A., D'Apolito, M., Petillo, O., Cipollaro, M., Fusco, F.R., and Melone, M.A. (2013). Efficient cultivation of neural stem cells with controlled delivery of FGF-2. Stem Cell Res 10, 85-94.
- Garcia-Campmany, L., and Marti, E. (2007). The TGFbeta intracellular effector Smad3 regulates neuronal differentiation and cell fate specification in the developing spinal cord. Development 134, 65-75.
- Gordon, K.J., and Blobel, G.C. (2008). Role of transforming growth factor-beta superfamily signaling pathways in human disease. Biochim Biophys Acta 1782, 197-228.
- Graham, F.L., Smiley, J., Russell, W.C., and Nairn, R. (1977). Characteristics of a human cell line transformed by DNA from human adenovirus type 5. J Gen Virol 36, 59-74.
- Haarhuis, J.H., Elbatsh, A.M., and Rowland, B.D. (2014). Cohesin and its regulation: on the logic of X-shaped chromosomes. Dev Cell 31, 7-18.
- Hah, N., Murakami, S., Nagari, A., Danko, C.G., and Kraus, W.L. (2013). Enhancer transcripts mark active estrogen receptor binding sites. Genome Res 23, 1210-1223.
- Hamburger, V., and Hamilton, H.L. (1951). A series of normal stages in the development of the chick embryo. 1951. Dev Dyn 195, 231-272.
- Hansen, A.S., Cattoglio, C., Darzacq, X., and Tjian, R. (2018). Recent evidence that TADs and chromatin loops are dynamic structures. Nucleus 9, 20-32.
- Hansen, A.S., Pustova, I., Cattoglio, C., Tjian, R., and Darzacq, X. (2017). CTCF and cohesin regulate chromatin loop stability with distinct dynamics. Elife 6.
- Harlen, K.M., and Churchman, L.S. (2017). The code and beyond: transcription regulation by the RNA polymerase II carboxy-terminal domain. Nat Rev Mol Cell Biol 18, 263-273.
- Harp, J.M., Hanson, B.L., Timm, D.E., and Bunick, G.J. (2000). Asymmetries in the nucleosome core particle at 2.5 A resolution. Acta Crystallogr D Biol Crystallogr 56, 1513-1534.
- Hawinkels, L.J., and Ten Dijke, P. (2011). Exploring anti-TGF-beta therapies in cancer and fibrosis. Growth Factors 29, 140-152.
- Heintzman, N.D., Hon, G.C., Hawkins, R.D., Kheradpour, P., Stark, A., Harp, L.F., Ye, Z., Lee, L.K., Stuart, R.K., Ching, C.W., *et al.* (2009). Histone modifications at human enhancers reflect global cell-type-specific gene expression. Nature 459, 108-112.

- Heintzman, N.D., Stuart, R.K., Hon, G., Fu, Y., Ching, C.W., Hawkins, R.D., Barrera, L.O., Van Calcar, S., Qu, C., Ching, K.A., *et al.* (2007). Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet* 39, 311-318.
- Hnisz, D., Abraham, B.J., Lee, T.I., Lau, A., Saint-Andre, V., Sigova, A.A., Hoke, H.A., and Young, R.A. (2013). Super-enhancers in the control of cell identity and disease. *Cell* 155, 934-947.
- Hodawadekar, S.C., and Marmorstein, R. (2007). Chemistry of acetyl transfer by histone modifying enzymes: structure, mechanism and implications for effector design. *Oncogene* 26, 5528-5540.
- Huang, J., Liu, X., Li, D., Shao, Z., Cao, H., Zhang, Y., Trompouki, E., Bowman, T.V., Zon, L.I., Yuan, G.C., *et al.* (2016). Dynamic Control of Enhancer Repertoires Drives Lineage and Stage-Specific Transcription during Hematopoiesis. *Dev Cell* 36, 9-23.
- Hufnagel, R.B., Le, T.T., Riesenberger, A.L., and Brown, N.L. (2010). Neurog2 controls the leading edge of neurogenesis in the mammalian retina. *Dev Biol* 340, 490-503.
- Imayoshi, I., Isomura, A., Harima, Y., Kawaguchi, K., Kori, H., Miyachi, H., Fujiwara, T., Ishidate, F., and Kageyama, R. (2013). Oscillatory control of factors determining multipotency and fate in mouse neural progenitors. *Science* 342, 1203-1208.
- Imayoshi, I., and Kageyama, R. (2014). bHLH factors in self-renewal, multipotency, and fate choice of neural progenitor cells. *Neuron* 82, 9-23.
- Ivanov, D., and Nasmyth, K. (2005). A topological interaction between cohesin rings and a circular minichromosome. *Cell* 122, 849-860.
- Iwafuchi-Doi, M., and Zaret, K.S. (2014). Pioneer transcription factors in cell reprogramming. *Genes Dev* 28, 2679-2692.
- Janknecht, R., Wells, N.J., and Hunter, T. (1998). TGF-beta-stimulated cooperation of smad proteins with the coactivators CBP/p300. *Genes Dev* 12, 2114-2119.
- Jepsen, K., Solum, D., Zhou, T., McEvelly, R.J., Kim, H.J., Glass, C.K., Hermanson, O., and Rosenfeld, M.G. (2007). SMRT-mediated repression of an H3K27 demethylase in progression from neural stem cell to neuron. *Nature* 450, 415-419.
- Jiang, W., Wang, J., and Zhang, Y. (2013). Histone H3K27me3 demethylases KDM6A and KDM6B modulate definitive endoderm differentiation from human ESCs by regulating WNT signaling pathway. *Cell research* 23, 122-130.
- Jin, F., Li, Y., Dixon, J.R., Selvaraj, S., Ye, Z., Lee, A.Y., Yen, C.A., Schmitt, A.D., Espinoza, C.A., and Ren, B. (2013). A high-resolution map of the three-dimensional chromatin interactome in human cells. *Nature* 503, 290-294.
- Jori, F.P., Galderisi, U., Napolitano, M.A., Cipollaro, M., Cascino, A., Giordano, A., and Melone, M.A. (2007). RB and RB2/P130 genes cooperate with extrinsic signals to promote differentiation of rat neural stem cells. *Mol Cell Neurosci* 34, 299-309.
- Kaikkonen, M.U., Spann, N.J., Heinz, S., Romanoski, C.E., Allison, K.A., Stender, J.D., Chun, H.B., Tough, D.F., Prinjha, R.K., Benner, C., *et al.* (2013). Remodeling of the enhancer landscape during macrophage activation is coupled to enhancer transcription. *Mol Cell* 51, 310-325.
- Karolchik, D., Hinrichs, A.S., Furey, T.S., Roskin, K.M., Sugnet, C.W., Haussler, D., and Kent, W.J. (2004). The UCSC Table Browser data retrieval tool. *Nucleic Acids Res* 32, D493-496.
- Kartikasari, A.E., Zhou, J.X., Kanji, M.S., Chan, D.N., Sinha, A., Grapin-Botton, A., Magnuson, M.A., Lowry, W.E., and Bhushan, A. (2013). The histone demethylase Jmjd3 sequentially associates with the transcription factors Tbx3 and Eomes to drive endoderm differentiation. *EMBO J* 32, 1393-1408.
- Katayama, Y., Nishiyama, M., Shoji, H., Ohkawa, Y., Kawamura, A., Sato, T., Suyama, M., Takumi, T., Miyakawa, T., and Nakayama, K.I. (2016). CHD8 haploinsufficiency results in autistic-like phenotypes in mice. *Nature* 537, 675-679.

Bibliography

- Kim, S.W., Yoon, S.J., Chuong, E., Oyolu, C., Wills, A.E., Gupta, R., and Baker, J. (2011). Chromatin and transcriptional signatures for Nodal signaling during endoderm formation in hESCs. *Dev Biol* 357, 492-504.
- Kim, T.K., Hemberg, M., Gray, J.M., Costa, A.M., Bear, D.M., Wu, J., Harmin, D.A., Laptewicz, M., Barbara-Haley, K., Kuersten, S., *et al.* (2010). Widespread transcription at neuronal activity-regulated enhancers. *Nature* 465, 182-187.
- Koch, C.M., Andrews, R.M., Flicek, P., Dillon, S.C., Karaoz, U., Clelland, G.K., Wilcox, S., Beare, D.M., Fowler, J.C., Couttet, P., *et al.* (2007). The landscape of histone modifications across 1% of the human genome in five human cell lines. *Genome Res* 17, 691-707.
- Kolovos, P., Knoch, T.A., Grosveld, F.G., Cook, P.R., and Papantonis, A. (2012). Enhancers and silencers: an integrated and simple model for their function. *Epigenetics Chromatin* 5, 1.
- Komada, M. (2012). Sonic hedgehog signaling coordinates the proliferation and differentiation of neural stem/progenitor cells by regulating cell cycle kinetics during development of the neocortex. *Congenit Anom (Kyoto)* 52, 72-77.
- Kooistra, S.M., and Helin, K. (2012). Molecular mechanisms and potential functions of histone demethylases. *Nat Rev Mol Cell Biol* 13, 297-311.
- Kornberg, R.D. (1974). Chromatin structure: a repeating unit of histones and DNA. *Science* 184, 868-871.
- Kouzarides, T. (2007). Chromatin modifications and their function. *Cell* 128, 693-705.
- Kowalczyk, M.S., Hughes, J.R., Garrick, D., Lynch, M.D., Sharpe, J.A., Sloane-Stanley, J.A., McGowan, S.J., De Gobbi, M., Hosseini, M., Vernimmen, D., *et al.* (2012). Intragenic enhancers act as alternative promoters. *Mol Cell* 45, 447-458.
- Kvon, E.Z. (2015). Using transgenic reporter assays to functionally characterize enhancers in animals. *Genomics* 106, 185-192.
- Kwak, H., and Lis, J.T. (2013). Control of transcriptional elongation. *Annu Rev Genet* 47, 483-508.
- Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680-685.
- Lam, M.T., Cho, H., Lesch, H.P., Gosselin, D., Heinz, S., Tanaka-Oishi, Y., Benner, C., Kaikkonen, M.U., Kim, A.S., Kosaka, M., *et al.* (2013). Rev-Erbs repress macrophage gene expression by inhibiting enhancer-directed transcription. *Nature* 498, 511-515.
- Lan, F., Bayliss, P.E., Rinn, J.L., Whetstone, J.R., Wang, J.K., Chen, S., Iwase, S., Alpatov, R., Issaeva, I., Canaani, E., *et al.* (2007). A histone H3 lysine 27 demethylase regulates animal posterior development. *Nature* 449, 689-694.
- Lan, F., and Shi, Y. (2009). Epigenetic regulation: methylation of histone and non-histone proteins. *Science in China Series C, Life sciences* 52, 311-322.
- Lanctot, C., Cheutin, T., Cremer, M., Cavalli, G., and Cremer, T. (2007). Dynamic genome architecture in the nuclear space: regulation of gene expression in three dimensions. *Nat Rev Genet* 8, 104-115.
- Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10, R25.
- Larisch, S., Yi, Y., Lotan, R., Kerner, H., Eimerl, S., Tony Parks, W., Gottfried, Y., Birkey Reffey, S., de Caestecker, M.P., Danielpour, D., *et al.* (2000). A novel mitochondrial septin-like protein, ARTS, mediates apoptosis dependent on its P-loop motif. *Nat Cell Biol* 2, 915-921.
- Larue, L., and Bellacosa, A. (2005). Epithelial-mesenchymal transition in development and cancer: role of phosphatidylinositol 3' kinase/AKT pathways. *Oncogene* 24, 7443-7454.
- Lawrence, M., Daujat, S., and Schneider, R. (2016). Lateral Thinking: How Histone Modifications Regulate Gene Expression. *Trends in genetics : TIG* 32, 42-56.

- Lee, D.Y., Hayes, J.J., Pruss, D., and Wolffe, A.P. (1993). A positive role for histone acetylation in transcription factor access to nucleosomal DNA. *Cell* 72, 73-84.
- Lee, K., Hsiung, C.C.-S., Huang, P., Raj, A., and Blobel, G.A. (2015). Dynamic enhancer–gene body contacts during transcription elongation. *Genes & Development* 29, 1992-1997.
- Lee, M.G., Villa, R., Trojer, P., Norman, J., Yan, K.P., Reinberg, D., Di Croce, L., and Shiekhattar, R. (2007). Demethylation of H3K27 regulates polycomb recruitment and H2A ubiquitination. *Science* 318, 447-450.
- Lenhard, B., Sandelin, A., and Carninci, P. (2012). Metazoan promoters: emerging characteristics and insights into transcriptional regulation. *Nature Reviews Genetics* 13, 233.
- Li, Q., Zou, J., Wang, M., Ding, X., Chepelev, I., Zhou, X., Zhao, W., Wei, G., Cui, J., Zhao, K., *et al.* (2014). Critical role of histone demethylase Jmjd3 in the regulation of CD4⁺ T-cell differentiation. *Nat Commun* 5, 5780.
- Li, W., Cogswell, C.A., and LoTurco, J.J. (1998). Neuronal Differentiation of Precursors in the Neocortical Ventricular Zone Is Triggered by BMP. *The Journal of Neuroscience* 18, 8853-8862.
- Li, W., Notani, D., Ma, Q., Tanasa, B., Nunez, E., Chen, A.Y., Merkurjev, D., Zhang, J., Ohgi, K., Song, X., *et al.* (2013). Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. *Nature* 498, 516-520.
- Li, W., Notani, D., and Rosenfeld, M.G. (2016). Enhancers as non-coding RNA transcription units: recent insights and future perspectives. *Nature Reviews Genetics* 17, 207.
- Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., *et al.* (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 326, 289-293.
- Lin, C., Garruss, A.S., Luo, Z., Guo, F., and Shilatifard, A. (2013). The RNA Pol II elongation factor EII3 marks enhancers in ES cells and primes future gene activation. *Cell* 152, 144-156.
- Liu, W., Ma, Q., Wong, K., Li, W., Ohgi, K., Zhang, J., Aggarwal, A., and Rosenfeld, M.G. (2013). Brd4 and JMJD6-associated anti-pause enhancers in regulation of transcriptional pause release. *Cell* 155, 1581-1595.
- Liu, X., Kraus, W.L., and Bai, X. (2015a). Ready, pause, go: regulation of RNA polymerase II pausing and release by cellular signaling pathways. *Trends Biochem Sci* 40, 516-525.
- Liu, Z., Lavis, L.D., and Betzig, E. (2015b). Imaging live-cell dynamics and structure at the single-molecule level. *Mol Cell* 58, 644-659.
- Long, H.K., Prescott, S.L., and Wysocka, J. (2016). Ever-Changing Landscapes: Transcriptional Enhancers in Development and Evolution. *Cell* 167, 1170-1187.
- Luger, K., Dechassa, M.L., and Tremethick, D.J. (2012). New insights into nucleosome and chromatin structure: an ordered state or a disordered affair? *Nature Reviews Molecular Cell Biology* 13, 436.
- Luse, D.S. (2014). The RNA polymerase II preinitiation complex. Through what pathway is the complex assembled? *Transcription* 5, e27050.
- Luyten, A., Zang, C., Liu, X.S., and Shivdasani, R.A. (2014). Active enhancers are delineated de novo during hematopoiesis, with limited lineage fidelity among specified primary blood cells. *Genes Dev* 28, 1827-1839.
- Macias, M.J., Martin-Malpartida, P., and Massague, J. (2015). Structural determinants of Smad function in TGF-beta signaling. *Trends Biochem Sci* 40, 296-308.
- Magnuson, B., Veloso, A., Kirkconnell, K.S., de Andrade Lima, L.C., Paulsen, M.T., Ljungman, E.A., Bedi, K., Prasad, J., Wilson, T.E., and Ljungman, M. (2015). Identifying transcription start sites and active enhancer elements using BruUV-seq. *Sci Rep* 5, 17978.
- Margueron, R., and Reinberg, D. (2011). The Polycomb complex PRC2 and its mark in life. *Nature* 469, 343-349.

Bibliography

- Marino-Ramirez, L., Kann, M.G., Shoemaker, B.A., and Landsman, D. (2005). Histone structure and nucleosome stability. *Expert review of proteomics* 2, 719-729.
- Massague, J. (1998). TGF-beta signal transduction. *Annu Rev Biochem* 67, 753-791.
- Massagué, J. (2012). TGFβ signalling in context. *Nature Reviews Molecular Cell Biology* 13, 616.
- Massague, J., Seoane, J., and Wotton, D. (2005). Smad transcription factors. *Genes Dev* 19, 2783-2810.
- Maston, G.A., Evans, S.K., and Green, M.R. (2006). Transcriptional Regulatory Elements in the Human Genome. *Annual Review of Genomics and Human Genetics* 7, 29-59.
- Matharu, N., and Ahituv, N. (2015). Minor Loops in Major Folds: Enhancer-Promoter Looping, Chromatin Restructuring, and Their Association with Transcriptional Regulation and Disease. *PLoS Genet* 11, e1005640.
- Matsui, T., Segall, J., Weil, P.A., and Roeder, R.G. (1980). Multiple factors required for accurate initiation of transcription by purified RNA polymerase II. *J Biol Chem* 255, 11992-11996.
- McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol* 28, 495-501.
- Meas, R., and Mao, P. (2015). Histone ubiquitylation and its roles in transcription and DNA damage response. *DNA Repair (Amst)* 36, 36-42.
- Metzger, E., Wissmann, M., Yin, N., Muller, J.M., Schneider, R., Peters, A.H., Gunther, T., Buettner, R., and Schule, R. (2005). LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. *Nature* 437, 436-439.
- Micucci, J.A., Sperry, E.D., and Martin, D.M. (2015). Chromodomain helicase DNA-binding proteins in stem cells and human developmental diseases. *Stem Cells Dev* 24, 917-926.
- Molyneaux, B.J., Arlotta, P., Menezes, J.R., and Macklis, J.D. (2007). Neuronal subtype specification in the cerebral cortex. *Nat Rev Neurosci* 8, 427-437.
- Morikawa, M., Koinuma, D., Miyazono, K., and Heldin, C.H. (2013). Genome-wide mechanisms of Smad binding. *Oncogene* 32, 1609-1615.
- Moses, H.L., and Serra, R. (1996). Regulation of differentiation by TGF-beta. *Curr Opin Genet Dev* 6, 581-586.
- Mullen, A.C., Orlando, D.A., Newman, J.J., Loven, J., Kumar, R.M., Bilodeau, S., Reddy, J., Guenther, M.G., DeKoter, R.P., and Young, R.A. (2011). Master transcription factors determine cell-type-specific responses to TGF-beta signaling. *Cell* 147, 565-576.
- Mutch, C.A., Funatsu, N., Monuki, E.S., and Chenn, A. (2009). Beta-catenin signaling levels in progenitors influence the laminar cell fates of projection neurons. *J Neurosci* 29, 13710-13719.
- Nagano, T., Lubling, Y., Varnai, C., Dudley, C., Leung, W., Baran, Y., Mendelson Cohen, N., Wingett, S., Fraser, P., and Tanay, A. (2017). Cell-cycle dynamics of chromosomal organization at single-cell resolution. *Nature* 547, 61-67.
- Nieto, M.A., Huang, R.Y., Jackson, R.A., and Thiery, J.P. (2016). EMT: 2016. *Cell* 166, 21-45.
- Nishiyama, M., Nakayama, K., Tsunematsu, R., Tsukiyama, T., Kikuchi, A., and Nakayama, K.I. (2004). Early embryonic death in mice lacking the beta-catenin-binding protein Duplin. *Mol Cell Biol* 24, 8386-8394.
- Nishiyama, M., Oshikawa, K., Tsukada, Y., Nakagawa, T., Iemura, S., Natsume, T., Fan, Y., Kikuchi, A., Skoultchi, A.I., and Nakayama, K.I. (2009). CHD8 suppresses p53-mediated apoptosis through histone H1 recruitment during early embryogenesis. *Nat Cell Biol* 11, 172-182.
- Nishiyama, M., Skoultchi, A.I., and Nakayama, K.I. (2012). Histone H1 recruitment by CHD8 is essential for suppression of the Wnt-beta-catenin signaling pathway. *Mol Cell Biol* 32, 501-512.
- Norega, M.A., Ovcharenko, I., Afzal, V., and Rubin, E.M. (2003). Scanning human gene deserts for long-range enhancers. *Science* 302, 413.

- Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N.L., Meisig, J., Sedat, J., *et al.* (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485**, 381-385.
- Ohlsson, R., Renkawitz, R., and Lobanenkov, V. (2001). CTCF is a uniquely versatile transcription regulator linked to epigenetics and disease. *Trends in genetics : TIG* **17**, 520-527.
- Ohtaka-Maruyama, C., and Okado, H. (2015). Molecular Pathways Underlying Projection Neuron Production and Migration during Cerebral Cortical Development. *Front Neurosci* **9**, 447.
- Ohtani, K., Zhao, C., Dobрева, G., Manavski, Y., Kluge, B., Braun, T., Rieger, M.A., Zeiher, A.M., and Dimmeler, S. (2013). Jmjd3 controls mesodermal and cardiovascular differentiation of embryonic stem cells. *Circ Res* **113**, 856-862.
- Ornitz, D.M., and Itoh, N. (2015). The Fibroblast Growth Factor signaling pathway. *Wiley Interdiscip Rev Dev Biol* **4**, 215-266.
- Ozair, M.Z., Kintner, C., and Brivanlou, A.H. (2013). Neural induction and early patterning in vertebrates. *Wiley Interdisciplinary Reviews: Developmental Biology* **2**, 479-498.
- Parada, L.A., McQueen, P.G., Munson, P.J., and Misteli, T. (2002). Conservation of relative chromosome positioning in normal and cancer cells. *Curr Biol* **12**, 1692-1697.
- Park, D.H., Hong, S.J., Salinas, R.D., Liu, S.J., Sun, S.W., Sgualdino, J., Testa, G., Matzuk, M.M., Iwamori, N., and Lim, D.A. (2014). Activation of neuronal gene expression by the JMJD3 demethylase is required for postnatal and adult brain neurogenesis. *Cell Rep* **8**, 1290-1299.
- Patthey, C., and Gunhaga, L. (2014). Signaling pathways regulating ectodermal cell fate choices. *Exp Cell Res* **321**, 11-16.
- Perlman, R., Schiemann, W.P., Brooks, M.W., Lodish, H.F., and Weinberg, R.A. (2001). TGF-beta-induced apoptosis is mediated by the adapter protein Daxx that facilitates JNK activation. *Nat Cell Biol* **3**, 708-714.
- Platt, R.J., Zhou, Y., Slaymaker, I.M., Shetty, A.S., Weisbach, N.R., Kim, J.A., Sharma, J., Desai, M., Sood, S., Kempton, H.R., *et al.* (2017). Chd8 Mutation Leads to Autistic-like Behaviors and Impaired Striatal Circuits. *Cell Rep* **19**, 335-350.
- Pohlers, D., Brenmoehl, J., Loffler, I., Muller, C.K., Leipner, C., Schultze-Mosgau, S., Stallmach, A., Kinne, R.W., and Wolf, G. (2009). TGF-beta and fibrosis in different organs - molecular pathway imprints. *Biochim Biophys Acta* **1792**, 746-756.
- Pollard, S.M., Conti, L., Sun, Y., Goffredo, D., and Smith, A. (2006). Adherent neural stem (NS) cells from fetal and adult forebrain. *Cereb Cortex* **16 Suppl 1**, i112-120.
- Pott, S., and Lieb, J.D. (2014). What are super-enhancers? *Nature Genetics* **47**, 8.
- Pradeepa, M.M., Grimes, G.R., Kumar, Y., Olley, G., Taylor, G.C., Schneider, R., and Bickmore, W.A. (2016). Histone H3 globular domain acetylation identifies a new class of enhancers. *Nat Genet* **48**, 681-686.
- Proudhon, C., Snetkova, V., Raviram, R., Lobry, C., Badri, S., Jiang, T., Hao, B., Trimarchi, T., Kluger, Y., Aifantis, I., *et al.* (2016). Active and Inactive Enhancers Cooperate to Exert Localized and Long-Range Control of Gene Regulation. *Cell Rep* **15**, 2159-2169.
- Qiu, X., Kumari, G., Gerasimova, T., Du, H., Labaran, L., Singh, A., De, S., Wood, W.H., 3rd, Becker, K.G., Zhou, W., *et al.* (2018). Sequential Enhancer Sequestration Dysregulates Recombination Center Formation at the IgH Locus. *Mol Cell* **70**, 21-33 e26.
- Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841-842.
- Rada-Iglesias, A., Bajpai, R., Prescott, S., Brugmann, S.A., Swigut, T., and Wysocka, J. (2012). Epigenomic annotation of enhancers predicts transcriptional regulators of human neural crest. *Cell Stem Cell* **11**, 633-648.

Bibliography

- Rada-Iglesias, A., Bajpai, R., Swigut, T., Bruggmann, S.A., Flynn, R.A., and Wysocka, J. (2011). A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* **470**, 279-283.
- Ramadoss, S., Chen, X., and Wang, C.Y. (2012). Histone demethylase KDM6B promotes epithelial-mesenchymal transition. *J Biol Chem* **287**, 44508-44517.
- Ramirez, F., Ryan, D.P., Gruning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dunder, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* **44**, W160-165.
- Ran, F.A., Hsu, P.D., Lin, C.Y., Gootenberg, J.S., Konermann, S., Trevino, A.E., Scott, D.A., Inoue, A., Matoba, S., Zhang, Y., *et al.* (2013). Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* **154**, 1380-1389.
- Rao, S.S., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., *et al.* (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665-1680.
- Rao, S.S.P., Huang, S.C., Glenn St Hilaire, B., Engreitz, J.M., Perez, E.M., Kieffer-Kwon, K.R., Sanborn, A.L., Johnstone, S.E., Bascom, G.D., Bochkov, I.D., *et al.* (2017). Cohesin Loss Eliminates All Loop Domains. *Cell* **171**, 305-320 e324.
- Raposo, A.A., Vasconcelos, F.F., Drechsel, D., Marie, C., Johnston, C., Dolle, D., Bithell, A., Gillotin, S., van den Berg, D.L., Ettwiller, L., *et al.* (2015). Ascl1 Coordinately Regulates Gene Expression and the Chromatin Landscape during Neurogenesis. *Cell Rep*.
- Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B.D., Sun, Z.W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C.P., Allis, C.D., *et al.* (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* **406**, 593-599.
- Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. *Nat Biotechnol* **29**, 24-26.
- Rodriguez-Paredes, M., Ceballos-Chavez, M., Esteller, M., Garcia-Dominguez, M., and Reyes, J.C. (2009). The chromatin remodeling factor CHD8 interacts with elongating RNA polymerase II and controls expression of the cyclin E2 gene. *Nucleic Acids Res* **37**, 2449-2460.
- Rosa, S., and Shaw, P. (2013). Insights into chromatin structure and dynamics in plants. *Biology (Basel)* **2**, 1378-1410.
- Ross, S.E., Greenberg, M.E., and Stiles, C.D. (2003). Basic helix-loop-helix factors in cortical development. *Neuron* **39**, 13-25.
- Sanborn, A.L., Rao, S.S., Huang, S.C., Durand, N.C., Huntley, M.H., Jewett, A.I., Bochkov, I.D., Chinnappan, D., Cutkosky, A., Li, J., *et al.* (2015). Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proceedings of the National Academy of Sciences of the United States of America* **112**, E6456-6465.
- Satoh, T., Takeuchi, O., Vandenbon, A., Yasuda, K., Tanaka, Y., Kumagai, Y., Miyake, T., Matsushita, K., Okazaki, T., Saitoh, T., *et al.* (2010). The Jmjd3-Irf4 axis regulates M2 macrophage polarization and host responses against helminth infection. *Nat Immunol* **11**, 936-944.
- Sawicka, A., and Seiser, C. (2014). Sensing core histone phosphorylation - a matter of perfect timing. *Biochim Biophys Acta* **1839**, 711-718.
- Schuettengruber, B., and Cavalli, G. (2009). Recruitment of polycomb group complexes and their role in the dynamic regulation of cell fate choice. *Development* **136**, 3531-3542.
- Schwindt, T.T., Motta, F.L., Gabriela, F.B., Cristina, G.M., Guimaraes, A.O., Calcagnotto, M.E., Pesquero, J.B., and Mello, L.E. (2009). Effects of FGF-2 and EGF removal on the differentiation of mouse neural precursor cells. *An Acad Bras Cienc* **81**, 443-452.
- Seaberg, R.M., and van der Kooy, D. (2003). Stem and progenitor cells: the premature desertion of rigorous definitions. *Trends Neurosci* **26**, 125-131.

- Seoane, J. (2006). Escaping from the TGFbeta anti-proliferative control. *Carcinogenesis* 27, 2148-2156.
- Sexton, T., Yaffe, E., Kenigsberg, E., Bantignies, F., Leblanc, B., Hoichman, M., Parrinello, H., Tanay, A., and Cavalli, G. (2012). Three-dimensional folding and functional organization principles of the Drosophila genome. *Cell* 148, 458-472.
- Shah, N.M., Groves, A.K., and Anderson, D.J. (1996). Alternative neural crest cell fates are instructively promoted by TGFbeta superfamily members. *Cell* 85, 331-343.
- Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J.R., Cole, P.A., Casero, R.A., and Shi, Y. (2004). Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* 119, 941-953.
- Shi, Y., and Massague, J. (2003). Mechanisms of TGF-beta signaling from cell membrane to the nucleus. *Cell* 113, 685-700.
- Shimojo, H., Ohtsuka, T., and Kageyama, R. (2008). Oscillations in notch signaling regulate maintenance of neural progenitors. *Neuron* 58, 52-64.
- Shull, M.M., Ormsby, I., Kier, A.B., Pawlowski, S., Diebold, R.J., Yin, M., Allen, R., Sidman, C., Proetzel, G., Calvin, D., *et al.* (1992). Targeted disruption of the mouse transforming growth factor-beta 1 gene results in multifocal inflammatory disease. *Nature* 359, 693-699.
- Siegel, P.M., and Massague, J. (2003). Cytostatic and apoptotic actions of TGF-beta in homeostasis and cancer. *Nat Rev Cancer* 3, 807-821.
- Siegenthaler, J.A., and Miller, M.W. (2005). Transforming growth factor beta 1 promotes cell cycle exit through the cyclin-dependent kinase inhibitor p21 in the developing cerebral cortex. *J Neurosci* 25, 8627-8636.
- Sims, R.J., 3rd, Mandal, S.S., and Reinberg, D. (2004). Recent highlights of RNA-polymerase-II-mediated transcription. *Curr Opin Cell Biol* 16, 263-271.
- Slattey, M., Zhou, T., Yang, L., Dantas Machado, A.C., Gordan, R., and Rohs, R. (2014). Absence of a simple code: how transcription factors read the genome. *Trends Biochem Sci* 39, 381-399.
- Song, L., Zhang, Z., Grasfeder, L.L., Boyle, A.P., Giresi, P.G., Lee, B.K., Sheffield, N.C., Graf, S., Huss, M., Keefe, D., *et al.* (2011). Open chromatin defined by DNaseI and FAIRE identifies regulatory elements that shape cell-type identity. *Genome Res* 21, 1757-1767.
- Soutourina, J. (2017). Transcription regulation by the Mediator complex. *Nature Reviews Molecular Cell Biology* 19, 262.
- Splinter, E., de Wit, E., van de Werken, H.J., Klous, P., and de Laat, W. (2012). Determining long-range chromatin interactions for selected genomic sites using 4C-seq technology: from fixation to computation. *Methods* 58, 221-230.
- Srinivasan, S., Dorigi, K.M., and Tamkun, J.W. (2008). Drosophila Kismet regulates histone H3 lysine 27 methylation and early elongation by RNA polymerase II. *PLoS Genet* 4, e1000217.
- Stadhouders, R., Kolovos, P., Brouwer, R., Zuin, J., van den Heuvel, A., Kockx, C., Palstra, R.J., Wendt, K.S., Grosveld, F., van Ijcken, W., *et al.* (2013). Multiplexed chromosome conformation capture sequencing for rapid genome-scale high-resolution detection of long-range chromatin interactions. *Nat Protoc* 8, 509-524.
- Sugathan, A., Biagioli, M., Golzio, C., Erdin, S., Blumenthal, I., Manavalan, P., Ragavendran, A., Brand, H., Lucente, D., Miles, J., *et al.* (2014). CHD8 regulates neurodevelopmental pathways associated with autism spectrum disorder in neural progenitors. *Proceedings of the National Academy of Sciences of the United States of America* 111, E4468-4477.
- Sun, D., Zhou, X., Yu, H.L., He, X.X., Guo, W.X., Xiong, W.C., and Zhu, X.J. (2017). Regulation of neural stem cell proliferation and differentiation by Kinesin family member 2a. *PLoS One* 12, e0179047.

Bibliography

- Sun, G., Reddy, M.A., Yuan, H., Lanting, L., Kato, M., and Natarajan, R. (2010). Epigenetic histone methylation modulates fibrotic gene expression. *J Am Soc Nephrol* 21, 2069-2080.
- Sun, J., Rockowitz, S., Xie, Q., Ashery-Padan, R., Zheng, D., and Cvekl, A. (2015). Identification of in vivo DNA-binding mechanisms of Pax6 and reconstruction of Pax6-dependent gene regulatory networks during forebrain and lens development. *Nucleic Acids Res* 43, 6827-6846.
- Sun, P.D., and Davies, D.R. (1995). The cystine-knot growth-factor superfamily. *Annu Rev Biophys Biomol Struct* 24, 269-291.
- Sydow, J.F., and Cramer, P. (2009). RNA polymerase fidelity and transcriptional proofreading. *Curr Opin Struct Biol* 19, 732-739.
- Tachibana, I., Imoto, M., Adjei, P.N., Gores, G.J., Subramaniam, M., Spelsberg, T.C., and Urrutia, R. (1997). Overexpression of the TGFbeta-regulated zinc finger encoding gene, TIEG, induces apoptosis in pancreatic epithelial cells. *J Clin Invest* 99, 2365-2374.
- Taverna, E., Gotz, M., and Huttner, W.B. (2014). The cell biology of neurogenesis: toward an understanding of the development and evolution of the neocortex. *Annu Rev Cell Dev Biol* 30, 465-502.
- Temple, S. (2001). The development of neural stem cells. *Nature* 414, 112-117.
- Theus, M.H., Ricard, J., and Liebl, D.J. (2012). Reproducible expansion and characterization of mouse neural stem/progenitor cells in adherent cultures derived from the adult subventricular zone. *Curr Protoc Stem Cell Biol* Chapter 2, Unit 2D 8.
- Thompson, B.A., Tremblay, V., Lin, G., and Bochar, D.A. (2008). CHD8 is an ATP-dependent chromatin remodeling factor that regulates beta-catenin target genes. *Mol Cell Biol* 28, 3894-3904.
- Thongjuea, S., Stadhouders, R., Grosveld, F.G., Soler, E., and Lenhard, B. (2013). r3Cseq: an R/Bioconductor package for the discovery of long-range genomic interactions from chromosome conformation capture and next-generation sequencing data. *Nucleic Acids Res* 41, e132.
- Tiberi, L., Vanderhaeghen, P., and van den Aemele, J. (2012). Cortical neurogenesis and morphogens: diversity of cues, sources and functions. *Curr Opin Cell Biol* 24, 269-276.
- Tiedje, C., Lubas, M., Tehrani, M., Menon, M.B., Ronkina, N., Rousseau, S., Cohen, P., Kotlyarov, A., and Gaestel, M. (2015). p38MAPK/MK2-mediated phosphorylation of RBM7 regulates the human nuclear exosome targeting complex. *RNA* 21, 262-278.
- Towbin, H., Staehelin, T., and Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proceedings of the National Academy of Sciences of the United States of America* 76, 4350-4354.
- Tsukada, Y., Fang, J., Erdjument-Bromage, H., Warren, M.E., Borchers, C.H., Tempst, P., and Zhang, Y. (2006). Histone demethylation by a family of JmjC domain-containing proteins. *Nature* 439, 811-816.
- Uhlen, M., Fagerberg, L., Hallstrom, B.M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, A., Kampf, C., Sjostedt, E., Asplund, A., *et al.* (2015). Proteomics. Tissue-based map of the human proteome. *Science* 347, 1260419.
- Vacik, T., and Raska, I. (2017). Alternative intronic promoters in development and disease. *Protoplasma* 254, 1201-1206.
- Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A., *et al.* (2001). The sequence of the human genome. *Science* 291, 1304-1351.
- Visel, A., Blow, M.J., Li, Z., Zhang, T., Akiyama, J.A., Holt, A., Plajzer-Frick, I., Shoukry, M., Wright, C., Chen, F., *et al.* (2009). ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* 457, 854-858.
- Visel, A., Minovitsky, S., Dubchak, I., and Pennacchio, L.A. (2007). VISTA Enhancer Browser--a database of tissue-specific human enhancers. *Nucleic Acids Res* 35, D88-92.

- Visel, A., Taher, L., Girgis, H., May, D., Golonzhka, O., Hoch, R.V., McKinsey, G.L., Pattabiraman, K., Silberberg, S.N., Blow, M.J., *et al.* (2013). A high-resolution enhancer atlas of the developing telencephalon. *Cell* 152, 895-908.
- Vissers, L.E., van Ravenswaaij, C.M., Admiraal, R., Hurst, J.A., de Vries, B.B., Janssen, I.M., van der Vliet, W.A., Huys, E.H., de Jong, P.J., Hamel, B.C., *et al.* (2004). Mutations in a new member of the chromodomain gene family cause CHARGE syndrome. *Nat Genet* 36, 955-957.
- Vogel, T., Ahrens, S., Buttner, N., and Kriegelstein, K. (2010). Transforming growth factor beta promotes neuronal cell fate of mouse cortical and hippocampal progenitors in vitro and in vivo: identification of Nedd9 as an essential signaling component. *Cereb Cortex* 20, 661-671.
- Wang, A.H., Juan, A.H., Ko, K.D., Tsai, P.F., Zare, H., Dell'Orso, S., and Sartorelli, V. (2017). The Elongation Factor Spt6 Maintains ESC Pluripotency by Controlling Super-Enhancers and Counteracting Polycomb Proteins. *Mol Cell* 68, 398-413 e396.
- Wapinski, O.L., Lee, Q.Y., Chen, A.C., Li, R., Corces, M.R., Ang, C.E., Treutlein, B., Xiang, C., Baubet, V., Suchy, F.P., *et al.* (2017). Rapid Chromatin Switch in the Direct Reprogramming of Fibroblasts to Neurons. *Cell Rep* 20, 3236-3247.
- Wapinski, O.L., Vierbuchen, T., Qu, K., Lee, Q.Y., Chanda, S., Fuentes, D.R., Giresi, P.G., Ng, Y.H., Marro, S., Neff, N.F., *et al.* (2013). Hierarchical mechanisms for direct reprogramming of fibroblasts to neurons. *Cell* 155, 621-635.
- Wasylyk, B., and Chambon, P. (1979). Transcription by eukaryotic RNA polymerases A and B of chromatin assembled in vitro. *European journal of biochemistry* 98, 317-327.
- Weiss, A., and Attisano, L. (2013). The TGFbeta superfamily signaling pathway. *Wiley Interdiscip Rev Dev Biol* 2, 47-63.
- Weth, O., and Renkawitz, R. (2011). CTCF function is modulated by neighboring DNA binding factors. *Biochem Cell Biol* 89, 459-468.
- Whyte, W.A., Orlando, D.A., Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H., Rahl, P.B., Lee, T.I., and Young, R.A. (2013). Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* 153, 307-319.
- Wilkinson, G., Dennis, D., and Schuurmans, C. (2013). Proneural genes in neocortical development. *Neuroscience* 253, 256-273.
- Williams, K., Christensen, J., Rappsilber, J., Nielsen, A.L., Johansen, J.V., and Helin, K. (2014). The histone lysine demethylase JMJD3/KDM6B is recruited to p53 bound promoters and enhancer elements in a p53 dependent manner. *PLoS One* 9, e96545.
- Williamson, M.P. (1994). The structure and function of proline-rich regions in proteins. *Biochem J* 297 (Pt 2), 249-260.
- Wilson, S.W., and Houart, C. (2004). Early steps in the development of the forebrain. *Dev Cell* 6, 167-181.
- Xi, Q., He, W., Zhang, X.H., Le, H.V., and Massague, J. (2008). Genome-wide impact of the BRG1 SWI/SNF chromatin remodeler on the transforming growth factor beta transcriptional program. *J Biol Chem* 283, 1146-1155.
- Yan, X., Liu, Z., and Chen, Y. (2009). Regulation of TGF-beta signaling by Smad7. *Acta Biochim Biophys Sin (Shanghai)* 41, 263-272.
- Yang, X.J., and Seto, E. (2007). HATs and HDACs: from structure, function and regulation to novel strategies for therapy and prevention. *Oncogene* 26, 5310-5318.
- Zabidi, M.A., and Stark, A. (2016). Regulatory Enhancer-Core-Promoter Communication via Transcription Factors and Cofactors. *Trends in genetics : TIG* 32, 801-814.
- Zaret, K.S., and Carroll, J.S. (2011). Pioneer transcription factors: establishing competence for gene expression. *Genes Dev* 25, 2227-2241.
- Zentner, G.E., and Scacheri, P.C. (2012). The chromatin fingerprint of gene enhancer elements. *J Biol Chem* 287, 30888-30896.

Bibliography

- Zerbino, D.R., Achuthan, P., Akanni, W., Amode, M.R., Barrell, D., Bhai, J., Billis, K., Cummins, C., Gall, A., Giron, C.G., *et al.* (2018). Ensembl 2018. *Nucleic Acids Res* 46, D754-D761.
- Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., *et al.* (2008). Model-based analysis of ChIP-Seq (MACS). *Genome Biol* 9, R137.
- Zinzen, R.P., Girardot, C., Gagneur, J., Braun, M., and Furlong, E.E. (2009). Combinatorial binding predicts spatio-temporal cis-regulatory activity. *Nature* 462, 65-70.

Addendums

1. Publications

In this addendum I would like to list the publications in which I have participated during this doctoral thesis. Due to copyright restrictions, I am not allowed to include the complete documents. Nonetheless, I include the research article corresponding to this thesis acknowledging to *Nucleic Acids Research Journal* its original publication.

1.1. Publications as a first author

- Fueyo, R., S. Iacobucci, S. Pappa, C. Estaras, S. Lois, M. Vicioso-Mantis, C. Navarro, S. Cruz-Molina, J. C. Reyes, A. Rada-Iglesias, X. de la Cruz and M. A. Martinez-Balbas (2018). "Lineage specific transcription factors and epigenetic regulators mediate TGFbeta-dependent enhancer activation." *Nucleic Acids Res* 46(7): 3351-3365.
- Asensio-Juan, E*, R. Fueyo*, S. Pappa, S. Iacobucci, C. Badosa, S. Lois, M. Balada, L. Bosch-Presegue, A. Vaquero, S. Gutierrez, C. Caelles, C. Gallego, X. de la Cruz and M. A. Martinez-Balsas (2017). "The histone demethylase PHF8 is a molecular safeguard of the IFNgamma response." *Nucleic Acids Res* 45(7): 3800-3811.
- Fueyo, R.* , M. A. Garcia* and M. A. Martinez-Balbás (2015). "Jumonji family histone demethylases in neural development." *Cell Tissue Res* 359(1): 87-98

1.2. Other publications

- Carbonell, A., R. Fueyo, A. Izquierdo-Bouldstridge, C. Moreta and A. Jordan (2018). "Epigenetic mechanisms in health and disease: BCEC 2017." *Epigenetics* 13(3): 331-341.
- Cruz-Molina, S., P. Respuela, C. Tebartz, P. Kolovos, M. Nikolic, R. Fueyo, W. F. J. van Ijcken, F. Grosveld, P. Frommolt, H. Bazzi and A. Rada-Iglesias (2017). "PRC2 Facilitates the Regulatory Topology Required for Poised Enhancer Function during Pluripotent Stem Cell Differentiation." *Cell Stem Cell* 20(5): 689-705 e689.
- Sanchez-Cid, L., M. Pons, J. J. Lozano, N. Rubio, M. Guerra-Rebollo, A. Soriano, L. Paris-Coderch, M. F. Segura, R. Fueyo, J. Arguimbau, E. Zodda, R. Bermudo, I. Alonso, X. Caparros, M. Cascante, A. Rafii, Y. Kang, M. Martinez-Balbás, S. J. Weiss, J. Blanco, M. Munoz, P. L. Fernandez and T. M. Thomson (2017). "MicroRNA-200, associated with metastatic breast cancer, promotes traits of mammary luminal progenitor cells." *Oncotarget* 8(48): 83384-83406.
- Akizu, N., M. A. Garcia, C. Estaras, R. Fueyo, C. Badosa, X. de la Cruz and M. A. Martinez-Balbás (2016). "EZH2 regulates neuroepithelium structure and neuroblast proliferation by repressing p21." *Open Biol* 6(4): 150227.

Addendums

- Estaras, C., R. Fueyo, N. Akizu, S. Beltran and M. A. Martinez-Balbás (2013). "RNA polymerase II progression through H3K27me3-enriched gene bodies requires JMJD3 histone demethylase." *Mol Biol Cell* 24(3): 351-360.

Lineage specific transcription factors and epigenetic regulators mediate TGF β -dependent enhancer activation

Raquel Fueyo¹, Simona Iacobucci¹, Stella Pappa¹, Conchi Estarás¹, Sergio Lois², Marta Vicioso-Mantis¹, Claudia Navarro¹, Sara Cruz-Molina³, José Carlos Reyes⁴, Álvaro Rada-Iglesias³, Xavier de la Cruz² and Marian A. Martínez-Balbás^{1,*}

¹Department of Molecular Genomics. Instituto de Biología Molecular de Barcelona (IBMB), Consejo Superior de Investigaciones Científicas (CSIC), Barcelona 08028, Spain, ²Vall d'Hebron Institute of Research (VHIR), Passeig de la Vall d'Hebron, 119; E-08035 Barcelona, Spain. Institut Català per la Recerca i Estudis Avançats (ICREA), Barcelona 08018, Spain, ³Center for Molecular Medicine Cologne (CMMC), University of Cologne, Robert-Koch-Strasse 21, 50931 Cologne, Germany and ⁴Centro Andaluz de Biología Molecular y Medicina Regenerativa-CABIMER, Consejo Superior de Investigaciones Científicas (CSIC)-Universidad de Sevilla-Universidad Pablo de Olavide, Av. Americo Vespucio 41092 Seville, Spain

Received August 05, 2017; Revised January 29, 2018; Editorial Decision January 31, 2018; Accepted February 01, 2018

ABSTRACT

During neurogenesis, dynamic developmental cues, transcription factors and histone modifying enzymes regulate the gene expression programs by modulating the activity of neural-specific enhancers. How transient developmental signals coordinate transcription factor recruitment to enhancers and to which extent chromatin modifiers contribute to enhancer activity is starting to be uncovered. Here, we take advantage of neural stem cells as a model to unravel the mechanisms underlying neural enhancer activation in response to the TGF β signaling. Genome-wide experiments demonstrate that the proneural factor ASCL1 assists SMAD3 in the binding to a subset of enhancers. Once located at the enhancers, SMAD3 recruits the histone demethylase JMJD3 and the remodeling factor CHD8, creating the appropriate chromatin landscape to allow enhancer transcription and posterior gene activation. Finally, to analyze the phenotypical traits owed to cis-regulatory regions, we use CRISPR–Cas9 technology to demonstrate that the TGF β -responsive *Neurog2* enhancer is essential for proper neuronal polarization.

INTRODUCTION

During central nervous system development, multipotent neuroepithelial precursor cells originate specialized neurons and glial cells (1,2). The identity of the cells generated along neurogenesis is determined by the transcriptional programs operating on each cell type. The different gene expression outputs are the result of the interplay between developmental cues and epigenetic factors that control the activity of specific neural promoters and enhancers. Enhancers are distal *cis*-regulatory elements essential to control gene expression programs in a spatial-temporal manner during development (3). To do that, they integrate extrinsic and intrinsic signaling cues by containing clusters of recognition motifs for either lineage-specific transcription factors (TFs) or effectors of developmental signaling pathways (4). Moreover, TFs are dependent on the recruitment of coactivator proteins in order to activate transcription in the chromatin context (5). Historically, enhancers have been difficult to investigate due to the lack of identification criteria; however, recent epigenomic approaches to identify enhancer sequences at a genome-wide level have facilitated their study. Chromatin and TF signatures have allowed not only the unbiased detection of enhancers but also the classification of enhancers into active, primed and poised (6–8).

Despite the massive identification of enhancers that has been carried out in neural progenitors (9–11), the field still lacks data on how signaling cascades govern transient enhancer activations that lead to neural-fate changes. To address this question we investigate the mechanisms by which neural enhancers become activated by transforming growth

*To whom correspondence should be addressed. Tel: +34 93 403 4961; Fax: +34 93 403 4979; Email: mmbmc@ibmb.csic.es
Present address: Conchi Estarás, Regulatory Biology Laboratory, The Salk Institute for Biological Studies, La Jolla, CA 92037-1099, USA.

factor β (TGF β) pathway. As vertebrate neurogenesis is an asynchronized series of events, the study of molecular mechanisms involved in enhancer activation is very difficult in the developing embryo. As an alternative, we have used neural stem cells (NSCs) extracted from cerebral cortices of mouse embryos from E12.5. These NSCs adherent cultures are a very reliable model to study neurogenesis in comparison to the heterogeneity of the embryo (12). In response to TGF β , NSCs lose multipotency and commit to the neuronal lineage (13–15), although TGF β alone is not sufficient to differentiate NSCs into mature neurons. Mechanistically, TGF β transduces signals from the plasma membrane to the nucleus by interacting with the serine/threonine kinase-type I and type II receptors. TGF β binding leads to phosphorylation and activation of the effectors SMAD2 and SMAD3, that next interact with SMAD4 to enter the nucleus and regulate gene expression (16,17). At the cellular level, TGF β controls growth, differentiation, migration and adhesion, in a cell context-dependent manner (18). Thus, the biological output of TGF β action depends on the subset of genes and/or enhancers that are regulated (19), and this relies on the particular combination of co-factors participating in each cellular context. Indeed, several chromatin modifier enzymes have been identified to be associated with activated SMAD proteins (histone acetyltransferases CBP/p300, P/CAF or the remodeling factor BRG1) (17,20,21). Particularly, the TGF β effectors interact with the lysine demethylase (KDM) JMJD3 in embryonic stem cells (ESCs) (22,23) and in neural progenitors to facilitate neuronal differentiation induction (15).

Here, we show that TGF β signaling pathway activates a specific set of SMAD-responsive enhancers involved in neuronal commitment. Using genome-wide data, we identify that SMAD3 binding to neural enhancers coincides with the proneural achaete-scute family bHLH transcription factor 1 (ASCL1). SMAD3 binding to enhancers is associated with an increase in eRNA transcription that in turns correlates with gene activation. We uncover that this process is dependent on the action of the SMAD3 cofactor JMJD3 and the previously unknown partner CHD8 (chromodomain helicase DNA-binding protein 8). Importantly, we unequivocally identify a TGF β responsive enhancer that drives Neurogenin 2 (*Neurog2*) gene expression and proper neuronal differentiation of NSCs.

MATERIALS AND METHODS

Cell culture and differentiation

Mouse NSCs were dissected from cerebral cortices of C57BL/6J mouse fetal brains (E12.5) and cultured in poly-D-lysine (5 μ g/ml, 2 h 37°C) and laminin (5 μ g/ml 37°C, 4 h 37°C) precoated dishes following the previous published procedures (24). NSCs were grown with a medium prepared by mixing equal parts of DMEM F12 (without Phenol Red, Gibco) and Neural Basal Media (Gibco) containing Penicillin/Streptomycin, Glutamax (1%), N2 and B27 supplements (Gibco), non essential aminoacids (0.1 mM), sodium pyruvate (1 mM), HEPES (5 mM), Heparin (0.2 mg/l), bovine serum albumin (0.8 mg/l) and β -mercaptoethanol (0.01 mM) as previously described

(15). Fresh recombinant human Epidermal Growth Factor (EGF) (R&D systems) and Fibroblast Growth Factor (FGF) (Invitrogen) to 20 and 10 ng/ml final concentrations respectively were added to the growing media. Medium, supplements, EGF and FGF form the so-called expansion medium. Under these conditions, NSCs maintain the ability to self-renew and to generate a wide range of differentiated neural cell types (24–26). For NSCs differentiation experiments Parental and Δ *Neurog2* enh cells were plated in 24-well plates pre-coated with poly-D-lysine (5 μ g/ml, 2 h 37°C) and laminin (5 μ g/ml 37°C, 4 h 37°C) at a seeding density of 0.1×10^6 cells per well in NSCs expansion medium. After 24 h, expansion medium was replaced by differentiating medium, consisting in the same components of the expansion medium but without EGF and FGF (27–29). Fresh differentiating medium was supplied every 2 days and after 3, 6 or 8 days, cells were fixed and stained for indirect immunofluorescence. Under these conditions, NSCs differentiate toward neurons, astrocytes and oligodendrocytes (30). TGF β (Millipore) was used at a final concentration of 5 ng/ml. Human HEK293T cells were cultured in DMEM supplemented with 10% of fetal bovine serum (Gibco) and 1% of Penicillin/Streptomycin (31).

Antibodies and reagents

TGF β was acquired from Millipore (GF111). Antibodies used were anti: H3K27me3 (Millipore, 07449), H3K4me1 (Millipore, 07436), H3K4me2 (Millipore, 07030), H3K27me2 (Cell signaling, 9728S), H3K27ac (Abcam, ab4729), H3K4me3 (Abcam, ab8580), SMAD3 (Abcam, ab28379), phospho-SMAD3 (Cell Signaling, mAb9520), SMAD2/3 (BD Bioscience, 610842), ASCL1 (BD Pharmingen, 556604), JMJD3 (raised in the laboratory using amino acids 798–1095), CHD8 was raised in Dr José Carlos Reyes laboratory (32), β -TUBULIN III (TUJ1, Covance, MMS-435P), GFAP (Dako, z0334), NESTIN (Abcam, ab5968), HuC/D (MP, A21271), DAPI (ThermoFisher, D1306), β -TUBULIN (Millipore, MAB3408), VINCULIN (Sigma, V9131), HA tag (Abcam, ab20084) and MYC tag (Abcam, ab9132).

Plasmids and recombinant proteins

Previously published specific lentiviral vectors were either purchased from Sigma or cloned in pLKO.1 puro vector using AgeI and EcoRI sites, brackets indicate target sequence: pLKO-random (CAACAAGATGAAGAGCACC), pLKO-mSMAD3 (CCTTACCCTATCAGAGAGTA), pLKO-mASCL1 (CCACGGTCTTTGCTTCTGTTT), pLKO-mJMJD3 (CCTCTGTTCTTGAGGGACAAA), and pLKO-mCHD8 (TGCCTGGAAGAAATTGGAG). pCIG-HA-ASCL1, pCIG, pCIG-FLAG-SMAD3-S/D and pCIG-FLAG-SMAD3-S/A were kindly provided by Dr. Elisa Martí (33). pCIG-MYC-JMJD3 was described in (34). Luciferase pGL3-promoter and renilla pRL-TK vector were purchased in Promega. *Ctgf*(-102), *Nrip3*(-3,5) and *Neurog2*(-6) enhancer regions were extracted by PCR from mouse genomic DNA and cloned into luciferase reporter pGL3-promoter by using MluI and BglII restriction sites. Primers sequences can be found in Supplementary

Table S1. Empty backbone for CRISPR–Cas9 constructs was obtained from Addgene (#42230).

Chick *in ovo* electroporation

Eggs from White-Leghorn chickens were incubated at 38.5°C and 70% humidity. Embryos were staged following Hamburger and Hamilton (HH) (35). Chick embryos were electroporated with purified plasmid DNA at 1 µg/µl in H₂O with 50 ng/ml of Fast Green. Plasmid DNA was injected into the lumen of HH11–HH12 neural tubes, electrodes were placed at both sides of the neural tube and embryos were electroporated by an IntracelDual Pulse (TSS-100) electroporator delivering five 50 ms square pulses of 20–25 V.

In vivo luciferase assay

Enhancer activation by the TGFβ-pathway was assayed in chicken neural tubes. Chick embryos were electroporated at HH11–HH12 with pCIG-SMAD3-S/D, pCIG-SMAD3-S/A or empty pCIG, together with the luciferase reporter constructs and renilla for ooelectroporation efficiency normalization. Embryos were harvested after 48 h incubation *in ovo* and Dual Luciferase Reporter Assay System (Promega) was utilized to lyse neural tubes and measure luciferase and renilla activities.

Lentiviral transduction

Lentiviral transduction was carried out as previously described (36). Extended protocol is provided in Supplementary Materials and Methods.

CoIP and ChIP assays

Coimmunoprecipitation (CoIP) experiments were performed as previously described (34). Chromatin immunoprecipitation (ChIP) assays were essentially performed as described (37,38) with modifications: 1 × 10⁶ NSCs untreated or treated with TGFβ (5 ng/ml, for the indicated times) were fixed with formaldehyde 1% 10 min. Fixation was stopped by addition of 0.125 M glycine diluted in H₂O. Cells were lysed in 1% SDS lysis buffer (1% SDS; 10 mM EDTA pH 8.0; 50 mM Tris–HCl pH 8.1). Sonication step was performed in a Bioruptor sonicator and shredded chromatin was used for each immunoprecipitation. ChIP DNA was analyzed by qPCR with SYBR Green (Roche) in a LightCycler 480 PCR system (Roche) using specific primers (see Supplementary Table S1). Percentage of input was used for the quantification of the immunoprecipitated material with respect to the total starting chromatin. See expanded protocol in Supplementary Materials and Methods.

RNA extraction and qPCR

TRIZOL reagent (Invitrogen) was used to extract RNA, following the manufacturer instructions. Reverse transcription was performed with 2 µg of RNA using High Capacity cDNA reverse transcription kit (Invitrogen) and qPCR was performed with SYBR Green (Roche) in a LightCycler

480 (Roche) using specific primer pairs (see Supplementary Table S1). Extended protocol is provided in Supplementary Materials and Methods.

Indirect immunofluorescence and cell counting

Cells were fixed for 20 min in 4% paraformaldehyde (diluted in phosphate buffer 0.1 M, pH 7.4) and permeabilized with PBS-Triton X-100 (0.1%) before blocking at room temperature for 1 h in 1% BSA (in PBS with 0.1% Triton X-100) before overnight incubation at 4°C with primary antibodies. Finally, cells were incubated for 2 h at room temperature with Alexa-conjugated secondary IgG antibodies (Jackson ImmunoResearch) and 0.1 ng/µl DAPI (ThermoFisher, D1306). Images were captured by Leica SP5 confocal microscope using LAS-AF software.

Number of neurites per cell and percentage of uni/bipolar or multipolar neurons were quantified by direct counting of 10 randomly selected fields considering as multipolar the neurons with more than three neurites. Data show mean of $n = 60$ cells. Measure of the average length of the longest neurite in Parental and Δ *Neurog2* enh neurons was performed on representative fields using LAS AF Leica Microsystems Version: 1.8.2 build 1465 software.

Western blot

Immunoblotting was performed using standard procedures and visualized by means of an ECL kit (Amersham).

Size-exclusion chromatography

Size exclusion chromatography was performed with whole NSCs extracts in a Superose-6 10/300 gel filtration column (GE Healthcare) on AKTA purifier system (GE Healthcare). Proteins were detected by Western Blot.

ChIP-seq data acquisition and analysis

ChIP-seq data were downloaded from Gene Expression Omnibus (<https://www.ncbi.nlm.nih.gov/geo/>) (Accessions used in this paper are specified in Supplementary Table S2). For all of the accessions excepting for ASCL1, SMAD3 and JMJD3, bed files were already available and were used for enhancer identification and Venn diagram construction. For enhancer identification we used -intersectbed command from BEDTools with a minimum of overlapping base pairs of 50 (39). ChIP-seq captions were obtained from IGV genome browser (40). Extended protocol is provided in Supplementary Materials and Methods.

CRISPR–Cas9

In order to delete *Neurog2*(-6) enhancer, primer pairs of gRNA (Supplementary Table S1) were designed flanking the mm9 coordinates chr3:127 326 051–127 334 232 using the online tool <http://crispr.mit.edu/>. Selected pair of primers has a score of 88 (left) and 87 (right) and specificity was assessed observing that the highest off-target score for each pair of primers was theoretically low (1.4 left,

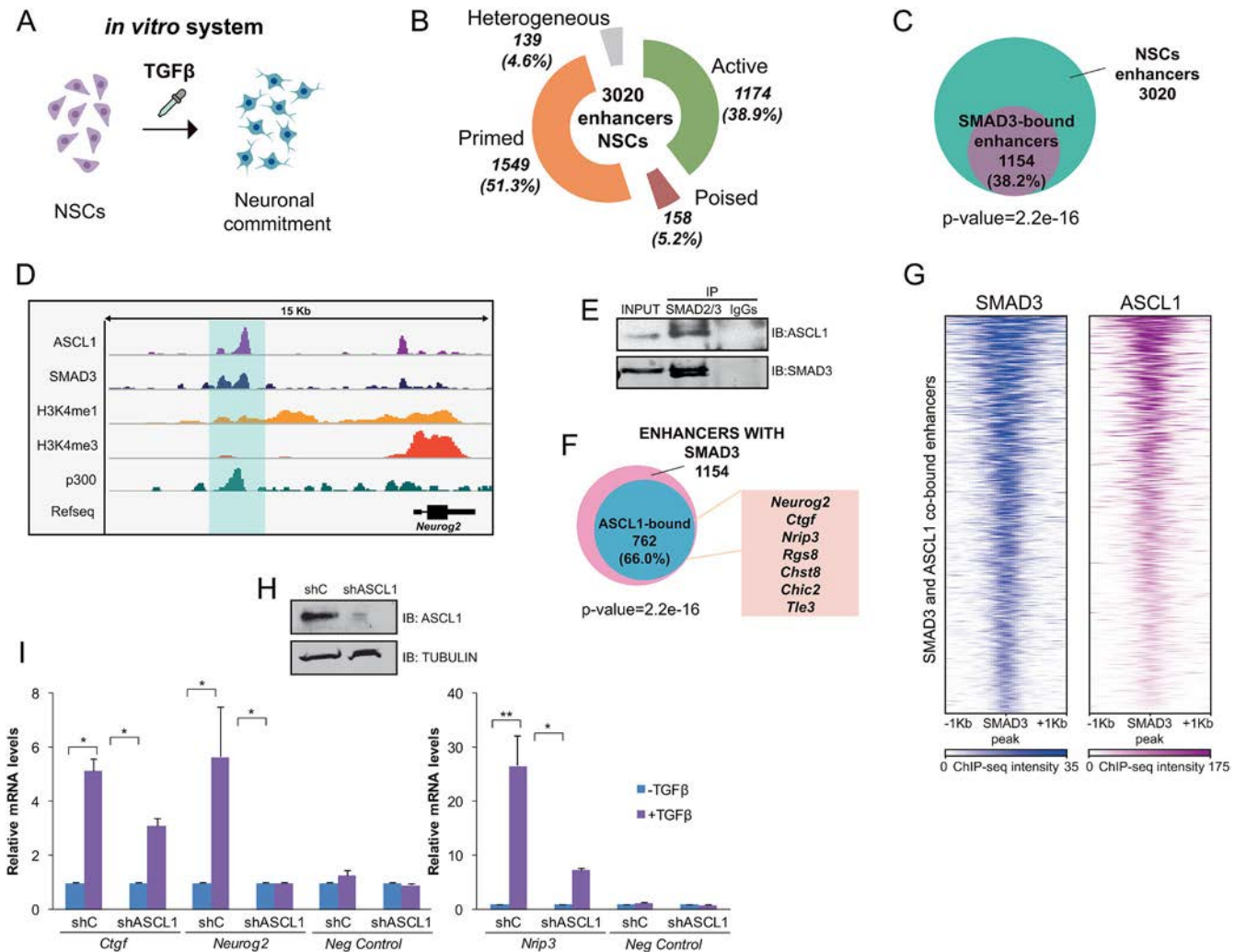


Figure 1. Epigenomic identification of neural enhancers. (A) Schematic view of the model used to study neural enhancer activation upon the TGF β differentiation signal. (B) Number and type of identified enhancers using ChIP-seq data for p300, H3K4me1, H3K4me3, H3K27ac and H3K27me3. The percentage related to the total number of enhancers is displayed in the graphic. (C) Venn diagram showing the number of neural enhancers bound by SMAD3 in NSCs treated with TGF β for 30 min. *P*-value is the result of an equal proportions test performed between SMAD3-bound enhancers and a random set. (D) IGV capture showing the chromatin landscape around the *Neurog2* gene. *Neurog2*(-6) enhancer is highlighted in blue. Tracks display ChIP-seq in NSCs treated for 30 min with TGF β (SMAD3) or untreated cells (ASCL1, p300, H3K4me1 and H3K4me3). (E) Endogenous SMAD2/3 was immunoprecipitated from NSCs and the presence of SMAD3 and ASCL1 in the immunopellet was determined by immunoblot with the antibodies indicated on the right part of the figure. Figure is representative of at least two biological independent experiments. (F) Venn diagram showing the number of neural SMAD3-bound enhancers upon TGF β treatment for 30 min that contain ASCL1 TF in untreated NSCs. The names of some genes putatively regulated by these enhancers are indicated. (G) Heatmap representation of SMAD3 (left) and ASCL1 (right) binding on the SMAD3 and ASCL1 co-occupied enhancers. Scales indicate ChIP-seq intensities. (H) NSCs were infected with lentivirus expressing shRNA control (shC) or shRNA specific for ASCL1 (shASCL1) cloned into pLKO vector. Forty eight hours later, total protein extracts were prepared and the ASCL1 and TUBULIN levels were determined by immunoblot. (I) shC and shASCL1 cells were treated with TGF β for 6 h (3 h for *Ctgf*) and mRNA levels of TGF β -responsive genes associated to the enhancers bound by SMAD3 and ASCL1 were analyzed by qPCR. Data were normalized to *Rps23* housekeeping gene and figure shows values relative to shC samples. *Ccne3* gene, a non TGF β -responsive gene was used as a control. Results are the mean of three biological independent experiments. Error bars indicate SD. **P* < 0.05; ***P* < 0.01 (Student's *t*-test).

1.3 right). gRNA were cloned in pX330-U6-Chimeric_BB-CBh-hSpCas9 vector using BbsI sites. Left and right-cutting plasmids were nucleofected in NSCs with an Amaxa Nucleofector (Lonza) following manufacturer instructions. After puromycin selection (2 μ g/ml) and detection analysis with conventional PCR, heterogeneous population carrying a majority of homozygotic deletions was used for experiments. Adequate CRISPR-Cas9 deletion was additionally

assessed by Sanger sequencing and SnapGene viewer was used to generate sequence pictures.

Statistical analysis

Quantitative data were expressed as mean and standard deviation (SD) (for immunofluorescence countings and RNA transcription experiments) and as mean and standard error of the mean (SEM) (for ChIPs). At least two or three biologically independent experiments were performed. The sig-

nificance of differences between groups was assessed using the Student's *t*-test (**P* < 0.05; ***P* < 0.01; ****P* < 0.001).

MIQE guidelines

This manuscript provides all the information recommended by MIQE Guidelines (Supplementary Dataset 1). Details on MIQE Guidelines for qPCR experiments are mainly provided in the RNA isolation and qPCR methods section and figure legends. MIQE check-list and primer list has been included as Supplementary Dataset 1 and Supplementary Table S1 respectively. RF is the lab member who performed all the qPCR assays.

RESULTS

Epigenetic identification of neural enhancers

To evaluate the functional relevance of enhancers during neural differentiation, we utilized NSCs from cortices of mouse embryos (E12.5) to create neural progenitor cells as a model (Figure 1A). Using previously published epigenomic data and following the well-established enhancer identification criteria (8) we analyzed the enhancer landscape in NSCs. For that purpose, we combined ChIP-seq of p300, H3K4me1, H3K4me3, H3K27ac and H3K27me3 generated in cortex progenitors (11) to classify the enhancers in active enhancers: presence of p300, H3K4me1 and H3K27ac; poised enhancers: binding of p300 and presence of H3K4me1 and H3K27me3 modifications; primed enhancers: marked with H3K4me1 and p300 and heterogeneous enhancers: presence of p300, H3K4me1 and both H3K27ac and H3K27me3. We identified 3020 putative enhancers in NSCs (defined as presence of p300, H3K4me1 and low levels of H3K4me3) (Figure 1B and D, and Supplementary Dataset 2). Among these, we found 1174 (38.9%) active enhancers, 158 (5.2%) poised enhancers, 1549 (51.3%) primed enhancers and 139 (4.6%) heterogeneous enhancers (Figure 1B). The lower number of poised enhancers in NSCs compared to ESCs is in agreement with previous reports indicating that the poised enhancers seem to be especially abundant in pluripotent cells (41). Nevertheless, in order to understand how relevant the poised enhancers are at this stage we analyzed how many of them become activated in terminally differentiated neurons. By comparing previously published datasets of active enhancers in neurons from E16.5 embryos (42), we found that 74 out of the 158 poised enhancers in progenitors (46.8%) became active after the stage-transition (Supplementary Figure S1A). Notably, some classical neuronal genes associated to these enhancers are included in this list (*Camk2n1*, *Kcnd3*, *Tle3*, *Amigo1*...). As H3K4me2 histone mark has also been associated to enhancers (43), we also cataloged the putative neural enhancers using this modification instead of H3K4me1 (together with p300 binding and low levels of H3K4me3). A similar set of data, although reduced in number and enriched in active enhancers was generated (Supplementary Figure S1B and Supplementary Dataset 2). Moreover, the number of mouse promoters co-localizing with the enhancers identified with H3K4me1/2 was very low (8 and 9, respectively), (Supplementary Figure S1C), validating the enhancer analysis.

Once identified the putative enhancers in NSCs, we sought to understand the mechanism underlying the dynamic activation of the enhancers in response to a developmental signal involved in neuronal fate. To do that, we studied the response of the enhancers to the well-characterized TGFβ pathway. This signaling cascade is already active under basal conditions to allow progenitor proliferation (14,15). Further TGFβ stimulation leads to the full gene activation that drives neuronal commitment *in vitro* and *in vivo* (14,15,44,45) (Figure 1A) (Supplementary Figure S1D and Supplementary Dataset 3). To decipher whether TGFβ pathway is directly implicated in NSCs enhancer activation, we analyzed the SMAD3 genomic distribution in NSCs upon TGFβ signaling with our previously published SMAD3 ChIP-seq (15). After combining the data, we identified 1154 (38.2%) putative NSCs enhancers that were bound by SMAD3 (*P*-value 2.2e-16, equal proportions test against a random sample) (Figure 1C), reinforcing the potential role of TGFβ-signaling in NSCs.

ASCL1 interacts with SMAD3 at NSCs enhancers

It has been described that SMAD genome-binding pattern shares many targets with cell-type specific TFs important for cell identity (46,47). We then, questioned about the neural specific factor/s that could cooperate with SMAD3 to target neural enhancers. Basic helix-loop-helix (bHLH) proteins have largely been demonstrated to be crucial players in chromatin regulation (48). Concretely, the bHLH protein ASCL1 has been shown to exert a critical role regulating neural gene expression by binding and opening the chromatin structure at enhancers (49). Thus, we tested whether ASCL1 and SMAD3 physically interact by CoIP assays. Figure 1E shows that endogenous ASCL1 and SMAD3 interact together. Next, we analyzed the colocalization of SMAD3 and ASCL1 at the genome wide level using previously published ChIP-seq data from NSCs (15,49). Doing that, we identified 762 (66%) SMAD3-bound enhancers that also contained ASCL1 (*P*-value 2.2e-16, equal proportion test against a random sample), (Figure 1F and G). Following these findings we decided to test whether TGFβ and ASCL1 had any functional association. To do that, ASCL1 protein levels were transiently depleted by transduction of lentivirus containing specific ASCL1 shRNA into the NSCs (Figure 1H and Supplementary Figure S1E). The transient removal of ASCL1 did not affect SMAD3 levels, proliferation rate or differentiation status of the NSCs (Supplementary Figure S1E, F and G). After viral transduction, the expression of some TGFβ-responsive genes associated to enhancers bound by ASCL1 and SMAD3 (Figure 1F) was tested by qPCR in the shASCL1 and control cell lines. We chose *Ctgf*, *Nrip3* and *Neurog2* genes that cover the spectrum of transcriptional levels in our previously published microarray data (15,50). *Ccne3* gene, which does not respond to TGFβ, was used as a negative control. Results in Figure 1I demonstrate that ASCL1 is essential to fully activate TGFβ-targets in NSCs. In ASCL1 depleted cells, the response to TGFβ of *Neurog2* was totally abolished and *Nrip3* and *Ctgf* induction was severely decreased. All things considered, these data point to the proneural factor ASCL1

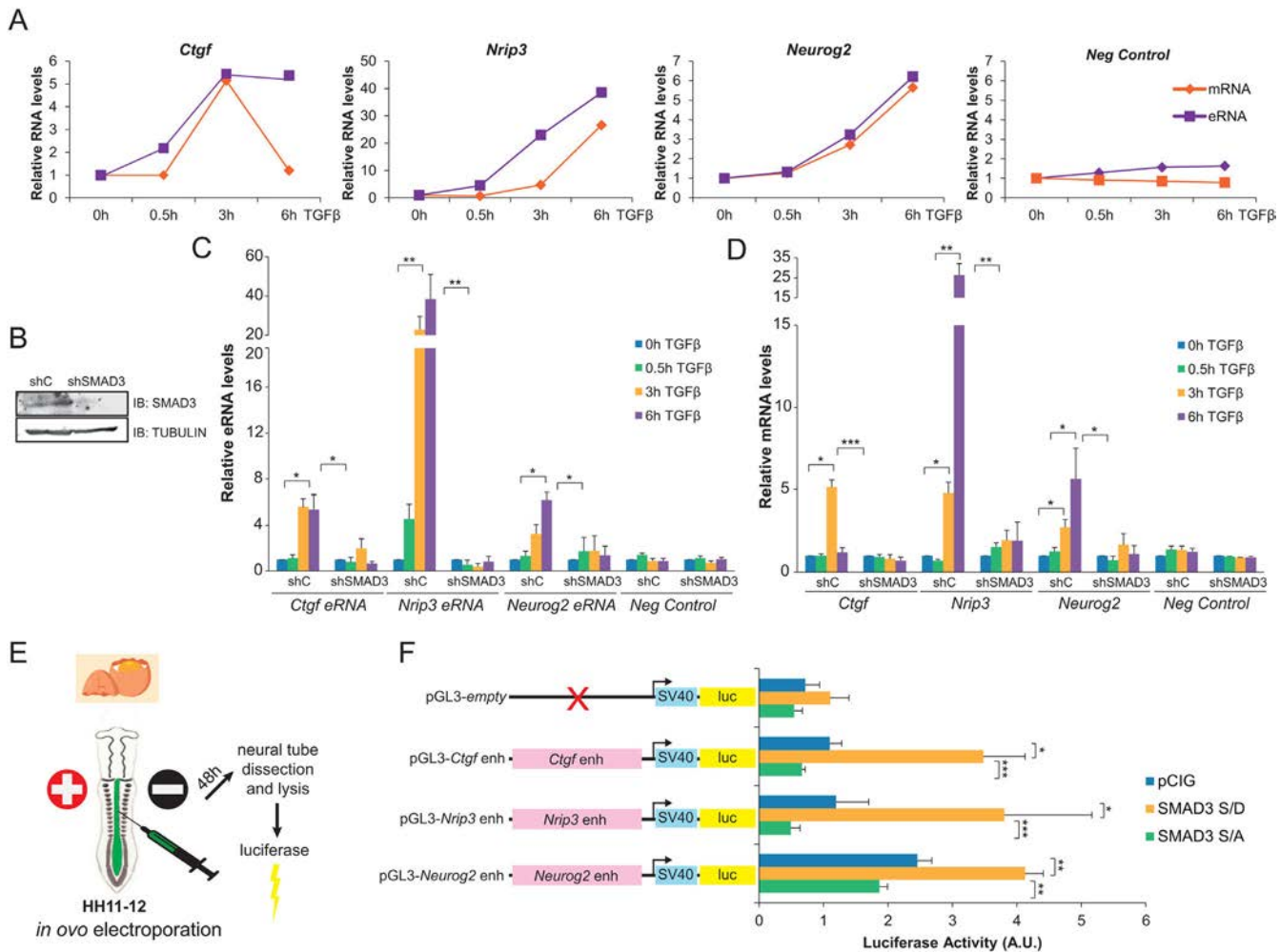


Figure 2. SMAD3-bound enhancers respond to TGF β . (A) NSCs cells were treated with TGF β . eRNA from the indicated enhancers was determined by qPCR and the mRNA of the associated genes was analyzed and represented together with the eRNA levels. Transcription values were normalized to the housekeeping gene *Rps23* and figure shows values relative to shC samples. Progesterone-responsive *Fabp4* gene and eRNA were used as negative controls. Results are the mean of three biological independent experiments. Errors bars represent SD. * $P < 0.05$; ** $P < 0.01$ (Student's *t*-test). (B) NSCs were infected with lentivirus expressing shRNA control (shC) or shRNA specific for SMAD3 (shSMAD3). Forty eight hours later, total protein extracts were prepared and the SMAD3 and TUBULIN levels were determined by immunoblot. (C and D) Control (shC) or SMAD3 depleted (shSMAD3) NSCs cells were treated for the indicated time with TGF β . Then, total RNA was prepared and the levels of the indicated eRNA (C) or mRNA (D) were determined by qPCR. Transcription values were normalized to the housekeeping gene *Rps23* and figure shows values relative to time 0h. Progesterone-responsive *Fabp4* eRNA or mRNA were used as negative controls. Results are the mean of three biological independent experiments. Errors bars represent SD. * $P < 0.05$; ** $P < 0.01$. (E) Schematic representation of the *in vivo* ooelectroporation and luciferase experiment. (F) HH11–12 embryos were electroporated *in ovo* with pGL3-empty or the pGL3-enhancer fusions together with pCIG (empty vector), SMAD3-S/D (pseudo-phosphorylated mutant) or SMAD3-S/A (mutant that cannot be phosphorylated). 48h-PE neural tubes were dissected, tissue was disaggregated and the luciferase activity was measured using the Promega dual kit. Data represent ratios between luciferase and renilla in arbitrary units. Values are the mean of three experiments from four to six embryos. Error bars indicate SD. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$ (Student's *t*-test).

as a new partner of SMAD3 in forebrain enhancer recognition.

Neural enhancers are dynamically activated by TGF β

Once identified the SMAD3-ASCL1 bound neural enhancers, we investigated whether they became fully active upon TGF β signaling. To this end, we evaluated the transcription from the enhancers, measuring the enhancer RNAs (eRNAs) which serve as readout of enhancer activation (51). We chose for posterior analysis the *Nrip3*(–3.5), *Ctgf*(–102) and *Neurog2*(–6) enhancers. These enhancers are associated to the previously analyzed genes (Figure 1I)

and represent the two major functional enhancer categories: active (*Nrip3* and *Ctgf*) and poised (*Neurog2*). In NSCs, eRNAs were detected as early as 30 min after TGF β stimulation (Figure 2A). Interestingly, we observed a strong correlation on both the magnitude of the transcription and the stability among the eRNAs and their corresponding mRNAs (Figure 2A). To further confirm that enhancer activation was dependent on TGF β pathway and thus, of SMAD3 binding, NSCs depleted of SMAD3 were generated (Figure 2B) and eRNA synthesis was tested. In concordance with the aforementioned results, eRNAs were hardly induced in shSMAD3 cells compared to the control cell line (Figure

2C). Similarly, mRNA of the associated genes was not produced upon TGF β addition in the shSMAD3 NSCs (Figure 2D).

We also analyzed the levels of H3K27ac and H3K4me2 chromatin marks by ChIP assays. Data in Supplementary Figure S2A and B shows that the analyzed enhancers display H3K27ac and H3K4me2. After TGF β addition, a time-dependent increase on the H3K27ac and H3K4me2 levels is observed at *Ctgf*(-102) and *Nrip3*(-3.5) enhancers but not at the intergenic region used as a negative control. No changes on H3K4me1 and H3K4me3 levels were detected (Supplementary Figure S2C and D), yet a clear increase on H3K4me3 was observed at the promoter of the analyzed genes (Supplementary Figure S2E) correlating with full activation in response to TGF β (Figure 2D). Altogether, these data demonstrate that TGF β signaling pathway stimulation results in enhancer activation and subsequent transcription of associated genes *in vitro*.

The above described findings support the idea of TGF β pathway activating relevant enhancers for neural development. Therefore, we tested whether SMAD3 regulates the analyzed enhancers in an *in vivo* model of neurogenesis, the chick embryo neural tube (Figure 2E). This model was chosen because it has been previously shown that activation of TGF β pathway by overexpression of a constitutively active pseudo-phosphorylated form of SMAD3 (SMAD3-S/D) promotes neuronal differentiation (15,33). With this goal, *Ctgf*(-102), *Nrip3*(-3.5) and *Neurog2*(-6) enhancers were cloned into the pGL3 promoter vector that contains the SV40 promoter fused to the luciferase reporter gene (Figure 2F). Then, chick embryos were electroporated *in ovo* (EP) at HH11-12 stage with the luciferase reporter constructs together with either the empty vector (pCIG) or SMAD3-S/D expressing plasmid (Figure 2E). Figure 2F shows that the expression of the luciferase downstream the enhancers was increased by co-EP of the SMAD3-S/D constitutively active mutant, while the expression of the luciferase with the promoter alone (pGL3-empty) was not affected. On the contrary, co-EP of a SMAD3 mutant that cannot be phosphorylated (SMAD3-S/A) and acts as a dominant negative form of SMAD3 (33), blocked the TGF β induction of the analyzed enhancers. The results demonstrate that the investigated enhancers are activated *in vivo* in response to TGF β signaling pathway.

JMJD3 is recruited to neural enhancers in a TGF β -dependent manner

Next, we sought to identify the molecular machinery that collaborates with ASCL1/SMAD3 for full enhancer activation after TGF β signaling. Previous work in our laboratory has demonstrated that JMJD3 interacts and cooperates with SMAD3 at promoters to induce the TGF β -neurogenic program *in vitro* and *in vivo* (15,50). Thus, we wondered whether JMJD3 also affects TGF β transcriptional response by contributing to enhancer activation. To address this hypothesis, we first identified the JMJD3 bound enhancers upon TGF β using our reported ChIP-seq data (15). Results in Figure 3A and C shows that JMJD3 was found in 30.1% of total forebrain enhancers. Interestingly, among these, 66.3% contained also

ASCL1 and SMAD3 (20% of total enhancers), showing a strong colocalization of the three proteins at enhancers (Figure 3A-C). SMAD3 and JMJD3 were found at the three types of enhancers (active, primed and poised). However, there is an increase in the percentage of active enhancers when the three partners, ASCL1/SMAD3/JMJD3, are occupying the regions (*P*-value $1.1e-15$, equal proportions test between the ASCL1/SMAD3/JMJD3-bound enhancers and the total enhancers), suggesting that ASCL1 and JMJD3 contribute to enhancer activation in response to TGF β (Figure 3D). Moreover, GREAT analysis (52) of the ASCL1/SMAD3/JMJD3-bound enhancers returned categories involved in neurogenesis (Figure 3E). The fact that JMJD3 co-occupies enhancers together with ASCL1 prompted us to test whether JMJD3 interacts with ASCL1 in NSCs. CoIP experiments showed a clear interaction between these proteins (Figure 3F), supporting the idea that JMJD3, SMAD3 and ASCL1 form a functional complex at neural enhancers.

JMJD3 contributes to full enhancer activation in response to TGF β

Once identified the neural enhancers bound by ASCL1, SMAD3 and JMJD3 we evaluated the contribution of this KDM to full neural enhancer activation upon TGF β signaling. To do that, we efficiently depleted JMJD3 from NSCs using lentivirus expressing specific shRNAs (Figure 3G) and the activity of *Ctgf*(-102), *Nrip3*(-3.5) and *Neurog2*(-6) enhancers in response to TGF β was analyzed. Figure 3H shows that the synthesis of eRNAs was profoundly diminished in shJMJD3 NSCs upon TGF β stimulation, correlating with the lack of induction of the associated genes (Figure 3I). Notably, transcriptional level of *Utx*, another member of the JMJD3 KDM family, was unchanged in the tested conditions (Supplementary Figure S3A). Within the chromatin context, TGF β treatment was unable to increase neither H3K27ac nor H3K4me2 levels in the *Ctgf*(-102) enhancer and H3K27ac in the *Nrip3*(-3.5) enhancer in JMJD3 depleted cells (Supplementary Figure S3B and C). These data strongly suggest that JMJD3 is required for full enhancer activation upon TGF β stimulation.

As JMJD3 is a KDM responsible for H3K27me_{2/3} demethylation (53,54), we wondered whether JMJD3-bound enhancers were demethylated during TGF β induced activation. To answer this question, we selected three H3K27me_{2/3}-marked poised enhancers from genes that are activated upon TGF β -stimulation (Supplementary Figure S4F) (15): *Neurog2*(-6), *Chic2*(-26), *Tle3*(-114) and *Ctgf*(-102) as an active enhancer devoid of H3K27me_{2/3} marks. Then, we determined the levels of H3K27me_{2/3} in control and JMJD3 depleted NSCs before and upon TGF β stimulation by ChIP-qPCR. Our results did not show any significant change on the H3K27me₃ levels associated to *Neurog2*(-6) enhancer. Only an increase of H3K27me₃ was observed in JMJD3 depleted cell line upon TGF β addition (Supplementary Figure S4A and B). In the case of *Chic2*(-26) and *Tle3*(-114) H3K27me₃ decreased upon stimulation; however, these changes were independent of JMJD3 (Supplementary Figure S4A and B). The active enhancer *Ctgf*(-102) did not suffered any methylation.

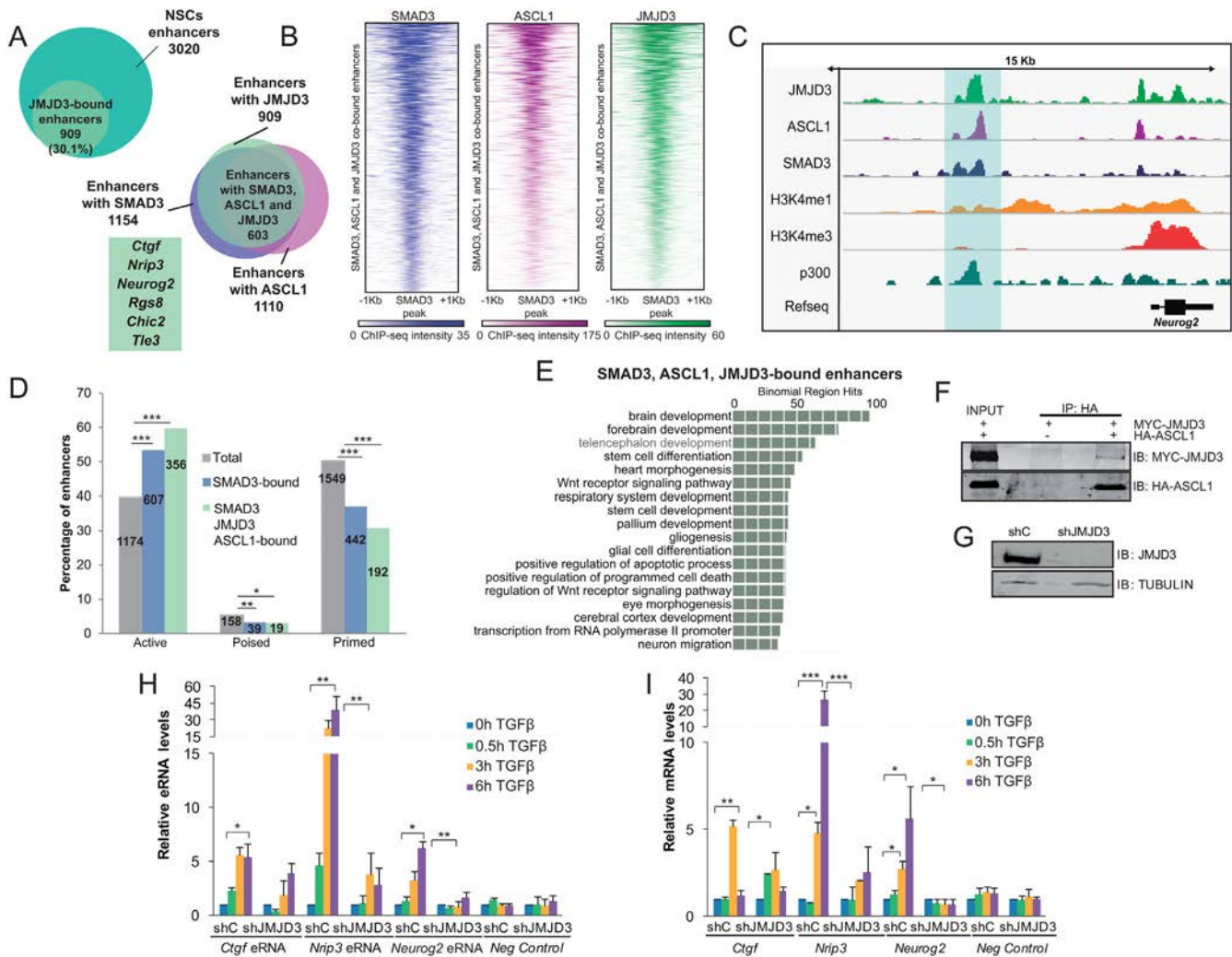


Figure 3. JMJD3 interacts with SMAD3 and ASCL1 at neural enhancers. (A) Venn diagrams showing the number and percentage of JMJD3-bound enhancers after 0.5 h of TGF β treatment in NSCs (upper panel) and number of enhancers with overlapping ASCL1 (before TGF β), SMAD3 and JMJD3 peaks (upon 30 min of TGF β addition) in NSCs. Genes associated to the identified enhancers are indicated. (B) Heatmap showing JMJD3 binding to neural enhancers co-bound by ASCL1, SMAD3 and JMJD3. Scales depict ChIP-seq intensities. (C) IGV capture showing the chromatin landscape around the *Neurog2* gene. Enhancer is highlighted in blue. Tracks display ChIP-seq in NSCs treated for 30 min with TGF β (SMAD3 and JMJD3) or untreated cells (ASCL1, p300, H3K4me1 and H3K4me3). (D) The percentage of active, poised and primed total enhancers, and those bound by SMAD3 or SMAD3/ASCL1/JMJD3 are depicted. Numbers inside the bars indicate absolute number of enhancers. An equal proportions test was performed and asterisks show *P*-values: **P* < 0.05; ***P* < 0.01; ****P* < 0.001. (E) GREAT analysis showing GO Biological Process of the enhancers co-occupied by ASCL1/SMAD3/JMJD3, analysis was performed using as a background the whole *Mus musculus* genome. (F) 293T cells were transfected with HA-ASCL1 and MYC-JMJD3 as indicated. ASCL1 was precipitated using the HA tag antibody and the presence of JMJD3 and ASCL1 in the immunopellet was determined by immunoblot with MYC and HA antibodies respectively. Figure is representative of at least three biological independent experiments. (G) NSCs were infected with lentivirus expressing shRNA control (shC) or shRNA specific for JMJD3 (shJMJD3). Forty eight hours later, total protein extracts were prepared and the JMJD3 and TUBULIN levels were determined by immunoblot. (H and I) shC or shJMJD3 NSCs were treated for the indicated times with TGF β . Levels of the indicated eRNA (J) or mRNA (K) were determined by qPCR. Transcription values were normalized to the housekeeping gene *Rps23* and figure shows values relative to time 0 h. Progesterone-responsive *Fabp4* eRNA or mRNA were used as negative controls. Results are the mean of three biological independent experiments. Errors bars represent SD. **P* < 0.05; ***P* < 0.01 (Student's *t*-test).

tion change in any of the conditions tested (Supplementary Figure S4A and B). Interestingly, the promoters associated to these poised enhancers were efficiently demethylated upon TGF β (Supplementary Figure S4C); accordingly, levels of the PRC2 complex subunit EZH2 decreased and the H3K4me3 levels resulted incremented (Supplementary Figure S4D and E). Correspondingly, at the mRNA level, treatment with TGF β boosted gene activation (Supplementary Figure S4F). These results suggest that H3K27 methy-

lated enhancers exhibit different behaviour upon TGF β -stimulation and that JMJD3 main role at enhancers is neither the demethylation of H3K27me3 nor the anti-H3K27me2 accumulation effect, suggesting a co-existence of catalytic-dependent and independent roles.

CHD8, a novel SMAD/JMJD3 partner

CHD8 belongs to a wide family of ATP-dependent remodelers that bind and open chromatin at enhancer re-

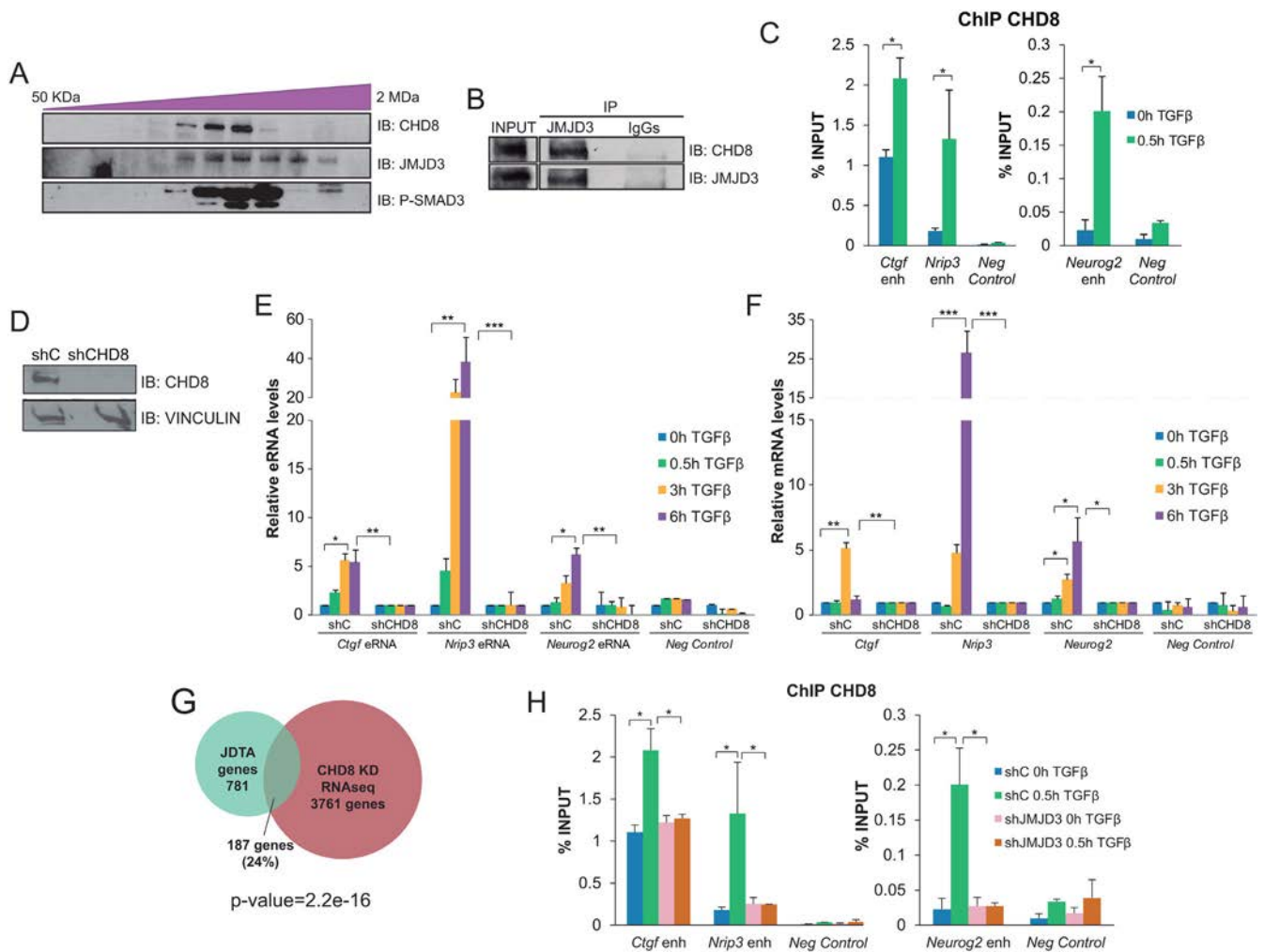


Figure 4. CHD8 facilitates enhancer activation. (A) Size-exclusion chromatography immunoblots depicting the co-elution of CHD8, JMJD3 and phospho-SMAD3. (B) Endogenous JMJD3 was precipitated from NSCs using JMJD3 antibody and the presence of CHD8 in the immunopellet was determined by immunoblot with the antibodies indicated on the right part of the figure. IgGs were used as negative control. Figure is representative of at least three biological independent experiments. (C) ChIP of CHD8 in NSCs treated for 0 and 0.5 h with TGFβ and analyzed by qPCR at the indicated enhancers. An intergenic region devoid of histone marks was used as negative control. Results are the mean of two biological independent experiments. Errors bars represent SEM. * $P < 0.05$ (Student's *t*-test). (D) Immunoblot showing the CHD8 levels in NSCs infected with lentivirus expressing shRNA control (shC) or shRNA specific for CHD8 (shCHD8). Total protein extracts were prepared from NSCs 48 h after infection and CHD8 and VINCULIN levels were determined by immunoblot. (E and F) shC or shCHD8 NSCs cells were treated for the indicated time with TGFβ. Then, total RNA was purified and the levels of the indicated eRNA (E) and mRNA (F) were determined by qPCR. Expression values were normalized to the housekeeping gene *Rps23*, and figure shows values relative to time 0h. Progesterone-responsive *Fabp4* eRNA or mRNA respectively were used as negative controls. Results are the mean of three biological independent experiments. Errors bars represent SD. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$ (Student's *t*-test). (G) Venn diagram showing the number and percentage of genes that are regulated by TGFβ pathway in a JMJD3 dependent manner (JDTA genes) in NSCs that are also regulated by CHD8 in the RNA-seq of cortical progenitors. (H) ChIP of CHD8 in control (shC) or JMJD3 depleted (shJMJD3) NSCs cells treated for 0 and 0.5 h with TGFβ and analyzed by qPCR at the indicated enhancers. An intergenic region was used as negative control. Results are the mean of two biological independent experiments. Errors bars represent SEM. * $P < 0.05$ (Student's *t*-test).

gions (55,56). Moreover, CHD8 is essential for early neurogenesis (57,58). Thus, we considered the possibility that CHD8 could contribute to full neural enhancer activation upon TGFβ signaling. To investigate this hypothesis, we first checked whether CHD8 could be physically contacting SMAD3 and JMJD3. For that purpose, we used size-exclusion chromatography on whole NSCs extracts and we observed that CHD8 co-eluted with the phosphorylated form of SMAD3. Consistently, JMJD3 co-eluted in the same fractions (Figure 4A). Additionally, by CoIP experiments we demonstrated that endogenous CHD8 in-

teracted with JMJD3 (Figure 4B). Then, we investigated whether CHD8 recruitment to the neural enhancer was dependent on TGFβ. Using ChIP-qPCR assays we observed that CHD8 was already bound to these active enhancers (due to the basal activity of TGFβ pathway) and TGFβ stimulation significantly increased at the analyzed enhancers upon TGFβ-stimulation (Figure 4C). Next, we tested whether CHD8 was required for the full enhancer activation induced by TGFβ. Depletion of CHD8 protein levels using specific shRNAs (Figure 4D) blocked the enhancer activity (Figure 4E), and the full activation of the

associated genes (Figure 4F). Additionally, genome-wide analysis showed that 24% of the genes regulated by TGF β that depend on JMJD3 (JDTA) (15) were also regulated by CHD8 (P -value $2.2e-16$, equal proportions test against a random sample of genes) (59). (Figure 4G). Then, we speculate that JMJD3 could be contributing to CHD8 high-affinity binding. To test this hypothesis we investigated the role of JMJD3 in CHD8 binding to neural enhancers upon TGF β . To do that, we analyzed the chromatin association of CHD8 in JMJD3 depleted cells. The results showed a clear decrease on CHD8 recruitment to the analyzed enhancers upon TGF β stimulation in shJMJD3 cells (Figure 4H). Altogether, these data demonstrate that CHD8 is a new SMAD3/JMJD3 partner at neural enhancers that is essential to full enhancer activation upon TGF β signaling.

The TGF β -responsive enhancer *Neurog2(-6)* is essential for proper neuronal differentiation

To demonstrate the physiological contribution of TGF β -responsive enhancers to neuronal commitment, we used CRISPR-Cas9 technology to delete an enhancer associated to a relevant gene for neurogenesis, *Neurog2(-6)* (Figure 5A and Supplementary Figure S5A). This regulatory region has been functionally tested *in vivo* by transgenic mouse assays that were positive for enhancer activity in forebrain and midbrain (60). Importantly, *Neurog2* is an essential bHLH TF that promotes neuronal differentiation, blocks glial differentiation and is essential for proper neuronal morphogenesis and migration (61–64). Thus, by using CRISPR-Cas9 technology we deleted *Neurog2(-6)* enhancer (Δ *Neurog2* enh) as demonstrated by multiplex PCR (Figure 5A) and genomic sequencing (Supplementary Figure S5B). Next, the parental and the Δ *Neurog2* enh NSCs were treated with TGF β and the expression of *Neurog2* was evaluated by RT-qPCR. Results show that *Neurog2* gene remained silenced in non-treated parental and Δ *Neurog2* enh NSCs (Figure 5B), suggesting that this poised enhancer is not responsible of maintaining *Neurog2* gene in an inactive state before differentiation induction, and, hence, does not act as a silencer. After TGF β addition, a clear *Neurog2* activation was observed in the parental line but not in Δ *Neurog2* enh cells (Figure 5B). Accordingly, synthesis of *Neurog2* eRNA was not observed in Δ *Neurog2* enh cell line (Figure 5B). These effects on *Neurog2* gene induction were not due to a lack of TGF β response, because other TGF β targets, such as *Nrip3*, were clearly upregulated in these cells upon TGF β treatment (Figure 5B). These data demonstrate that the *Neurog2(-6)* enhancer plays an essential and non-redundant role during the induction of its target gene upon TGF β . To further investigate the developmental relevance of this enhancer, we evaluated the consequences of lacking *Neurog2(-6)* enhancer during neuronal differentiation and morphogenesis *in vitro*. To do that, we tested the differentiation capacity of parental and Δ *Neurog2* enh NSCs by immunostaining using the neural progenitor marker NESTIN, (present in dividing cells and downregulated upon differentiation) and specific markers for neurons (TUBB3, known as TUJ1, and HuC/D) (Figure 5C and Supplementary Figure S5C). After three days in differentiating medium, both cell lines stopped proliferating (Figure 5Cii and vi and

Supplementary Figure S5D) and the number of cells that expressed neuronal markers were similar in parental and Δ *Neurog2* enh cells (Figure 5Cx and xiv and Supplementary Figure S5E). Nevertheless, Δ *Neurog2* enh cells showed lower TUJ1 intensity than parental NSCs (Figure 5C, e.g. x versus xiv). Interestingly, Δ *Neurog2* enh cells failed to properly differentiate into neurons (Figure 5C, e.g. xii versus xvi). After 6 days in differentiating medium, the number of neurites per cell was clearly reduced when compared with neurons from the parental cell line, going from 4–5 in control to 2–3 in Δ *Neurog2* enh TUJ1+ cells (Figure 5C e.g. xvii versus xviii; Supplementary Figure S5C e.g. i versus ii, quantified in Figure 5D). In addition, the neurite length was also reduced (Figure 5C e.g. xvii versus xviii; quantified in Supplementary Figure S5F) and interestingly, in Δ *Neurog2* enh pseudoneurons, a high number of cells presented only one, two (uni/bipolar neurons) or none neurites (Figure 5C e.g. xvii versus xviii; quantified in Figure 5E). These morphological alterations could be pointing to a misregulation of the transcription of cytoskeleton genes such as *Rnd2*, *Cdc42* or *Dcx*. For that reason, we tested their expression in parental and Δ *Neurog2* enh cell lines and we observed that the latter showed a clear misinduction of these genes upon differentiation (Figure 5F). These results are in agreement with previous reports indicating that NEUROG2 protein controls the expression of these cytoskeleton regulators (24,33,62).

These results demonstrate that the *Neurog2(-6)* enhancer is sufficient to modulate the acquisition of phenotypical traits during neuronal commitment. Altogether, these data show that TGF β -responsive enhancers play an essential role for the induction of a major regulator of neuronal differentiation of NSCs.

DISCUSSION

In this paper we provide a detailed molecular description of enhancer activation in response to the TGF β signaling pathway during neurogenesis. Our data uncover an unforeseen interplay between the TFs SMAD3 and ASCL1 and the chromatin modifiers JMJD3 and CHD8 that modulates neural enhancer activity towards the exit of the stem state upon TGF β signaling. We demonstrate that neural enhancers work as platforms to integrate developmental signals, cell-type specific TFs and epigenetic regulators to fine-tune neuronal commitment. A major caveat in the field is the understanding of the specificity of SMAD3 transcriptional response. In particular, the identity of the SMAD3 partners in different lineages is an intriguing issue. Recent studies have shown that master TFs, such as OCT4 in ESCs, MYOD1 in myotubes and PU.1 in pro-B cells select cell-type-specific response to TGF β signaling providing specificity to the TGF β response in these particular cellular contexts (46). Our work describes for the first time an interplay between TGF β -pathway and the bHLH ASCL1 (Figure 1). Our results demonstrate that the proneural TF, ASCL1, assists SMAD3 in chromatin binding at enhancers and their ulterior activation in NSCs. ASCL1, as well as other bHLH, is a main regulator of neurogenesis (65,66). Interestingly, ASCL1 works as a pioneer factor in a neurogenic context and its binding to enhancers results in disruptions in the

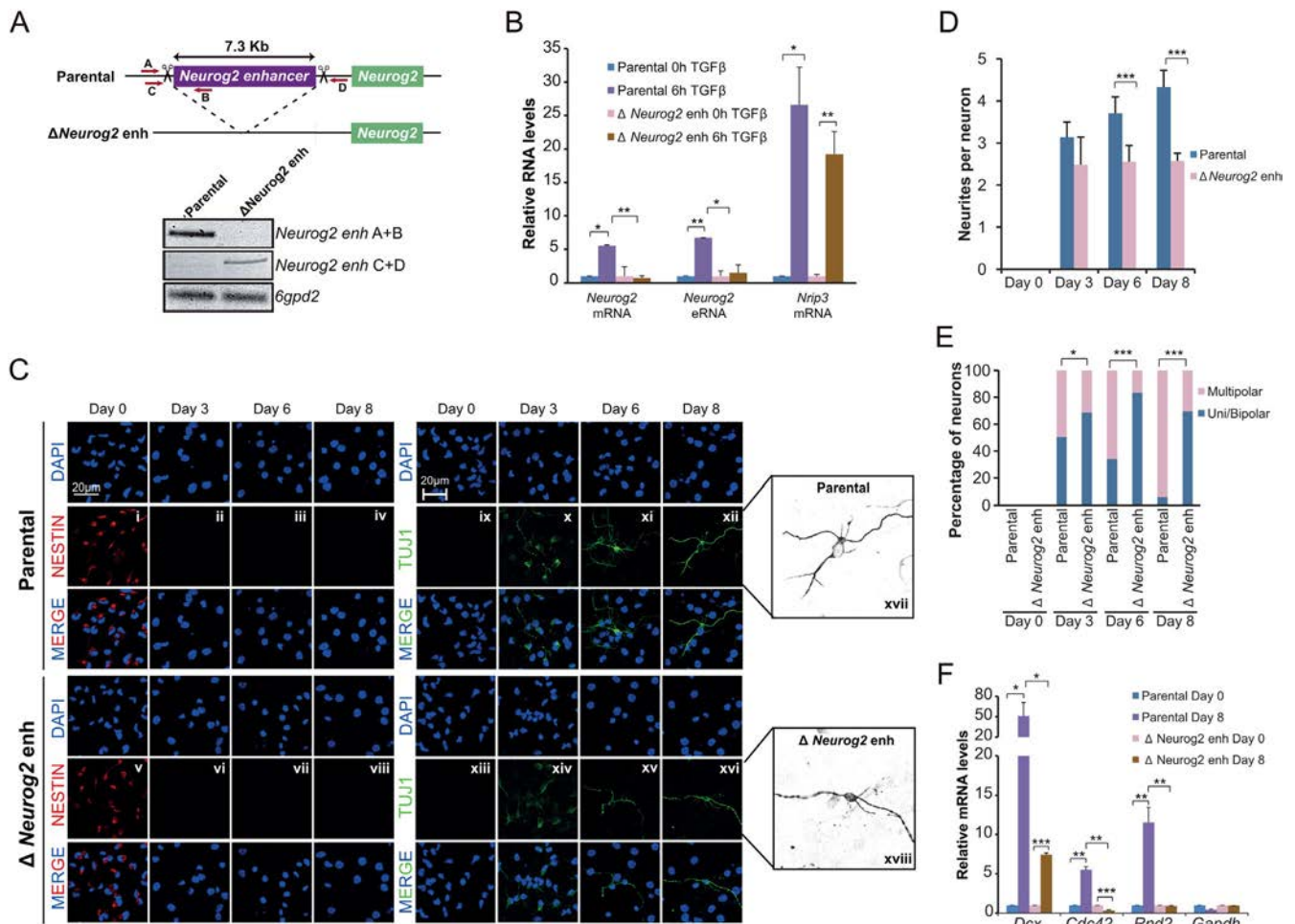


Figure 5. TGF β -responsive enhancers are essential for neuronal differentiation. (A) Schematic representation of the CRISPR/Cas9 experimental approach used to delete *Neurog2*(-6) enhancer in NSCs. Two gRNAs flanking *Neurog2*(-6) enhancer region were used to create the deletion. Red arrows represent primers to test the deletion. PCR using *Neurog2* deletion and *6gpd2* pairs of primers is shown in the top of the figure in parental and Δ *Neurog2* enh NSC lines. (B) Parental and Δ *Neurog2* enh cell lines were treated with TGF β for 6 h. Total RNA was prepared and the expression levels of *Neurog2* mRNA or eRNA were determined by qPCR. mRNA level of *Nrip3* was used as a TGF β response control. Transcription values were normalized to the housekeeping gene *Rps23*, and figure shows values relative to time 0 h. Errors bars represent SD. Results are representatives of two biological independent experiments. * $P < 0.05$; ** $P < 0.01$ (Student's *t*-test). (C–E) Parental and Δ *Neurog2* enh cell lines were cultured in differentiating medium. After 3, 6 or 8 days cells were fixed and stained with NESTIN, TUJ1 antibody and DAPI (C). Number of neurites per cell (D) and the percentage of uni/bipolar or multipolar pseudoneurons (E) were quantified by direct counting of 10 randomly selected fields. Data show mean of $n = 60$ cells. Error bars indicate SD. * $P < 0.05$; *** $P < 0.001$ (Student's *t*-test). (F) Parental and Δ *Neurog2* enh cell lines were maintained in differentiating medium for 8 days. Total RNA was purified and the expression levels of the indicated genes were determined by qPCR. Transcription values were normalized to the housekeeping gene *Ubc*, and figure shows values relative to Day 0 samples. Errors bars represent SD. Results are representatives of two biological independent experiments. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$ (Student's *t*-test).

chromatin landscape, allowing the posterior binding of signaling factors and cofactors (49). The functional cooperation between ASCL1 and SMAD3 provides specificity to TGF β response in neural context and opens new avenues to understand the functional interplay between intrinsic factors and extrinsic signals during development.

In addition to the factors that help SMAD3 recruitment, TGF β plasticity is dependent on the coactivator proteins, mainly chromatin acting factors, which regulate transcription in the genomic context. Although a large number of SMAD3 cofactors have been previously described, how they provide specificity to TGF β response is still unknown. Our studies extend the list of SMAD3 cofactors, by show-

ing that JMJD3 and CHD8 are essential to activate TGF β -responsive enhancers in NSCs.

JMJD3 is an essential cofactor during neural fate establishment (15,21). Our lab has previously demonstrated that JMJD3 interacts with SMAD3 at promoters in NSCs (15,50,67). In this study, we went further on that cooperation demonstrating that JMJD3 is essential to fully activate TGF β -responsive enhancers (Figure 3). The molecular link between our previously published SMAD3/JMJD3 interaction at promoters and the one now described at enhancers remains unclear. However, we hypothesize that the recruitment of TGF β cofactors is dependent on the three-dimensional structure of the chromatin and that the chromatin topology might constrain or facilitate the TF

and cofactor interplay and the potential enhancer-promoter contacts. Another critical player for neural enhancer activation in response to TGF β is CHD8 (Figure 4). This protein is an ATP-dependent chromatin remodeler factor *in vitro* and *in vivo*. In concordance, CHD8 binds enhancers and facilitated their remodeling in response to progesterone (55). Markedly, CHD8 plays an essential role in neurogenesis (57,58). Recently, CHD8 has called much attention due to its relevance in autism spectrum disorder (ASD). Functional analysis demonstrate that CHD8 regulates many ASD risk genes involved in neurodevelopment, synaptic function and WNT and p53 signaling pathways (58,59,68–72). Our data suggest that in addition to WNT and p53 pathways, TGF β signaling might be also orchestrating CHD8 targeting during neurogenesis. This new discovered partnership could be related to ASD or other developmental or neurodegenerative disorders in which TGF β is involved. Although more experiments are required to fully understand the interaction between TGF β and CHD8, the discovery of the functional axis SMAD3/JMJD3/CHD8 provides insights into how histone modifier enzymes and chromatin remodelers are coordinated to fully activate SMAD3 targeted neural enhancers in a temporal specific manner during neural development.

Using CRISPR–Cas9 genetic deletions, we show that *Neurog2*(–6) poised enhancer is necessary for the induction of its target gene upon TGF β signaling activation (Figure 5). Interestingly, *Neurog2*(–6) enhancer deletion did not result in the activation of the gene in NSCs, supporting the idea that the poised enhancers have not repressive activity (73) as it has been previously proposed (74–76). Major developmental and cell-identity genes are frequently regulated by multiple and sometimes redundant enhancers (77,78). However, our results clearly demonstrate that the deletion of a single poised enhancer totally blocks the expression of *Neurog2* after TGF β and compromise normal neuronal differentiation. This extends previous observations indicating that certain enhancers can control gene expression in a non-redundant manner in different cellular contexts (73,79,80).

The contribution of H3K27me3 demethylation by JMJD3 to enhancer activation is an intriguing question. Our results indicate that H3K27me2/3 levels are not altered in the analyzed enhancers during activation in a JMJD3-dependent manner. Studying the data, two mechanisms might be envisioned to explain the role of JMJD3 and H3K27me3 at the analyzed poised enhancers (Supplementary Figure S4A). In *Chic2*(–26) and *Tle3*(–114) enhancers, the H3K27me3 levels decreased upon TGF β in a JMJD3-independent manner. At these enhancers, JMJD3 function might not be related to H3K27me3, and could be linked to demethylation of other locus like the promoters (Supplementary Figure S4C) or associated with RNA-Polymerase II release at transcriptional starting sites (50,81). In the case of *Neurog2*(–6), the increase observed in H3K27me3 upon TGF β in shJMJD3 cells might be related to the ability of the PRC2 complex to bind this enhancer. It has been recently shown that PRC2 might be working as an activator of neural poised enhancers, by facilitating loop formation (73). Thus, it is possible that in the absence of JMJD3, EZH2 could be targeting more efficiently the enhancer increasing the levels of

H3K27me3 in response to TGF β signal. In addition to the poised enhancers, JMJD3 is actively recruited to some non-H3K27-methylated enhancers such as *Ctgf*(–102). This data suggest that in addition to demethylation, other JMJD3 catalytic-independent functions might be involved in TGF β -responsive enhancer activation as it has been previously proposed at promoters and gene bodies in different cellular contexts (50,82–85). These results also open the possibility that other essential factors different than histone H3 might be targeted by JMJD3 KDM activity upon TGF β to facilitate transcription activation. To fully clarify these issues more work will be required.

In summary, our results highlight enhancers as TF-binding platforms where different modifying enzymes coordinate their activities to induce faithful gene activation. This study uncovers the molecular mechanism responsible for full enhancer activation in response to TGF β signaling in a neural stem cell context. This involves the action of JMJD3 and CHD8 cofactors, which, by remodeling enhancers, previously pre-marked by ASCL1, activate the neuronal commitment program. Due to the broad range of TGF β functions in areas of cancer and neurodegenerative disorders, this work paves the way for investigating the ASCL1/SMAD/JMJD3/CHD8 contribution to transcriptional regulation in other cellular contexts and helps to move forward our understanding of the myriad of crosstalk between epigenetics and developmental programs.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Dr Mariona Arbonès and Dr Sonia Najas for technical help dissecting NSCs. We also thank Dr Elisa Martí for reagents. This work has been performed within the framework of the doctoral program in Biochemistry and Molecular Biology of the Universitat Autònoma de Barcelona.

FUNDING

Spanish Ministry of Education and Science [BFU2015-69248-P, BFU20012-34261 to M.M.B.]; Fundació La Marató de TV3 [090210 to M.M.B.]; Jérôme Lejeune Foundation; FPU fellowship and a travelling fellowship from the Company of Biologists (to R.F.). Funding for open access charge: Spanish Ministry of Education and Science [BFU2015-69248].

Conflict of interest statement. None declared.

REFERENCES

1. Temple, S. (2001) The development of neural stem cells. *Nature*, **414**, 112–117.
2. Tiberi, L., Vanderhaeghen, P. and van den Aemele, J. (2012) Cortical neurogenesis and morphogens: diversity of cues, sources and functions. *Curr. Opin. Cell Biol.*, **24**, 269–276.
3. Schaffner, W. (2015) Enhancers, enhancers - from their discovery to today's universe of transcription enhancers. *Biol. Chem.*, **396**, 311–327.

4. Buecker, C. and Wysocka, J. (2012) Enhancers as information integration hubs in development: lessons from genomics. *Trends Genet.*, **28**, 276–284.
5. Weake, V.M. and Workman, J.L. (2010) Inducible gene expression: diverse regulatory mechanisms. *Nat. Rev. Genet.*, **11**, 426–437.
6. Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A. *et al.* (2010) Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 21931–21936.
7. Zentner, G.E., Tesar, P.J. and Scacheri, P.C. (2011) Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions. *Genome Res.*, **21**, 1273–1283.
8. Rada-Iglesias, A., Bajpai, R., Swigut, T., Brugmann, S.A., Flynn, R.A. and Wysocka, J. (2011) A unique chromatin signature uncovers early developmental enhancers in humans. *Nature*, **470**, 279–283.
9. Visel, A., Blow, M.J., Li, Z., Zhang, T., Akiyama, J.A., Holt, A., Plajzer-Frick, I., Shoukry, M., Wright, C., Chen, F. *et al.* (2009) ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature*, **457**, 854–858.
10. Visel, A., Taher, L., Girgis, H., May, D., Golonzhka, O., Hoch, R.V., McKinsey, G.L., Pattabiraman, K., Silberberg, S.N., Blow, M.J. *et al.* (2013) A high-resolution enhancer atlas of the developing telencephalon. *Cell*, **152**, 895–908.
11. Sun, J., Rockowitz, S., Xie, Q., Ashery-Padan, R., Zheng, D. and Cvekl, A. (2015) Identification of *in vivo* DNA-binding mechanisms of Pax6 and reconstruction of Pax6-dependent gene regulatory networks during forebrain and lens development. *Nucleic Acids Res.*, **43**, 6827–6846.
12. Conti, L. and Cattaneo, E. (2010) Neural stem cell systems: physiological players or *in vitro* entities? *Nat. Rev. Neurosci.*, **11**, 176–187.
13. Roussa, E., Wiehle, M., Dunker, N., Becker-Katins, S., Oehlke, O. and Kriegstein, K. (2006) Transforming growth factor beta is required for differentiation of mouse mesencephalic progenitors into dopaminergic neurons *in vitro* and *in vivo*: ectopic induction in dorsal mesencephalon. *Stem Cells*, **24**, 2120–2129.
14. Vogel, T., Ahrens, S., Buttner, N. and Kriegstein, K. (2010) Transforming growth factor beta promotes neuronal cell fate of mouse cortical and hippocampal progenitors *in vitro* and *in vivo*: identification of Nedd9 as an essential signaling component. *Cereb. Cortex*, **20**, 661–671.
15. Estaras, C., Akizu, N., Garcia, A., Beltran, S., de la Cruz, X. and Martinez-Balbas, M.A. (2012) Genome-wide analysis reveals that Smad3 and JMJD3 HDM co-activate the neural developmental program. *Development*, **139**, 2681–2691.
16. Shi, Y. and Massague, J. (2003) Mechanisms of TGF-beta signaling from cell membrane to the nucleus. *Cell*, **113**, 685–700.
17. Feng, X.H. and Derynck, R. (2005) Specificity and versatility in tgf-beta signaling through Smads. *Annu. Rev. Cell Dev. Biol.*, **21**, 659–693.
18. Yang, L. and Moses, H.L. (2008) Transforming growth factor beta: tumor suppressor or promoter? Are host immune cells the answer? *Cancer Res.*, **68**, 9107–9111.
19. Massague, J. (2000) How cells read TGF-beta signals. *Nat. Rev. Mol. Cell Biol.*, **1**, 169–178.
20. Massague, J., Seoane, J. and Wotton, D. (2005) Smad transcription factors. *Genes Dev.*, **19**, 2783–2810.
21. Burgold, T., Spreafico, F., De Santa, F., Totaro, M.G., Prosperini, E., Natoli, G. and Testa, G. (2008) The histone H3 lysine 27-specific demethylase Jmjd3 is required for neural commitment. *PLoS One*, **3**, e3034.
22. Dahle, O., Kumar, A. and Kuehn, M.R. (2010) Nodal signaling recruits the histone demethylase Jmjd3 to counteract polycomb-mediated repression at target genes. *Sci. Signal.*, **3**, ra48.
23. Kim, S.W., Yoon, S.J., Chuong, E., Oyolu, C., Wills, A.E., Gupta, R. and Baker, J. (2011) Chromatin and transcriptional signatures for Nodal signaling during endoderm formation in hESCs. *Dev. Biol.*, **357**, 492–504.
24. Curre, D.S., Hu, J.S., Kolski-Andreaco, A. and Monuki, E.S. (2007) Culture of mouse neural stem cell precursors. *J. Vis. Exp.*, **152**, doi:10.3791/152.
25. Pollard, S.M., Conti, L., Sun, Y., Goffredo, D. and Smith, A. (2006) Adherent neural stem (NS) cells from fetal and adult forebrain. *Cereb. Cortex*, **16**(Suppl. 1), ii12–ii20.
26. Theus, M.H., Ricard, J. and Liebl, D.J. (2012) Reproducible expansion and characterization of mouse neural stem/progenitor cells in adherent cultures derived from the adult subventricular zone. *Curr. Protoc. Stem Cell Biol.*, doi:10.1002/9780470151808.sc02d08s20.
27. Galderisi, U., Peluso, G., Di Bernardo, G., Calarco, A., D'Apollito, M., Petillo, O., Cipollaro, M., Fusco, F.R. and Melone, M.A. (2013) Efficient cultivation of neural stem cells with controlled delivery of FGF-2. *Stem Cell Res.*, **10**, 85–94.
28. Schwindt, T.T., Motta, F.L., Gabriela, F.B., Cristina, G.M., Guimaraes, A.O., Calcagnotto, M.E., Pesquero, J.B. and Mello, L.E. (2009) Effects of FGF-2 and EGF removal on the differentiation of mouse neural precursor cells. *An. Acad. Bras. Cienc.*, **81**, 443–452.
29. Sun, D., Zhou, X., Yu, H.L., He, X.X., Guo, W.X., Xiong, W.C. and Zhu, X.J. (2017) Regulation of neural stem cell proliferation and differentiation by Kinesin family member 2a. *PLoS One*, **12**, e0179047.
30. Jori, F.P., Galderisi, U., Napolitano, M.A., Cipollaro, M., Cascino, A., Giordano, A. and Melone, M.A. (2007) RB and RB2/P130 genes cooperate with extrinsic signals to promote differentiation of rat neural stem cells. *Mol. Cell. Neurosci.*, **34**, 299–309.
31. Blanco-Garcia, N., Asensio-Juan, E., de la Cruz, X. and Martinez-Balbas, M.A. (2009) Autoacetylation regulates P/CAF nuclear localization. *J. Biol. Chem.*, **284**, 1343–1352.
32. Rodriguez-Paredes, M., Ceballos-Chavez, M., Esteller, M., Garcia-Dominguez, M. and Reyes, J.C. (2009) The chromatin remodeling factor CHD8 interacts with elongating RNA polymerase II and controls expression of the cyclin E2 gene. *Nucleic Acids Res.*, **37**, 2449–2460.
33. Garcia-Campmany, L. and Marti, E. (2007) The TGFbeta intracellular effector Smad3 regulates neuronal differentiation and cell fate specification in the developing spinal cord. *Development*, **134**, 65–75.
34. Akizu, N., Estaras, C., Guerrero, L., Marti, E. and Martinez-Balbas, M.A. (2010) H3K27me3 regulates BMP activity in developing spinal cord. *Development*, **137**, 2915–2925.
35. Hamburger, V. and Hamilton, H.L. (1951) A series of normal stages in the development of the chick embryo. 1951. *Dev. Dyn.*, **195**, 231–272.
36. Asensio-Juan, E., Fueyo, R., Pappa, S., Iacobucci, S., Badosa, C., Lois, S., Balada, M., Bosch-Presegue, L., Vaquero, A., Gutierrez, S. *et al.* (2017) The histone demethylase PHF8 is a molecular safeguard of the IFNgamma response. *Nucleic Acids Res.*, **45**, 3800–3811.
37. Frank, S.R., Schroeder, M., Fernandez, P., Taubert, S. and Amati, B. (2001) Binding of c-Myc to chromatin mediates mitogen-induced acetylation of histone H4 and gene activation. *Genes Dev.*, **15**, 2069–2082.
38. Valls, E., Blanco-Garcia, N., Aquizu, N., Piedra, D., Estaras, C., de la Cruz, X. and Martinez-Balbas, M.A. (2007) Involvement of chromatin and histone deacetylation in SV40 T antigen transcription regulation. *Nucleic Acids Res.*, **35**, 1958–1968.
39. Quinlan, A.R. and Hall, I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, **26**, 841–842.
40. Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G. and Mesirov, J.P. (2011) Integrative genomics viewer. *Nat. Biotechnol.*, **29**, 24–26.
41. Rada-Iglesias, A., Bajpai, R., Prescott, S., Brugmann, S.A., Swigut, T. and Wysocka, J. (2012) Epigenomic annotation of enhancers predicts transcriptional regulators of human neural crest. *Cell Stem Cell*, **11**, 633–648.
42. Kim, T.K., Hemberg, M., Gray, J.M., Costa, A.M., Bear, D.M., Wu, J., Harmin, D.A., Laptewicz, M., Barbara-Haley, K., Kuersten, S. *et al.* (2010) Widespread transcription at neuronal activity-regulated enhancers. *Nature*, **465**, 182–187.
43. Luyten, A., Zang, C., Liu, X.S. and Shivdasani, R.A. (2014) Active enhancers are delineated *de novo* during hematopoiesis, with limited lineage fidelity among specified primary blood cells. *Genes Dev.*, **28**, 1827–1839.
44. Eden, E., Navon, R., Steinfeld, I., Lipson, D. and Yakhini, Z. (2009) GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics*, **10**, 48.

45. Supek, F., Bosnjak, M., Skunca, N. and Smuc, T. (2011) REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One*, **6**, e21800.
46. Mullen, A.C., Orlando, D.A., Newman, J.J., Loven, J., Kumar, R.M., Bilodeau, S., Reddy, J., Guenther, M.G., DeKoter, R.P. and Young, R.A. (2011) Master transcription factors determine cell-type-specific responses to TGF-beta signaling. *Cell*, **147**, 565–576.
47. Trompouki, E., Bowman, T.V., Lawton, L.N., Fan, Z.P., Wu, D.C., DiBiase, A., Martin, C.S., Cech, J.N., Sessa, A.K., Leblanc, J.L. *et al.* (2011) Lineage regulators direct BMP and Wnt pathways to cell-specific programs during differentiation and regeneration. *Cell*, **147**, 577–589.
48. Castro, D.S. and Guillemot, F. (2011) Old and new functions of proneural factors revealed by the genome-wide characterization of their transcriptional targets. *Cell Cycle*, **10**, 4026–4031.
49. Raposo, A.A., Vasconcelos, F.F., Drechsel, D., Marie, C., Johnston, C., Dolle, D., Bithell, A., Gillotin, S., van den Berg, D.L., Ettwiller, L. *et al.* (2015) Ascl1 coordinately regulates gene expression and the chromatin landscape during neurogenesis. *Cell Rep.*, **9**, 1544–1556.
50. Estaras, C., Fueyo, R., Akizu, N., Beltran, S. and Martinez-Balbas, M.A. (2013) RNA polymerase II progression through H3K27me3-enriched gene bodies requires JMJD3 histone demethylase. *Mol. Biol. Cell*, **24**, 351–360.
51. Li, W., Notani, D. and Rosenfeld, M.G. (2016) Enhancers as non-coding RNA transcription units: recent insights and future perspectives. *Nat. Rev. Genet.*, **17**, 207–223.
52. McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M. and Bejerano, G. (2010) GREAT improves functional interpretation of cis-regulatory regions. *Nat. Biotechnol.*, **28**, 495–501.
53. De Santa, F., Totaro, M.G., Prosperini, E., Notarbartolo, S., Testa, G. and Natoli, G. (2007) The histone H3 lysine-27 demethylase Jmjd3 links inflammation to inhibition of polycomb-mediated gene silencing. *Cell*, **130**, 1083–1094.
54. Agger, K., Cloos, P.A., Christensen, J., Pasini, D., Rose, S., Rappsilber, J., Issaeva, I., Canaani, E., Salcini, A.E. and Helin, K. (2007) UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. *Nature*, **449**, 731–734.
55. Ceballos-Chavez, M., Subtil-Rodriguez, A., Giannopoulou, E.G., Soronellas, D., Vazquez-Chavez, E., Vicent, G.P., Elemento, O., Beato, M. and Reyes, J.C. (2015) The chromatin Remodeler CHD8 is required for activation of progesterone receptor-dependent enhancers. *PLoS Genet.*, **11**, e1005174.
56. Marfella, C.G. and Imbalzano, A.N. (2007) The Chd family of chromatin remodelers. *Mutat. Res.*, **618**, 30–40.
57. Nishiyama, M., Nakayama, K., Tsunematsu, R., Tsukiyama, T., Kikuchi, A. and Nakayama, K.I. (2004) Early embryonic death in mice lacking the beta-catenin-binding protein Duplin. *Mol. Cell Biol.*, **24**, 8386–8394.
58. Platt, R.J., Zhou, Y., Slaymaker, I.M., Shetty, A.S., Weisbach, N.R., Kim, J.A., Sharma, J., Desai, M., Sood, S., Kempton, H.R. *et al.* (2017) Chd8 mutation leads to autistic-like behaviors and impaired striatal circuits. *Cell Rep.*, **19**, 335–350.
59. Durak, O., Gao, F., Kaeser-Woo, Y.J., Rueda, R., Martorell, A.J., Nott, A., Liu, C.Y., Watson, L.A. and Tsai, L.H. (2016) Chd8 mediates cortical neurogenesis via transcriptional regulation of cell cycle and Wnt signaling. *Nat. Neurosci.*, **19**, 1477–1488.
60. Visel, A., Minovitsky, S., Dubchak, I. and Pennacchio, L.A. (2007) VISTA Enhancer Browser—a database of tissue-specific human enhancers. *Nucleic Acids Res.*, **35**, D88–92.
61. Sauvageot, C.M. and Stiles, C.D. (2002) Molecular mechanisms controlling cortical gliogenesis. *Curr. Opin. Neurobiol.*, **12**, 244–249.
62. Hand, R., Bortone, D., Mattar, P., Nguyen, L., Heng, J.I., Guerrier, S., Boutt, E., Peters, E., Barnes, A.P., Parras, C. *et al.* (2005) Phosphorylation of Neurogenin2 specifies the migration properties and the dendritic morphology of pyramidal neurons in the neocortex. *Neuron*, **48**, 45–62.
63. Ge, W., He, F., Kim, K.J., Bianchi, B., Coskun, V., Nguyen, L., Wu, X., Zhao, J., Heng, J.I., Martinowich, K. *et al.* (2006) Coupling of cell migration with neurogenesis by proneural bHLH factors. *Proc. Natl. Acad. Sci. U.S.A.*, **103**, 1319–1324.
64. Heng, J.I., Nguyen, L., Castro, D.S., Zimmer, C., Wildner, H., Armant, O., Skowronska-Krawczyk, D., Bedogni, F., Matter, J.M., Hevner, R. *et al.* (2008) Neurogenin 2 controls cortical neuron migration through regulation of Rnd2. *Nature*, **455**, 114–118.
65. Bertrand, N., Castro, D.S. and Guillemot, F. (2002) Proneural genes and the specification of neural cell types. *Nat. Rev. Neurosci.*, **3**, 517–530.
66. Wilkinson, G., Dennis, D. and Schuurmans, C. (2013) Proneural genes in neocortical development. *Neuroscience*, **253**, 256–273.
67. Fueyo, R., Garcia, M.A. and Martinez-Balbas, M.A. (2015) Jumonji family histone demethylases in neural development. *Cell Tissue Res.*, **359**, 87–98.
68. De Rubeis, S., He, X., Goldberg, A.P., Poultney, C.S., Samocha, K., Cicek, A.E., Kou, Y., Liu, L., Fromer, M., Walker, S. *et al.* (2014) Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature*, **515**, 209–215.
69. Sugathan, A., Biagioli, M., Golzio, C., Erdin, S., Blumenthal, I., Manavalan, P., Ragavendran, A., Brand, H., Lucente, D., Miles, J. *et al.* (2014) CHD8 regulates neurodevelopmental pathways associated with autism spectrum disorder in neural progenitors. *Proc. Natl. Acad. Sci. U.S.A.*, **111**, E4468–E4477.
70. Cotney, J., Muhle, R.A., Sanders, S.J., Liu, L., Willsey, A.J., Niu, W., Liu, W., Klei, L., Lei, J., Yin, J. *et al.* (2015) The autism-associated chromatin modifier CHD8 regulates other autism risk genes during human neurodevelopment. *Nat. Commun.*, **6**, 6404.
71. Wilkinson, B., Grepo, N., Thompson, B.L., Kim, J., Wang, K., Evgrafov, O.V., Lu, W., Knowles, J.A. and Campbell, D.B. (2015) The autism-associated gene chromodomain helicase DNA-binding protein 8 (CHD8) regulates noncoding RNAs and autism-related genes. *Transl. Psychiatry*, **5**, e568.
72. Katayama, Y., Nishiyama, M., Shoji, H., Ohkawa, Y., Kawamura, A., Sato, T., Suyama, M., Takumi, T., Miyakawa, T. and Nakayama, K.I. (2016) CHD8 haploinsufficiency results in autistic-like phenotypes in mice. *Nature*, **537**, 675–679.
73. Cruz-Molina, S., Respuela, P., Tebartz, C., Kolovos, P., Nikolic, M., Fueyo, R., van Ijcken, W.F.J., Grosveld, F., Frommolt, P., Bazzi, H. *et al.* (2017) PRC2 facilitates the regulatory topology required for poised enhancer function during pluripotent stem cell differentiation. *Cell Stem Cell*, **20**, 689–705.
74. Spitz, F. and Furlong, E.E. (2012) Transcription factors: from enhancer binding to developmental control. *Nat. Rev. Genet.*, **13**, 613–626.
75. Bonn, S., Zinnen, R.P., Girardot, C., Gustafson, E.H., Perez-Gonzalez, A., Delhomme, N., Ghavi-Helm, Y., Wilczynski, B., Riddell, A. and Furlong, E.E. (2012) Tissue-specific analysis of chromatin state identifies temporal signatures of enhancer activity during embryonic development. *Nat. Genet.*, **44**, 148–156.
76. Entrevan, M., Schuettengruber, B. and Cavalli, G. (2016) Regulation of genome architecture and function by polycomb proteins. *Trends Cell Biol.*, **26**, 511–525.
77. Frankel, N., Davis, G.K., Vargas, D., Wang, S., Payre, F. and Stern, D.L. (2010) Phenotypic robustness conferred by apparently redundant transcriptional enhancers. *Nature*, **466**, 490–493.
78. Whyte, W.A., Orlando, D.A., Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H., Rahl, P.B., Lee, T.I. and Young, R.A. (2013) Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell*, **153**, 307–319.
79. Huang, J., Liu, X., Li, D., Shao, Z., Cao, H., Zhang, Y., Trompouki, E., Bowman, T.V., Zon, L.I., Yuan, G.C. *et al.* (2016) Dynamic control of enhancer repertoires drives lineage and stage-specific transcription during hematopoiesis. *Dev. Cell*, **36**, 9–23.
80. Shin, H.Y., Willi, M., Yoo, K.H., Zeng, X., Wang, C., Metser, G. and Hennighausen, L. (2016) Hierarchy within the mammary STAT5-driven Wap super-enhancer. *Nat. Genet.*, **48**, 904–911.
81. Chen, S., Ma, J., Wu, F., Xiong, L.J., Ma, H., Xu, W., Lv, R., Li, X., Villen, J., Gygi, S.P. *et al.* (2012) The histone H3 Lys 27 demethylase JMJD3 regulates gene expression by impacting transcriptional elongation. *Genes Dev.*, **26**, 1364–1375.
82. De Santa, F., Narang, V., Yap, Z.H., Tusi, B.K., Burgold, T., Austenaa, L., Bucci, G., Caganova, M., Notarbartolo, S., Casola, S. *et al.* (2009) Jmjd3 contributes to the control of gene expression in LPS-activated macrophages. *EMBO J.*, **28**, 3341–3352.
83. Miller, S.A., Mohn, S.E. and Weinmann, A.S. (2010) Jmjd3 and UTX play a demethylase-independent role in chromatin remodeling to regulate T-box family member-dependent gene expression. *Mol. Cell*, **40**, 594–605.

84. Kang,S.C., Kim,S.K., Chai,J.C., Kim,S.H., Won,K.J., Lee,Y.S., Jung,K.H. and Chai,Y.G. (2015) Transcriptomic profiling and H3K27me3 distribution reveal both demethylase-dependent and independent regulation of developmental gene transcription in cell differentiation. *PLoS One*, **10**, e0135276.
85. Xun,J., Wang,D., Shen,L., Gong,J., Gao,R., Du,L., Chang,A., Song,X., Xiang,R. and Tan,X. (2017) JMJD3 suppresses stem cell-like characteristics in breast cancer cells by downregulation of Oct4 independently of its demethylase activity. *Oncotarget*, **8**, 21918–21929.