

Unravelling the epigenetic modifier Smchd1

by

Iromi Wanigasuriya
806410

[ORCID 0000-0002-2941-8764](https://orcid.org/0000-0002-2941-8764)

A thesis submitted in total fulfilment for the
degree of Doctor of Philosophy

in the

Department of Medical Biology
Faculty of Medicine, Dentistry and Health Sciences
THE UNIVERSITY OF MELBOURNE

October 2020

Abstract

Structural maintenance of chromosomes flexible hinge domain containing 1 (Smchd1) is an epigenetic modifier that plays an important role in X chromosome inactivation, autosomal gene silencing, and is also implicated in several diseases in humans [Blewitt et al., 2008, Mould et al., 2013, Gendrel et al., 2013, Lemmers et al., 2012, Gordon et al., 2017, Shaw et al., 2017, Jansz et al., 2018a]. Thus far the majority of work on Smchd1 has been carried out on its zygotic form. Therefore, I partnered allele-specific genomics with conditional deletion of Smchd1 in the oocyte to test the role of maternally derived Smchd1 in imprinted gene expression. I found that Smchd1 is a novel maternal effect gene involved in genomic imprinting. When Smchd1 is maternally deleted loss of imprinting is observed at ten imprinted genes, without affecting DNA methylation imprints. Additionally, I also discovered seven imprinted genes where zygotic Smchd1 plays a dose dependent role. Interestingly, almost all the imprinted genes affected by the loss of maternal Smchd1, including *Xist* possess H3K27me3 imprints. This, together with Smchd1's known role in long range chromatin interactions and function as insulator protein [Chen et al., 2015, Jansz et al., 2018a], lead me to hypothesise that maternal Smchd1 may carry out its function secondary to H3K27me3 imprints and establish a chromatin state required for imprinted expression.

To narrow down Smchd1's behaviour in its native environment and how this behaviour relates to its structure and function, I established a series of robust microscopy experiments. I used fluorescence recovery after photo bleaching (FRAP) and Lattice LightSheet microscopy to demonstrate for the first time, Smchd1 dynamics in the inactive X during interphase and mitosis. Through study of our unique gain of function mutant MommeD43 I found that the act of unbinding and reloading likely has a role in Smchd1's insulation effects. By using 3D-direct stochastic optical reconstruction microscopy (3D-dSTORM) I discovered that Smchd1 is approximately a 40 nm long protein in the nucleus. These novel techniques open up new avenues to explore higher order structures that Smchd1 may form in the nucleus related to its role in chromatin architecture.

Together these data not only revealed a new function for maternal Smchd1 but also established novel methods to unravel Smchd1's molecular mechanism.

Declaration of Authorship

I, Iromi Wanigasuriya, declare that this thesis titled, “Unravelling the epigenetic modifier Smchd1” and the work presented in it are my own. I confirm that:

- the thesis comprises only my original work towards Doctor of Philosophy except where indicated in the preface;
- due acknowledgement has been made in the text to all other material used; and
- the thesis is fewer than 100,000 words in length, exclusive of tables, maps, bibliographies and appendices as approved by the Research Higher Degrees Committee.

Signed:

Iromi Wanigasuriya

Date:

October 2020

Preface

In accordance with the regulations governing the degree of Doctor of Philosophy at the University of Melbourne, I hereby submit that:

- The author's contribution to Chapter 3 was 80%. Figure 3.1 - 3.15, Figure 6.1 and relevant figure legends in this thesis are extracts from publication accepted at eLife journal ([accessible via bioRxiv](#)). Subsection 3.3.11 and 3.3.12 in this thesis are also adapted from the same paper. I am a co-first author in this manuscript, I created all the relevant figures and contributed significantly to the writing of the manuscript. Ms Tamara Beck collected some embryos and placentae. RRBS libraries were prepared by Ms Tamara Beck. Immunofluorescence of ovary sections was carried out in collaboration with Dr Jessica Stringer and Dr Karla Hutt at Monash BDI. M&T-seq was carried out in collaboration with Ms Ellise Roper and Dr Heather Lee at the University of Newcastle. All genomic analyses were carried out in collaboration with Dr Quentin Gouil;
- The author's contribution to Chapter 4 was 80%. Micro-CT image processing was done by Dr Lachlan Whitehead (Centre for Dynamic Imaging, WEHI). Ms Tamara Beck collected some placentae. RRBS libraries were prepared by Ms Tamara Beck. M&T-seq was carried out in collaboration with Ms Ellise Roper and Dr Heather Lee at the University of Newcastle. All genomic analyses were carried out in collaboration with Dr Quentin Gouil;
- The author's contribution to Chapter 5 was 80%. Dr Michael Mlodzianoski carried out all dSTORM image acquisition and analysis. For FRAP experiments, technical expertise was provided by A/Prof Kelly Rodgers and Dr Niall Geoghegan. Dr Quentin Gouil helped to analyse FRAP data. For Lattice LightSheet microscopy Dr Niall Geoghegan provided technical expertise and reconstructed images. Dr Tracy Wilson produced the NLS-acGFP construct and H2B-mScarlet construct.

Therefore, the author's overall contribution to the work presented in this thesis was 80%.

- The publication status of all chapters presented;
 - Chapter 3 - Published in [*eLife*] on [13-11-2020] pre-print accessible via [*eLife*]
 - Chapter 4 - Unpublished material not submitted for publication
 - Chapter 5 - Unpublished material not submitted for publication
- The author was supported by a Melbourne University International Research Scholarship and the Walter and Eliza Hall Institute of Medical Research PhD top-up scholarship. This work was supported by a NHMRC grant to Marnie E Blewitt and Matthew E Ritchie (1098290). Additional support was provided by the Victorian State Government Operational Infrastructure Support, Australian National Health and Medical Research Council IRIISS grant (9000433).

Acknowledgements

As an undergraduate student I read a short story on Irving Stone. Throughout the years one quote from his story stood out to me and now at the brink of my PhD I can certainly relate to this quote more than ever.

“Inspiration does not have to come whole and complete, it can come in many segments, divided by years and miles” - Irving Stone

Certainly, a village of people throughout my life helped me to get where I am today. And I am grateful to each and every one of them. I can only hope that I can pay forward the kindness and support each and every one acknowledged here has shown me.

First and foremost I would like to thank my primary supervisor Prof Marnie Blewitt. Words fail to describe what an integral role Marnie played not only as my PhD advisor but also as a role model in my life. I’m thankful to Marnie for believing in me and taking a leap of faith and recruiting me as a PhD student after meeting online for the first time on discussion forums in her Coursera course. Marnie was extremely patient, kind and understanding throughout the last few years. Her amazing work ethic and managerial skills has certainly left a mark on me. I am so very grateful for various opportunities Marnie provided me to expand my scientific and technical skills and making sure I am ready to tackle any future career prospects.

I am so grateful to my secondary supervisors, Dr Andrew Kenriy and Dr Matt Ritchie. Both were excellent mentors who were always enthusiastic and supportive. I would especially like to thank Andrew for taking his time to teach me the ropes of bioinformatics and other lab techniques, and Matt for attending all of my supervisor meetings even when they were not bioinformatics related and providing helpful insights. I would also like to thank my PhD advisory committee members: Prof Jane Visvader, Dr Samir Taoudi, A/Prof Kelly Rodgers, and Dr Sarah Kinkel for their valuable insights and support throughout the years.

I would particularly like to thank Dr Quentin Gouil for always being so enthusiastic about science and carrying out a significant amount of bioinformatic analysis presented in this thesis. It has been my absolute privilege to be co-first author with him in my first paper from my PhD. Quentin’s organisation skills, stimulating

ideas, and love for outdoors have certainly positively influenced me. Thank you for always suggesting great tools to improve my skills and making science fun and exciting. I certainly miss our pre-covid coffee and croissant catch-ups.

Heartfelt thank you to all the past and present Blewitt lab members for creating an extremely supportive environment to work in. First, Alexandra Gurzau and Andres Tapia del Fierro who closely shared the PhD experience with me, thank you for being amazing friends both inside and outside of the lab and making the past few years wholesome and enjoyable. I would like to acknowledge past PhD students Kelan Chen and Natasha Jansz who worked on Smchd1 before me and providing me with such excellent footsteps to follow. Thank you to Casey Ah-Cann who made lab environment fun, introduced me to "Art of Science" and for her helpful insights over coffee. Many thanks to Kelsey Breslin and Tamara Beck for helping me out with new techniques, carrying out dissections and genotyping, and also being such wonderful and kind people. Thank you to all the other past and present lab members Hannah Vannayi, Natalia Benetti, Jenny Steiner, Maya Harper, Ruoyan Wang, Joy Liu, Megan Imnitoff and Ayush Semwal for their support throughout this journey. It has been my absolute pleasure to work with each and every one of them.

Special mention to WEHI center for Dynamic Imaging which has grown as I grew in my PhD. A/Prof Kelly Rodgers who was not only one of my PhD committee members but also an excellent mentor who provided valuable expertise in imaging techniques. Special thanks to Dr Niall Geoghagen for not only contributing to chapter five of my thesis but also for our many wonderful chats during long imaging sessions. I really appreciate all encouragement provided throughout the years. Thank you to Dr Lachlan Whitehead for answering my last-minute imaging analysis questions, long hours spent at the Micro-CT and many insightful conversations about life outside of PhD, microscopy nightmares and Lego among other things. I would also like to thank Dr Michael Mlodzianoski for his help with super-resolution microscopy and chats about all things science and food.

I would also like to thank Dr Tracy Wilson for her assistance with making different constructs that have been vital for my imaging experiments. Dr Michal Pasternak, Dr Kate MacArthur, and Caleb Dawson for imaging insights. Jessica Martin and Andrea Briffa for assisting me with mice experiments and taking such good care of our mouse facility. Dr James Murphy for insights on protein function. Stephen Wilcox for sequencing help. Thank you to Sue Hardy, Fiona McGrath, Tanya de

Jong-Curtin and Sonja Gustin for administrative support throughout my PhD. Thank you to Dr Dawn Lin, Maggie Potts and Sara Tomei for being wonderful friends both inside and outside of WEHI. I would also like to acknowledge many members of former molecular medicine division and current epigenetics and development division for all their support throughout the years. I am grateful to the extremely supportive environment at WEHI, it has been my absolute pleasure and privilege to be a PhD student at such a wonderful institute. Special thanks to Prof Doug Hilton for making WEHI a wonderful environment for students and staff and leading by example.

Many thanks to all my collaborators who provided support in different aspects of my thesis. Dr Kirstin Elgass from Hudson Institute, Dr Heather Lee and Ellise Roper from University of Newcastle, Dr Jessica Stringer and Dr Karla Hutt from Monash BDI, and Dr Elizabeth Hinde and Dr Jieqiong Lou from University of Melbourne.

I would like to thank my mentors Dr Hasini Jayathilake and Ross Barrow for providing me insights on life outside of science and life in general. I'm grateful for their ongoing support. Many thanks to all my teachers and mentors throughout the years. Special mention to Prof Mohamed Noor from Duke University who conducted the first-ever Coursera course I took and laid the foundation to all my work that followed. Thank you for still supporting me and encouraging me after all these years. I have met a wonderful bunch of people through Mohamed's course and formed connections for life. Special thanks to Silvia, Christine, Lisa and Aysegul for being my biggest cheerleaders. I hope to meet you all in a post-Covid world.

Finally, I would also like to thank all my friends living all over the world right now, for their never-ending support and friendship that has withstood the test of time and distance. Thank you to Kirk for our many sessions, your insights and support certainly made this PhD journey possible. I'm grateful to my mum for providing me with strength and support throughout life. I would also like to thank my family both here in Melbourne and in Sri Lanka. My partner Janidu, thank you for being a pillar of strength for over a decade. Your unconditional love, silent support and strength, especially during these testing times, have been invaluable. Thank you for making delicious meals, providing me with support and comfort and making sure I have nothing to ask for. To Ollie, my furbaby, for being the best thing to happen to me during the final year of my PhD. For always making my day

and taking away all my stress and anxiety, making sure I get some exercise during long writing sessions and for making the extended Melbourne lockdown bearable. Thank you for filling my days with unconditional love.

Writing a thesis at any given time is a challenge and writing one through a global pandemic and the many obstacles that come with it is even more of a challenge. I certainly wouldn't have been able to do it without all of your support. Thank you from the bottom of my heart.

Contents

Declaration of Authorship	ii
Preface	iii
Acknowledgements	v
List of Figures	xiv
List of Tables	xvii
Abbreviations	xviii
1 Introduction	1
1.1 Epigenetic control of gene expression	1
1.2 Mechanisms of epigenetic regulation	5
1.2.1 DNA methylation	5
1.2.2 Histone modifications	7
1.2.2.1 Polycomb repressive complexes (PRC)	8
1.2.3 Chromatin remodelling proteins	9
1.2.4 Long non-coding RNAs	9
1.2.5 Higher order chromatin organisation	10
1.3 Epigenetic reprogramming	13
1.4 Monoallelic gene expression	16
1.4.1 X chromosome inactivation (XCI)	16
1.4.1.1 Mechanism of XCI	16
1.4.2 Conformation of Xi	19
1.4.2.1 XCI in mouse development	20
1.4.3 Genomic imprinting	22
1.4.3.1 Parental conflict theory	24
1.4.3.2 Imprinting disorders	25
1.5 Maternal effect proteins	27
1.5.1 Maternal effect proteins in imprinted gene regulation	27

1.6	Structural Maintenance of Chromosome Hinge Domain containing 1 (<i>Smchd1</i>)	28
1.6.1	Discovery of the epigenetic regulator <i>Smchd1</i>	28
1.6.2	<i>Smchd1</i> in the XCI	30
1.6.3	<i>Smchd1</i> in imprinted gene regulation	31
1.6.4	<i>Smchd1</i> in higher order chromatin structure	33
1.6.5	<i>Smchd1</i> protein structure	34
1.6.6	<i>Smchd1</i> mutations are implicated in disease	35
1.7	Aims	38
1.7.1	Aim 1: To evaluate whether <i>Smchd1</i> is a maternal effect protein that has a role in genomic imprinting	38
1.7.2	Aim 2: To unravel the <i>Smchd1</i> macro molecular structure and dynamics in the nucleus	38
2	Methods	39
2.1	Mouse strains	39
2.1.1	Ethics approvals	39
2.1.2	MMTV-Cre <i>Smchd1</i> ^{-fl} line	39
2.1.3	Zp3-Cre <i>Smchd1</i> ^{fl} line	40
2.1.4	CAST/EiJ line	40
2.1.5	<i>Smchd1</i> ^{GFP} line	40
2.1.6	Matings	40
2.1.7	Genotyping	40
2.2	Derivation and culture of cell lines	41
2.2.1	Mouse embryonic fibroblasts (MEFs)	41
2.2.2	Neuronal Stem Cells (NSCs)	41
2.3	Retrovirus production and transduction	42
2.3.1	Retrovirus production	42
2.3.2	Concentration of retrovirus with PEG	42
2.3.3	Transduction of MEFs	43
2.3.4	Transduction of NSCs	43
2.4	Microscopy	43
2.4.1	Immunofluorescence studies on mouse ovary sections	43
2.4.2	Native GFP fluorescence in preimplantation embryos	44
2.4.3	Immunofluorescence in preimplantation embryos	44
2.4.4	Micro-CT	44
2.4.5	Allele specific RNA-FISH on preimplantation embryos	45
2.4.5.1	Denhardt's treatment of coverslips	45
2.4.5.2	Probe preparation	45
2.4.5.3	Preparation of strand-specific <i>Xist</i> guide RNA probe	46
2.4.5.4	Preparation of pre-implantation embryos	46
2.4.5.5	Allele specific RNA-FISH	47
2.4.6	Fluorescence recovery after photobleaching (FRAP)	47
2.4.7	Mitotic shake-off	47

2.4.8	Lattice LightSheet microscopy	48
2.4.9	Direct stochastic optical reconstruction microscopy (dSTORM) sample preparation	49
2.5	Genomic assays	50
2.5.1	Embryos and placentae collection for bulk RNA-seq and DNA-seq	50
2.5.2	Library preparation - bulk RNA-seq	50
2.5.3	Library preparation - bulk reduced representation bisulfite sequencing (RRBS)	51
2.5.4	Single-preimplantation embryo sequencing - M&T-seq	51
3	<i>Smchd1</i> is a maternal effect gene required for genomic imprinting	54
3.1	Abstract	54
3.2	Introduction	55
3.3	Results	56
3.3.1	Maternal <i>Smchd1</i> expression in the early embryo	56
3.3.2	Timing of maternal <i>Smchd1</i> deletion during oogenesis in our maternal null models	57
3.3.3	Maternal deletion of <i>Smchd1</i> depletes <i>Smchd1</i> until the 16- cell embryonic stage	58
3.3.4	Maternal deletion of <i>Smchd1</i> does not cause pre-weaning lethality or altered postnatal weight in progeny	60
3.3.5	Maternal deletion of <i>Smchd1</i> does not have a striking effect on gene expression	61
3.3.6	No striking difference between <i>Smchd1</i> maternal deletion models - MMTV-Cre and Zp3-Cre	63
3.3.7	Maternal <i>Smchd1</i> is not required to regulate the imprinted expression at known <i>Smchd1</i> targets in the <i>Snrpn</i> and <i>Igf2r</i> clusters	65
3.3.8	Genes within the <i>Kcnq1</i> cluster are also under the control of zygotic <i>Smchd1</i>	67
3.3.9	Maternal deletion of <i>Smchd1</i> does not affect DNA methyla- tion at germline or secondary DMRs for canonical imprinted genes	68
3.3.10	Novel imprinted genes under the control of maternal <i>Smchd1</i>	69
3.3.11	Both maternal and zygotic <i>Smchd1</i> contribute to canonical imprinted gene expression	72
3.3.12	Both maternal and zygotic <i>Smchd1</i> contribute to non-canonical imprinted gene expression	73
3.3.13	Maternal <i>Smchd1</i> sets up imprinted expression pre-implantation	76
3.3.14	Maternal <i>Smchd1</i> does not affect methylation at DMRs dur- ing the preimplantation period	78
3.4	Discussion	80
4	Maternal <i>Smchd1</i> is required for imprinted XCI	84

4.1	Abstract	84
4.2	Introduction	85
4.3	Results	86
4.3.1	Upregulation of expression from the paternal X chromosome in pre-implantation embryos in the absence of maternal <i>Smchd1</i>	86
4.3.2	Loss of maternal <i>Smchd1</i> results in expression of maternal <i>Xist</i> in both male and female E2.75 embryos	88
4.3.3	Maternal <i>Smchd1</i> is required for faithful monoallelic <i>Xist</i> expression in preimplantation embryos	89
4.3.4	Maternal deletion of <i>Smchd1</i> results in failed iXCI in the placenta	92
4.3.5	Maternal deletion of <i>Smchd1</i> does not affect <i>Xist</i> expression in E14.5 placentae	93
4.3.6	Maternal deletion of <i>Smchd1</i> results in a widespread effect on DNA methylation on the Xi in E14.5 placentae	93
4.3.7	Maternal deletion of <i>Smchd1</i> does not affect methylation of CpG islands at XIC in placental samples	95
4.3.8	Maternal deletion of <i>Smchd1</i> results in a female-specific placental defect	98
4.3.9	Maternal deletion of <i>Smchd1</i> does not have a sex specific effect on the weight of the progeny	99
4.4	Discussion	101
5	Macro-molecular structure and dynamics of <i>Smchd1</i>	105
5.1	Abstract	105
5.2	Introduction	106
5.2.1	<i>Smchd1</i> gain of function mutant - MommeD43	106
5.2.2	Microscopy methods	108
5.2.2.1	Functional analysis by Fluorescence Recovery after Photobleaching (FRAP)	109
5.2.2.2	Fast volumetric imaging using lattice light sheet microscopy	111
5.2.2.3	Structural analysis using single molecule imaging methods	112
5.3	Results	114
5.3.1	<i>Smchd1</i> shows slow dynamics during interphase	114
5.3.2	<i>Smchd1</i> does not remain bound to Xi during mitosis	117
5.3.3	Macro-molecular structure of full length <i>Smchd1</i> in the nucleus	119
5.3.3.1	3D-STORM sample analysis	122
5.3.3.2	<i>Smchd1</i> is approximately 40 nm in length in the nucleus	123
5.4	Discussion	126
6	General Discussion	131

6.0.1	Aim 1: To evaluate whether Smchd1 is a maternal effect protein that has a role in genomic imprinting	131
6.0.2	Aim 2: To unravel the Smchd1 macro molecular structure and dynamics in the nucleus	135
A Differential imprinted expression in <i>Smchd1</i> maternal null vs wildtype male E2.75 embryos		137
Bibliography		142

List of Figures

1.1	Epigenetic landscape	2
1.2	Chromatin organisation	3
1.3	Epigenetic control of gene expression	4
1.4	The hierarchical organisation of the genome	11
1.5	DNA methylation reprogramming during development	15
1.6	Kinetics of X-inactivation	18
1.7	The ontogeny of XCI in mouse	21
1.8	Canonical imprinted gene regulation	24
1.9	ENU mutagenesis screen for modifiers of transgene silencing	29
1.10	Smchd1 regulates genes at Snrpn cluster	32
1.11	Smchd1 is Structurly similar to SMC family proteins	34
3.1	Maternal Smchd1 is detected in preimplantation embryos.	57
3.2	Maternal Smchd1 is deleted at different stages of oocyte development by MMTV-Cre and Zp3-Cre.	58
3.3	Zygotic Smchd1 is not detectable until around the 16-cell embryonic stage of embryonic development.	59
3.4	Maternal deletion of <i>Smchd1</i> does not result in pre-weaning lethality or altered postnatal weight	62
3.5	Maternal deletion of <i>Smchd1</i> does not have a major effect on gene expression in the embryo or the placenta	63
3.6	MTTV-Cre and Zp3-Cre shows consistant effects on imprinted expression	64
3.7	Maternal deletion of <i>Smchd1</i> does not affect imprinting at known Smchd1-sensitive clusters.	66
3.8	Haploinsufficiency for <i>Smchd1</i> results in relaxation of imprinting at the <i>Kcnq1</i> imprinted cluster	68
3.9	Maternal deletion of <i>Smchd1</i> does not affect DNA methylation at DMRs associated with canonical imprinted clusters sensitive to loss of Smchd1	70
3.10	Maternal deletion of <i>Smchd1</i> results in relaxation of imprinting non-canonical imprinted genes	71
3.11	Zygotic deletion of <i>Smchd1</i> results in complete loss of imprinting at proximal genes in the <i>Snrpn</i> cluster in embryos.	74

3.12	Zygotic deletion of <i>Smchd1</i> results in more severe loss of imprinting at canonical imprinted genes sensitive to heterozygosity for <i>Smchd1</i> in the placenta	75
3.13	Zygotic deletion of <i>Smchd1</i> has variable effects on non-canonical imprinted genes sensitive to maternal <i>Smchd1</i>	76
3.14	Maternal <i>Smchd1</i> is required for imprinted expression of select non-canonical imprinted genes in the preimplantation embryo.	77
3.15	Maternal <i>Smchd1</i> does not affect DMR methylation at canonical imprinted genes in the preimplantation period	79
4.1	Maternal deletion of <i>Smchd1</i> causes aberrant expression from the paternal X chromosome in female E2.75 preimplantation embryos	87
4.2	Maternal deletion of <i>Smchd1</i> results in <i>Xist</i> expression from the maternal allele in both male and female E2.75 preimplantation embryos	89
4.3	Maternal deletion of <i>Smchd1</i> results in biallelic expression of <i>Xist</i> from female embryos and <i>Xist</i> expression from male embryos	91
4.4	Maternal deletion of <i>Smchd1</i> does not affect methylation at XIC in female embryos	92
4.5	Maternal deletion of <i>Smchd1</i> results in failed silencing of the imprinted Xi in the placenta	94
4.6	Imprinted expression of <i>Xist</i> is unaffected in <i>Smchd1</i> maternal null placentae	95
4.7	Maternal deletion of <i>Smchd1</i> shows significant effect on CpG island hypermethylation on the imprinted Xi	96
4.8	Maternal deletion of <i>Smchd1</i> does not significantly affect methylation at XIC in E14.5 placentae	97
4.9	Maternal deletion of <i>Smchd1</i> results in smaller female placentae at E10.5	99
4.10	Maternal deletion of <i>Smchd1</i> does not have a sex specific effect on the progeny weights at P2	100
5.1	The observation of biology across a broad range of scales using different microscopy techniques.	107
5.2	Triangle of “Frustration”	110
5.3	Principals of Fluorescence Recovery After Photobleaching (FRAP)	111
5.4	Lattice lightsheet microscope set up compared to conventional microscopy	112
5.5	Schematic of basic principle of single molecule localisation microscopy	113

5.6	FRAP movie in Smchd1-GFP iMEF	115
5.7	Diffusion-uncoupled model used to analyse FRAP data . . .	116
5.8	Slow dynamics of Smchd1 cannot be attributed to GFP maturation	117
5.9	Smchd1 – MommeD43 mutant shows faster dynamics . . .	118
5.10	Smchd1 doesn't stay bound to to Xi during cell division . .	119
5.11	Lattice lightsheet movie showing loss of Smchd1 from the Xi chromatin during mitosis	120
5.12	Smchd1 macro molecular structure using 3D-dSTORM microscopy - experimental set up	121
5.13	3D-dSTORM microscopy - Image analysis	124
5.14	dSTORM fluorescent Probe separation	125
6.1	Maternal Smchd1 translates the imprints to establish a heritable chromatin state required for imprinted expression.	133

List of Tables

1.1	Imprinting disorders	26
2.1	Primers for genotyping	52
2.2	Primary and secondary antibodies	52
2.3	T3/T7 <i>Xist</i> primers for single stranded FISH probe template	52
2.4	<i>Xist</i> allele specific FISH probes	53
3.1	Smchd1 sensitive imprinted genes in mice	81
5.1	Smchd1 is approximately 40 nm in length in both primary and immortalised cells	125

Abbreviations

ATP	Adenosine triphosphate
AS	Angelman syndrome
bp	base pairs
BAMS	Bosma Arhinia and Microphthalmia Syndrome
Bl6	C57BL/6
BWS	Beckwith Wiedemann syndrome
C57	C57BL/6
Cast	CAST/EiJ
ChIP-seq	Chromatin immunoprecipitation sequencing
CGIs	CpG islands
CNVs	Copy number variations
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
DAPI	4',6'-diamidino-2-phenylindole
DNA	DeoxyriboNucleic Acid
DNA-seq	DNA sequencing
DMEM	Dulbecco's Modified Eagle Medium
DMRs	Differentially methylated regions
DMSO	dimethyl sulfoxide
DNMT	DeoxyriboNucleic Acid Mmethyltransferase
dNTPs	deoxynucleotides
dSTROM	Direct stochastic optical reconstruction microscopy
E	Embryonic day
EM	Electron microscopy
ENU	<i>N</i> -ethyl- <i>N</i> -nitrosourea

FCS	F luorescence C orrelation S pectroscopy
FBS	F oetal B ovine S erum
FDR	F alse D iscovery R ate
FISH	F luorescence i n s itu h ybridization
FLIM	F luorescence lifetime i maging
FRAP	F luorescence r ecovery a fter p hotobleaching
FRET	F örster r esonance e nergy t ransfer
FSHD	F acioscapulohumeral M uscular D ystrophy
GFP	G reen F luorescent P rotein
GHKL	G yrase, H sp90, histidine k inase, Mut L
H2AK119ub1	Monoubiquitinate lysine 119 of histone H2A
H3K27me3	Tri-methylation of Histone 3 Lysine 27
H3K4me3	Tri-methylation of Histone 3 Lysine 4
HAT	H istone a cetyl t ransferases
HBS	H EPES b uffered s aline
HDAC	H istone d eacetylase
HKMT	H istone lysine m ethyl t ransferase
IAPs	I ntercisternal A particles
ICRs	I mprinting c ontrol r egions
ICM	I nnner C ell M ass
iMEFs	I mmortalised M ouse E mbryonic F ibroblasts
iXCI	i mprinted X C hromosome I nactivation
KDM	Lysine d emethylase
LLPS	L iquid–liquid p hase s eparation
LLSM	L attice L ight S heet M icroscopy
m⁵C	5 methyl cytosine
logFC	l og F old C hange
MEFs	M ouse E mbryonic F ibroblasts
M&T-seq	M ethylation & T ranscriptome s equencing
MLID	M ulti L ocus I mprinting D isturbances
MMTV	M ammary tumor v irus

Mommes	M urine met astable epi alleles
NA	N umerical A perture
ncRNA	n oncoding RiboN ucleic A cid
NLS	N uclear l ocalisation s ignal
NSCs	N eural S tem C ells
PBS	P hosphate B uffered S aline
PCR	P olymerase c hain r eaction
PEG	P olyethylene g lycol
PFA	P ara f ormaldehyde
PGCs	P rimordial g erm c ells
PPPS	P olymer– p olymer p hase s eparation
PRC	P olycomb r epressive c omplex
PSF	P oint S pread F unctions
PWS	P raeder- W illi s yndrome
RNA	R ibonucleic A cid
RNA-seq	R NA s equencing
ROI	R egion o f i nterest
RRBS	R educed R epresentation B isulfite S equencing
rXCI	r andom X C hromosome I nactivation
SIM	S tructural i llumination m icroscopy
shRNA	s hort h airpin R NA
SMC	S tructural M aintenance of C hromosomes
SMCHD1	S tructural M aintenance of C hromosome H inge D omain containing 1
SNPs	S ingle N ucleotide P olymorphisms
SSC	S aline- S odium C itrate
STED	S timulated e mission d epletion
SWI/SNF	SW itch/ S ucrose N on- F ermentable
TADs	T opologically A ssociating D omains
TF	T ranscription F actor
UPDs	U niparental d isomies

WEHI	W alter and E liza H all I nstitute
XCI	X C hromosome I nactivation
Xi	Inactive X
XCI	X inactivation c enter
Xist	X -inactivation s pecific t ranscript
Xm	Maternal X
Xp	Paternal X
Zp3	Z ona p ellucida 3

Chapter 1

Introduction

1.1 Epigenetic control of gene expression

The term epigenetics as we know today was first conceived by embryologist Conrad Waddington in 1942. Waddington proposed the "epigenetic landscape" as a model (Figure 1.1) to illustrate how the cellular lineage (epigenesis) changes occur as a result of a complex developmental interplay between the genotype and phenotype of an organism [[Waddington, 1942](#), [Deichmann, 2016](#)].

This concept has undergone an evolution of its own, so much so that today's definition bears only conceptual similarity to Waddington's definition [[Greally, 2018](#)]. Epigenetics in the context of molecular biology can be defined as all non-genetic information that results in changes to gene expression without altering the underlying deoxyribonucleic acid (DNA) sequence. These epigenetic changes are mitotically heritable through cell divisions [[Russo et al., 1996](#), [Riggs, 2002](#), [Probst et al., 2009](#), [Oomen and Dekker, 2017](#)]. Epigenetic control of gene expression helps to give rise to all different complex cell types in a multi-cellular organism from a single cell; even though every single cell contains the same genetic information. It is the epigenetic landscape that helps to influence which genes are expressed at any given time. Therefore, while the definition has become refined molecularly, the concept has remained the same as Waddington's original description.

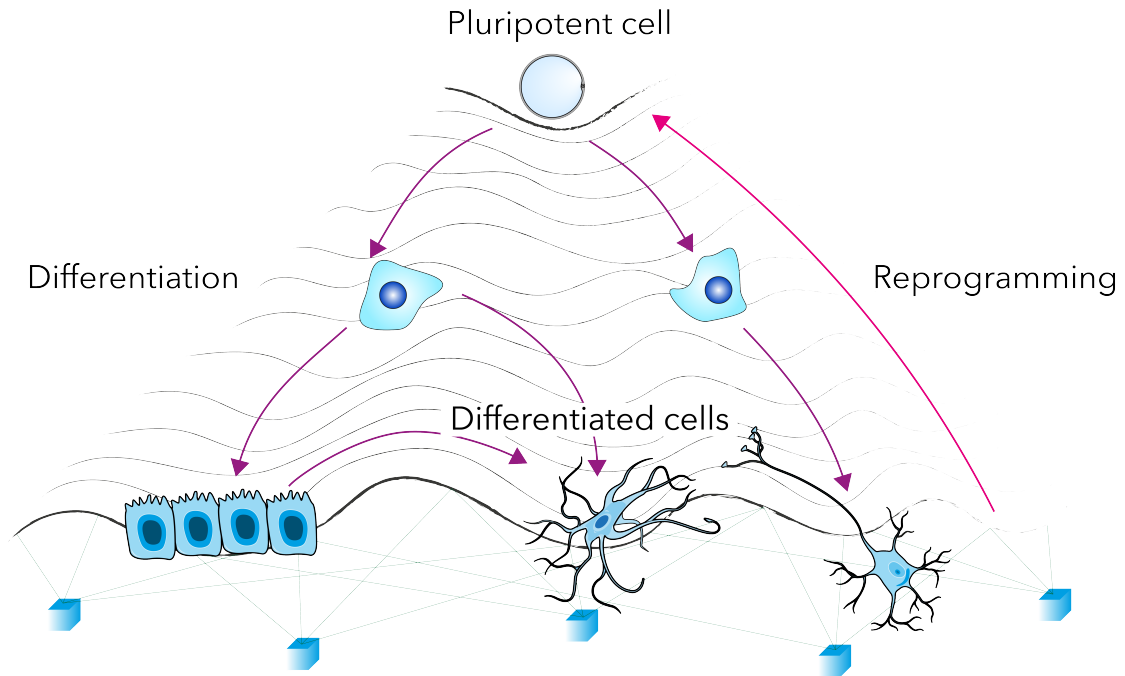


FIGURE 1.1: **Epigenetic landscape**

The epigenetic landscape is a phenomenon described by Conrad Waddington to explain how a pluripotent cell undergoes differentiation, committing to various somatic cell types. This process can be controlled by underlying epigenetic mechanisms. Differentiated cells are depicted in various formats depending on their cell type. Purple arrows indicate differentiation and blue cubes connected to the landscape by straight lines indicate the epigenetic network underpinning differentiation. The pink arrow represents reprogramming back to the original pluripotent cell.

Mammalian DNA is made up of over two billion bases coding for over 20 000 genes and contains all the necessary genetic information to sustain life. DNA is about two meters long and is tightly packed into the nucleus that is only a few micrometres in diameter. As organisms increase in complexity the packaging of the genome also needs to be more complex. Packaging is achieved through the interaction of negatively charged DNA with positively charged histone proteins[Alberts et al., 2002]

DNA and histones together are defined as chromatin (Figure 1.2). Some of the first visualisations of the packaging of chromatin came from electron micrograph images, that showed a structure described as "beads on a string" [Olins and Olins, 1974, Khorasanizadeh, 2004]. The beads in this structure are the nucleosomes, the smallest packing unit of chromatin. Nucleosomes are made up of two copies of four core histone proteins; H2A, H2B, H4, and H3 with 147 bp of DNA wrapped

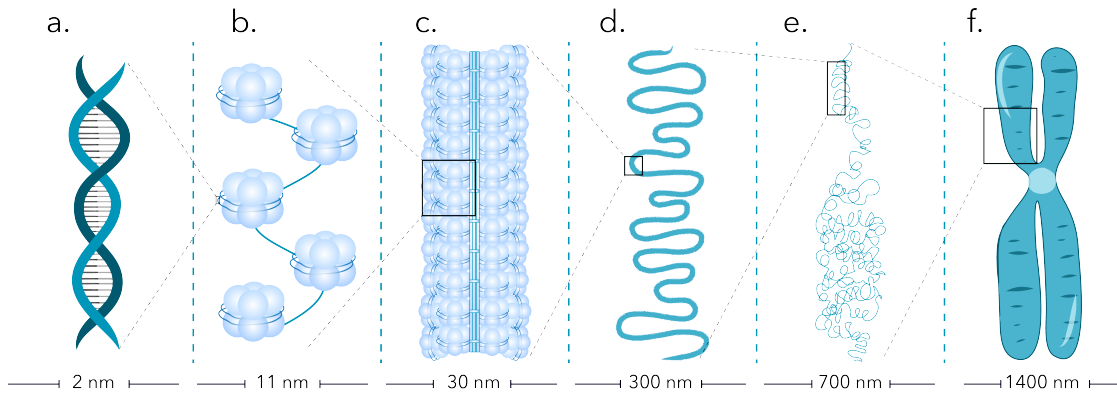


FIGURE 1.2: **Chromatin organisation**

Figure illustrates a. DNA b. DNA wrapped round histone octamers (nucleosomes) c. 30 nm chromatin fiber d. chromatin loops e. higher order chromatin structures and f. a chromosome. Figure adapted from [Felsenfeld and Groudine, 2003].

around the outside [Luger et al., 1997]. The string between the beads is DNA without histone proteins. The images and resulting ‘beads on a string’ model revealed a regular pattern of nucleosome binding along DNA, giving the first level of packaging that exists in chromatin (Figure 1.2b).

Histones possess N-terminal tails that protrude out of the core nucleosome, providing perfect surfaces for covalent modification by chromatin-modifying enzymes; for example methylation, ubiquitylation, acetylation, phosphorylation [Bannister and Kouzarides, 2011]. Chromatin binding proteins bind nucleosomes and may contribute to higher order secondary structures [Jacobs and Khorasanizadeh, 2002, Nuebler et al., 2018]. In general, these structures correlate with overall accessibility of the DNA to other factors such as the transcription machinery, and therefore contribute to gene expression control [Klemm et al., 2019].

The regions of chromatin that are more accessible or loosely packed are known as euchromatin. On the contrary inaccessible, more tightly packed regions are known as heterochromatin. Heterochromatin can be further subdivided into facultative and constitutive heterochromatin. Facultative heterochromatin contains transiently repressed genes that are differentially expressed in distinct cell types at specific developmental stages e.g. genes required in one cell type but not another.

On the other hand constitutive heterochromatin contains more permanently repressed genomic regions such as centromeres and telomeres that are required to be densely packaged to ensure the structural integrity of the genome [Allshire and Madhani, 2018].

The three dimensional structure of chromatin structure is achieved by an interplay between numerous epigenetic modifiers such as DNA methylation, histone modifications, chromatin binding proteins, long non-coding RNAs (long ncRNA) and also the transcription factors resulting in a highly complex higher order chromatin structure that regulates when and which genes are expressed in each and every different cell type in an organism [Allis and Jenuwein, 2016]. In culmination, the spatio-temporal distribution of the genome in the nucleus correlates to the function of the cell.

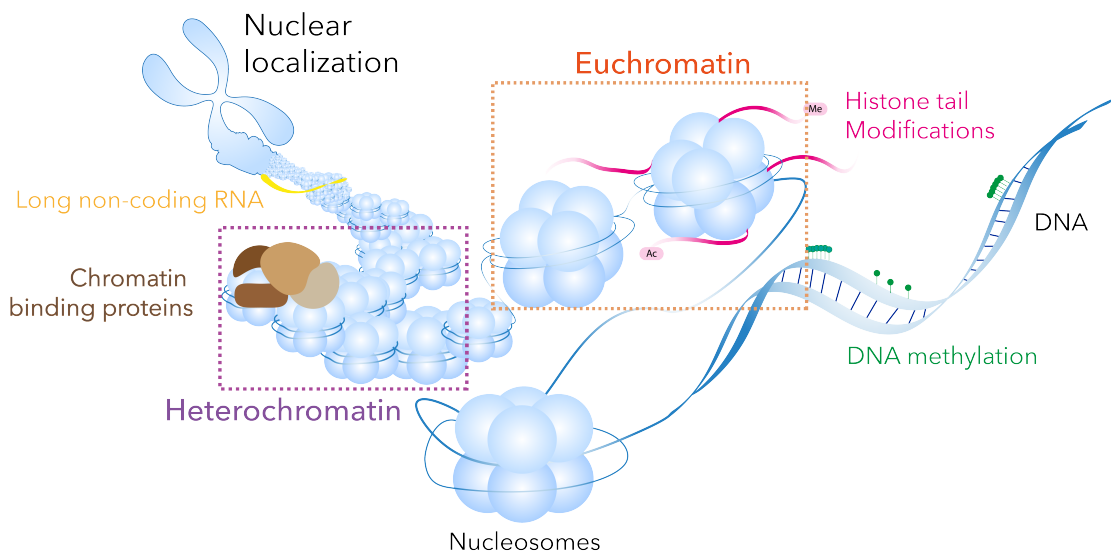


FIGURE 1.3: **Epigenetic control of gene expression**

Figure depicts different forms of epigenetic mechanisms that contribute to gene expression control. DNA contains **DNA methylation**, and histones contain **histone tail modifications**. DNA and Histones together form the chromatin. More open chromatin known as euchromatin is highlighted in an **orange dotted box** and closed heterochromatin is highlighted with a **purple dotted box**. Various epigenetic modifications are modulated by a host of **chromatin binding proteins** together with **long ncRNA**. Finally, the chromosomes are localised to different regions of the nucleus (not shown).

This exceptionally intricate network is known as epigenetic control of gene expression (Figure 1.3) and it is essential for the normal function of an organism

throughout its life. If things go awry, it can result in various developmental defects and diseases as highlighted in the following sections. This emphasizes the necessity to unravel these epigenetic modifiers, understand their mechanisms and study how things go wrong in disease.

1.2 Mechanisms of epigenetic regulation

1.2.1 DNA methylation

DNA methylation is undoubtedly the most well characterised epigenetic modification. Historically, it was first proposed in 1975 that DNA methylation is possibly involved in regulating gene expression [Holliday and Pugh, 1975, Riggs, 1975]. Since its discovery, the methylation at the fifth carbon, more commonly known as 5 methyl cytosine (m^5C) is one of the most well characterised and mechanistically understood epigenetic modifications [Hotchkiss, 1948, Li and Zhang, 2014]. Most of the m^5C in mammals is found at CpG sites. Cytosine residues of CpG dinucleotides can be methylated by *de novo* DNA methyltransferases (DNMTs). Importantly, the methylation can then be inherited mitotically, achieved by the maintenance methyltransferase DNMT1. During mitosis the hemimethylated CpG palindromes are recognised by DNMT3A and DNMT3B, which then replace the methyl group on the unmethylated strand of DNA [Lyko, 2018]. Global removal of *Dnmt1/Dnmt3b* results in embryonic lethality in mice, while loss of *Dnmt3a* results in postnatal lethality [Li et al., 1992, Okano et al., 1999]. Mutations in human DNA methyltransferases are linked to several diseases including cancer and other developmental disorders (for example: [Hansen et al., 1999, Shah and Licht, 2011, Winkelmann et al., 2012, Klein et al., 2011]).

In the mammalian genome about 70%-80% CpG sites are methylated. CpGs in repetitive elements, satellite DNA, and at introns and exons of genes tend to be methylated [Li and Zhang, 2014]. By contrast, CpGs found clustered at regions called CpG islands (CGIs) tend to be unmethylated. CGIs are roughly 1 kb sized

regions and are mainly found at gene promoters [McKeon et al., 1982, Bird et al., 1985, Craig and Bickmore, 1994]. These unmethylated CpG sites are permissive for transcription, although the genes may also be inactive due to silencing by other mechanisms such as histone modifications, described below.

Interestingly, a small number of CGIs do become methylated during development in special circumstances e.g. X chromosome inactivation and genomic imprinting in eutherian mammals [Wolf et al., 1984, Ferguson-Smith et al., 1993], which will be described in greater detail in [later sections](#). Methylated CpG sites can enable or disable the binding of factors that alter the packaging of the DNA. Furthermore, CpG sites can selectively attract specific transcription factors (TFs) that can bind to the underlying DNA sequence. Early studies have shown that removal of Sp1 transcription factor (TF) resulted in *de novo* methylation of the CpG sites [Macleod et al., 1994]. Recently, it was also shown that TFs are sensitive to DNA methylation [Yin et al., 2017]. Several pluripotent factors such as NANOG and POU5F1 (OCT4) preferentially bind to methylated-CpGs, which potentially could play an important role in [epigenetic reprogramming](#) [Yin et al., 2017]. Therefore, TF binding can block methylation or methylation can block TF binding, to contribute to gene silencing.

While m^5C at CpG sites is prevalent in mammals, non-CpG methylation (methylation of cytosines within the dinucleotides CpC, CpA, and CpT) more commonly found in plants and invertebrates is also present in mammals, specifically in certain cell types including embryonic stem cells (ESCs) and brain cells [Lindroth et al., 2001, Riggs, 1975]. Advanced technologies that allowed analysis of methylation of the whole genome at single base resolution showed that non-CpG methylation is far less abundant and less widely distributed compared to CpG methylation [Lister et al., 2009], but is still detectable. The exact function of non-CpG methylation is not known; however, it is shown that non-CpG methylation is associated with gene regulation in the brain and in the nervous system [Jang et al., 2017] .

Even though DNA methylation is mitotically heritable, it can also be removed by either active or passive demethylation mechanisms. Active demethylation occurs

via a multi-step process involving the family of enzymes known as Tet proteins [Tahiliani et al., 2009, Ito et al., 2010]. DNA methylation is processed to intermediates, formyl-cytosine and carboxy-cytosine, each of which can be subject to base excision repair to restore unmethylated cytosine [Wu and Zhang, 2010]. Whereas, cell division without maintenance of methylation is responsible for passive demethylation [Howell et al., 2001]. Particularly during early development the genome undergoes dynamic changes to DNA methylation which will be discussed later in the [epigenetic reprogramming](#) section.

1.2.2 Histone modifications

The long N-terminal flexible tails protruding from histones can be covalently modified. It was also found that residues in the core of the histones can also be modified [Tessarz and Kouzarides, 2014, Tropberger and Schneider, 2013]. Histones carry various post translational modifications in the form of methylation, acetylation, ubiquitylation, phosphorylation, sumoylation and yet more are being discovered [Kouzarides, 2007, Kebede et al., 2015]. In some cases, the histone tail modifications help to recruit protein complexes known as nucleosome remodelers to re-position nucleosomes thereby affecting inter-nucleosomal interactions, accessibility, thereby overall gene expression control [Lawrence et al., 2016].

Two of the best characterized histone modifications are acetyl and methyl group additions [Allfrey et al., 1964]. Acetylation occurs at lysine residues and histone methylation occurs at either arginine or lysine residues [Grunstein, 1997, Greer and Shi, 2012]. Acetylation is universally associated with active transcription, whereas the association of methylation transcriptional state depends on the particular residue that is methylated. Tri-methylation of Histone 3 Lysine 4 (H3K4me3) is associated with a permissive state, while tri-methylation of Histone 3 Lysine 27 (H3K27me3) is associated with a repressive state. These histone modifications are laid down and removed by different enzyme families e.g. histone acetyltransferases (HATs) and histone deacetylases (HDACs), histone lysine methyltransferases (HKMTs) and histone lysine demethylases (KDMs) [Bannister

and Kouzarides, 2011]. Some of the best characterized histone modifiers are the polycomb repressive complexes (PRCs).

1.2.2.1 Polycomb repressive complexes (PRC)

Polycomb group proteins (PcGs) were initially identified as transcriptional repressors of homeotic (*hox*) genes in genetic screens in *Drosophila* where mutants in PcGs failed to show *hox* genes silencing [Pirrotta, 1998, de Ayala Alonso et al., 2007]. Polycomb repressive complex 1 and 2 (PRC1 and PRC2) are complexes of PcGs that catalyse H2AK119ub and H3K27me3 histone marks, respectively. PRC1 complexes are E3 ubiquitin ligases that monoubiquitinate lysine 119 of histone H2A (H2AK119ub1), based on their catalytic components Ring1a or Ring1b [Wang et al., 2004]. PRC2 complexes target H3K27 for mono-, di- and trimethylation [Aranda et al., 2015], based on their catalytic components Ezh2 or Ezh1 [Cao et al., 2002, Kuzmichev et al., 2002].

PRCs bring about epigenetic repression via a number of mechanisms. PRC1 and PRC2 interact with each other's marks: such that H2AK119ub can recruit PRC2 and H3K27me3 can recruit PRC1 to various sites [Schwartz and Pirrotta, 2013]. It is known that H2AK119ub and PRC1 catalytic components interfere with the transcription elongation by RNA polymerase II [Zhou et al., 2008]. PRC1 is also involved in chromatin compaction at inactive *Hox* clusters and absence of PRC1 in cells results in failure of gene silencing [Grau et al., 2011, Eskeland et al., 2010, Kundu et al., 2017, Boyle et al., 2020]. PRCs can work with other epigenetic modifiers such as accessory proteins and long ncRNAs to establish and maintain a repressive heterochromatin environment [Blackledge et al., 2014, Jansz et al., 2018b, Pandey et al., 2008]. While knowledge of enzymatic activities of these groups are still expanding, how exactly these modifications bring about transcriptional repression and how these complexes functionally interact with one another and other epigenetic modifiers to accomplish its mechanism of action is still largely unknown.

1.2.3 Chromatin remodelling proteins

Chromatin remodellers are multi-protein complexes that use ATP to either open up the chromatin or close the chromatin by sliding or evicting nucleosomes [Becker and Workman, 2013]. Different chromatin modifications can attract these complexes to specific regions of the genome, after which the remodellers can modulate the nucleosome positions along the DNA to either allow or repress the transcription factor accessibility [Barozzi et al., 2014]. One such example of a chromatin remodeller is the SWItch/Sucrose Non-Fermentable (SWI/SNF) remodelling complex which functions to “open” the chromatin by recognising acetylated lysine modifications on histones [Hassan et al., 2002]. In contrast, the nucleosome remodeling deacetylase complex (NuRD)/Mi-2, deacetylates histones to “close” the chromatin structure [Denslow and Wade, 2007]. This attests to how chromatin remodellers actively contribute to the dynamic chromatin structure and precise control of gene expression.

1.2.4 Long non-coding RNAs

Non-Coding RNAs (ncRNAs) fall into two main categories: housekeeping ncRNAs and regulatory ncRNAs. Ribosomal, transfer, small nuclear and small nucleolar RNAs that are expressed constitutively fall into the housekeeping ncRNAs category [Ponting et al., 2009]. Regulatory ncRNAs are microRNAs, small interfering RNAs, Piwi-associated RNAs [Carthew and Sontheimer, 2009, Malone and Hannon, 2009] and long non-coding RNAs (long ncRNAs), that are over 200 bp. Long ncRNAs were first identified in early 2000s when sequencing genomes using large cDNA libraries [Okazaki et al., 2002].

Although most long ncRNAs are uncharacterised, some are known to be involved in gene expression control and are also implicated in disease [Uszczynska-Ratajczak et al., 2018, Marchese et al., 2017, Delas and Hannon, 2017]. Some very well studied and rather well characterised long ncRNAs in mammals are *Xist*, that is involved in X chromosome inactivation and *Airn*, *Snrpn*, *Kcnq1ot1*, and *Nespas*,

that are involved in regulation of imprinted gene clusters. These will be [further discussed below](#). It is also important to note that some long ncRNAs can interact with epigenetic modifiers. For example, *Xist* and *Kcnqot1* interact with [PRC2](#), thereby serving to recruit the PcG complex to the targeted genome region [[Pintacuda et al., 2017](#), [Bousard et al., 2019](#), [Hanna et al., 2019](#)]. However, all long ncRNAs can control the gene expression in additional ways. For example the act of transcription itself leading to the long ncRNA being produced can be important (e.g. *Airn*, *Snrpn*) or the transcript can unilaterally act on adjacent genomic regions and can recruit epigenetic modifiers to establish or maintain gene repression [[Mercer et al., 2009](#), [Yao et al., 2019](#)]

1.2.5 Higher order chromatin organisation

Recently, the higher order chromatin organisation or the 3D genome has become a focal point in epigenetic regulation of gene expression. With the fast paced development of advanced sequencing and imaging technologies that capture the chromatin conformation and chromatin profile, what we know about the 3D genome is evolving rapidly [[Lieberman-Aiden et al., 2009](#), [Hughes et al., 2014](#), [Skene and Henikoff, 2017](#), [Kempfer and Pombo, 2020](#), [Rust et al., 2006](#)].

The hierarchical organisation of the genome includes several scales of interactions; short-range chromatin loops, topologically associating domains (TADs), chromosome compartments, and finally chromosome territories inside the nucleus (Figure 1.4) [[Bonev and Cavalli, 2016](#)]. Chromatin loops act to bridge the distance between distal enhancers and promoters and bring them to close proximity allowing their interaction and thereby actuating gene expression [[Schoenfelder and Fraser, 2019](#)]. These interactions are often promoted by proteins such as: YY1 structural protein, CTCF, a structural protein, and cohesin, a ring-shaped DNA entrapping, ATPase dependent complex, that form loops by extrusion [[Weintraub et al., 2017](#), [Rao et al., 2017](#), [Davidson et al., 2019](#)]. Chromatin loops are formed to bring together distal regulatory elements such as enhancers and promoters and facilitate transcription. For example, in the *Igf2/H19* imprinted cluster parental

specific expression of the *Igf2* gene and the *H19* gene, which are separated by approximately 100 kb, is regulated by a shared set of enhancers driven by chromatin [Murrell et al., 2004].

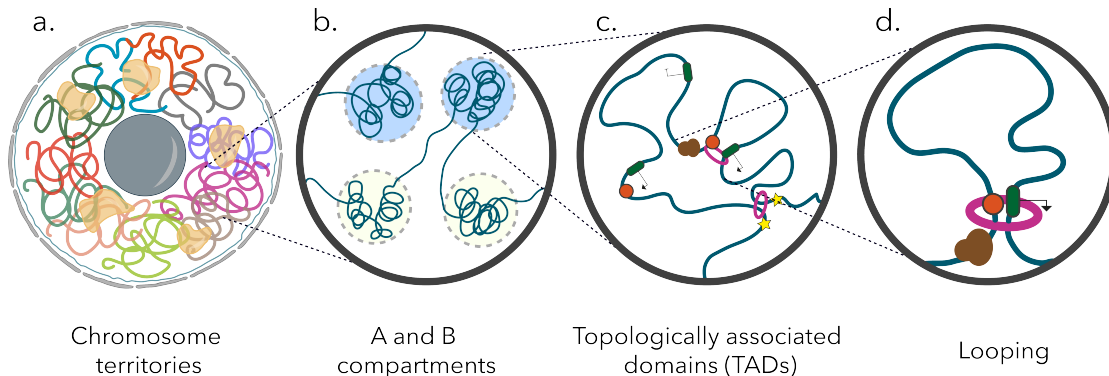


FIGURE 1.4: **The hierarchical organisation of the genome**

a. Each chromosome (depicted in different colours) occupies a distinct territory within the nucleus. Nuclear membrane (dashed circle) nuclear lamina, nuclear speckles, and the nucleolus mark regulatory landscapes within the nucleus. b. Theoretically chromatin is constricted into active "A" and repressed "B" compartments. c. Within these compartments reside topologically associated domains (TADs). TAD boundaries are enriched with CTCF insulator protein depicted as yellow stars. d. The smallest component of the hierarchical organisation, chromatin loops, facilitate long range interactions between promoters (green oblongs) and enhancers (orange circles). Loop formation can be mediated by different factors such as cohesin or PRCs.

Loops tend to be confined within TADs that are 0.2-1 Mb structures. TADs, a genomic region with high frequency of contacts, are separated by boundaries enriched with CTCF insulator protein, repeat elements and other epigenetic modifiers [Dixon et al., 2012, Rowley and Corces, 2018, Jansz et al., 2018b, Szabo et al., 2019]. A prime example for this process is the co-linear activation of *Hox* (*Hox* 1-13 in *HoxA* - *HoxD*) clusters which is important in anterior posterior (a-p) vertebral and limb patterning in vertebrates. Unsurprisingly, the *Hox* clusters adopt genomic architectures in the form of TADs and chromatin compartments for precise *Hox* gene regulation [Noordermeer et al., 2014, Rodriguez-Carballo et al., 2019]. Genes localised within a TAD tend to be co-regulated. Therefore, TADs seem to represent a contained, functionally active, hierarchical unit within the genome. Evidently, disruptions to the structural formation of TADs and its boundaries can lead to gene misexpression that results in development defects and cancer [Franke et al., 2016, Hnisz et al., 2016]. Therefore, how TADs are formed

and what constricts them are vital in elucidating proper gene regulation. While extensive studies are carried out on CTCF that mainly delimits TAD interactions, questions remain as to what and how exactly CTCF-independent TAD boundaries are driven and indeed which other proteins contribute to genome architecture and how this relates to gene expression.

Several TADs together form larger chromosome compartments that are either active (A compartments) or inactive (B compartments). A-compartments are enriched for active genes, whereas B compartments are enriched for repressed genes [Lieberman-Aiden et al., 2009, van Steensel and Belmont, 2017]. While active compartments are found in the middle of the nucleus, associating with nuclear speckles such as splicing speckles or transcription factories, inactive compartments are often found associating with nuclear periphery forming lamin-associated domains [Peric-Hupkes et al., 2010]. It was identified that the function of Lamin B receptor along with Lamin A and C proteins tether the gene poor heterocromatic B compartments to the nuclear periphery [Solovei et al., 2013].

There is speculation that the chromatin compartmentalisation can also occur via phase separation, but this concept is up for debate [Nuebler et al., 2018]. Two distinct mechanisms; Polymer–polymer phase separation (PPPS) and liquid–liquid phase separation (LLPS) have been proposed to explain how this might be achieved [Palikyras and Papantonis, 2019]. PPPS proposes cross linking between chromatin segments that can be destabilised (e.g CTCF, Hp1) [Zirkel et al., 2018, Larson et al., 2017, Strom et al., 2017]. LLPS describes a process where chromatin-associated proteins form a stable liquid like droplet (e.g nucleoli) that exhibits multivalent interactions with each other [Brangwynne et al., 2011, Frottin et al., 2019]. While phase separation is an extremely enticing way to model intricate chromatin organisation many questions including how exactly such phase separation is achieved and how it contributes to transcription dynamics remain to be answered.

Lastly, individual chromosomes occupy a distinct region within the interphase nucleus. The existence of such chromosome territories were brought to light with

chromosome painting and advanced microscopy technologies [Cremer and Cremer, 2010]. Interestingly, it was observed that these chromosome territories can be re-adjusted in disease [Matarazzo et al., 2007]. Exactly how these chromosome territories are established and maintained through dynamic cellular processes is yet to be unraveled.

How all of these interactions are facilitated or restricted may influence gene expression control. While a relationship between genome architecture and gene expression or repression is evident one can speculate whether genome architecture drives transcription or whether it is the gene transcription that drives 3D organisation of the genome. Even with the spike in recent advancement of technologies, we are yet to fully understand the 3D genome and appreciate all the complexities of higher order chromatin architecture.

On the whole, chromatin modifications, structural proteins and RNAs, and 3D genome architecture bring about an epigenetic state that influences gene expression. By diving deeply into how the intricate network of epigenetic regulation forms and functions, it is apparent that epigenetic mechanisms manifest a complex and coordinated symphony to regulate gene expression control. However, we are yet to fully appreciate all the fragile threads that combine to create this intricate network. This knowledge is essential as it will allow us firstly to more comprehensively understand normal development, but also to manipulate the epigenetic mechanisms to tackle disease.

1.3 Epigenetic reprogramming

Epigenetic reprogramming is a fundamental phenomenon that occurs in mammalian development. Just as epigenetic marks are laid on during development these marks need to be removed so that the progeny can lay on new epigenetic marks [Reik et al., 2001, Dean et al., 2003, Hemberger et al., 2009]. Thus, looking at the natural phenomena could give insights into how to manipulate it in somatic cell reprogramming and other applications [Dean et al., 2003, Kim et al., 2010].

The global scale epigenetic reprogramming during development involves DNA methylation, histone modifications, and nucleosome remodeling, that contributes to the establishment of transcriptionally repressive or permissive epigenetic landscapes at different stages of development. Genome wide epigenetic reprogramming occurs in two major waves: first, in the primordial germ cells (PGCs) in embryonic gonads. Second, in the zygote after fertilization. The epigenetic profile of the gametes assumes profoundly different chromatin organization to its somatic counterparts. In the male gamete canonical histones are replaced with transition proteins, histone variants, and protamines forming highly condensed DNA resulting in transcriptional repression [Oswald et al., 2000, Mayer et al., 2000, Sassone-Corsi, 2002, Miller et al., 2010, Jung et al., 2017]. PGCs undergo global losses of several histone modifications (e.g., H3K27me3, H3K9ac) and linker histone H1 during this first wave [Seki et al., 2005, Hajkova et al., 2008]. Again after zygote formation paternal pro-nuclei undergo extensive epigenetic changes once more to replace the non-canonical histones with the canonical histones [Liu et al., 2016]. Conversely, the maternal counterpart remains relatively stable apart from post translational modifications of histone tails that are involved in active gene expression such as arginine methylation and H3/H4 acetylation [Dahl et al., 2016, Liu et al., 2016, Burton and Torres-Padilla, 2010]. All these processes and subsequent cell divisions and differentiation process of the zygote are also facilitated by rapid remodelling of 3D genome that allows or restricts access to different regions of the genome [Gassler et al., 2017, Hug and Vaquerizas, 2018].

Global DNA CpG methylation reprogramming events are extensively studied [Santos et al., 2002, Feng et al., 2010]. It is important to note that the majority of CpG island promoters are not subjected to methylation reprogramming events because they are unmethylated at all relevant stages in development [Edwards et al., 2017]. During the first wave of reprogramming PGCs undergo modifications required for transcriptional repression [Guibert et al., 2012]. Active DNA demethylation co-occurs along with chromatin remodelling and is essential for this process [Popp et al., 2010, Hajkova et al., 2010]. Between embryonic day (E) 10.5 and E13.5 most of the active methylation is globally lost with the exception of intracisternal

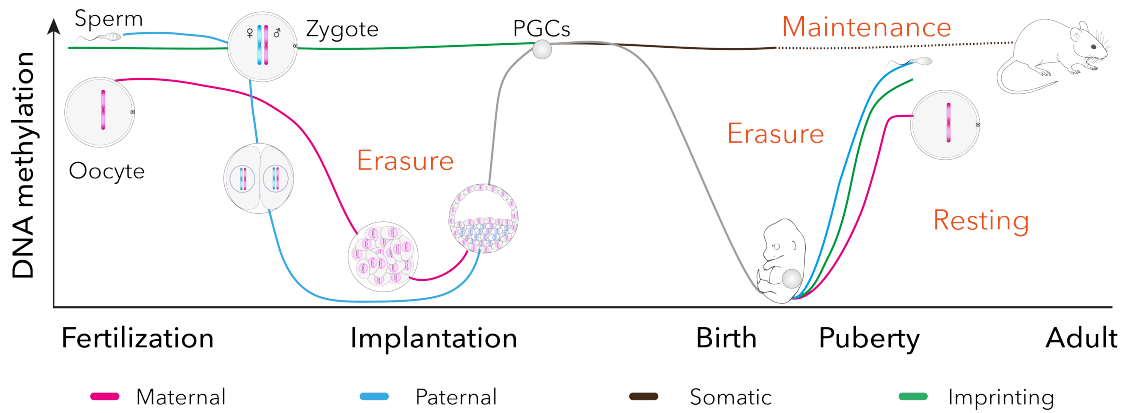


FIGURE 1.5: **DNA methylation reprogramming during development**

Schematic representation of DNA methylation reprogramming during mouse development. DNA demethylation of the zygote occurs passively in the maternal genome (pink), while the paternal genome (blue) undergoes active demethylation. During this time period methylation at imprinted genes (green) remains unchanged. Then the zygote obtains new methylation marks and they are maintained in somatic cells (brown) while these marks are removed in primordial germ cells (PGCs) in the second wave of epigenetic reprogramming

A particles (IAPs) that are protected from epigenetic reprogramming throughout [Lane et al., 2003, Feng et al., 2010]. This global loss of methylation is aided by the expression of *Stella*, *Tet1* and *Tet2* proteins in the PGCs [Hajkova et al., 2010]. The second wave of reprogramming occurs after fertilisation. During which the maternal genome assumes a bimodal distribution of methylation, where methylation is predominately passively lost while there is low level of active methylation [Santos et al., 2002]. The paternal genome which is highly methylated to begin with, undergoes rapid active demethylation (Figure 1.5) [Oswald et al., 2000, Mayer et al., 2000, Santos et al., 2002, Eckersley-Maslin et al., 2018]. Interestingly, specific regions of the genome such as differentially methylated regions in imprinted genes are specifically protected from epigenetic reprogramming during this second wave. This will be discussed in detail in later section.

1.4 Monoallelic gene expression

1.4.1 X chromosome inactivation (XCI)

One of the most remarkable mechanisms of epigenetic silencing is X chromosome inactivation (XCI) in females, which affects more than a thousand genes and is arguably the best example of long range, monoallelic gene silencing in therian mammals [Heard et al., 2004]. In 1949 Barr and Bertram saw the highly condensed sex chromatin body, now referred to as the “Barr body” when they were looking into a nucleus of feline neuronal cells [Barr and Bertram, 1949]. In 1951 Ohno and colleagues showed that the Barr body is actually one of the two X chromosomes in females [Ohno et al., 1959]. Notably, in 1961, Mary Lyon hypothesised that one of the two X chromosomes is inactivated in each cell in female mammals and this occurs as a form of dosage compensation compared to male cells that have only one X chromosome [Lyon, 1961, 1962, 1963]. In the pioneering studies in mice, Lyon hypothesised that XCI is a random phenomenon based on the mice coat colour patterning. Ensuing studies showed that while this process is random in most eutherian (placental) mammals, XCI is imprinted in marsupials and in some tissues in placental mammals such as rodents and cattle [Cooper et al., 1971, Takagi and Sasaki, 1975, Xue et al., 2002, Heard et al., 2004]. Close to half a century later since Lyon’s publication, the X inactivation field has grown exponentially, so much so that we can now infer how XCI manifests and how to manipulate this mechanism to study epigenetics in general.

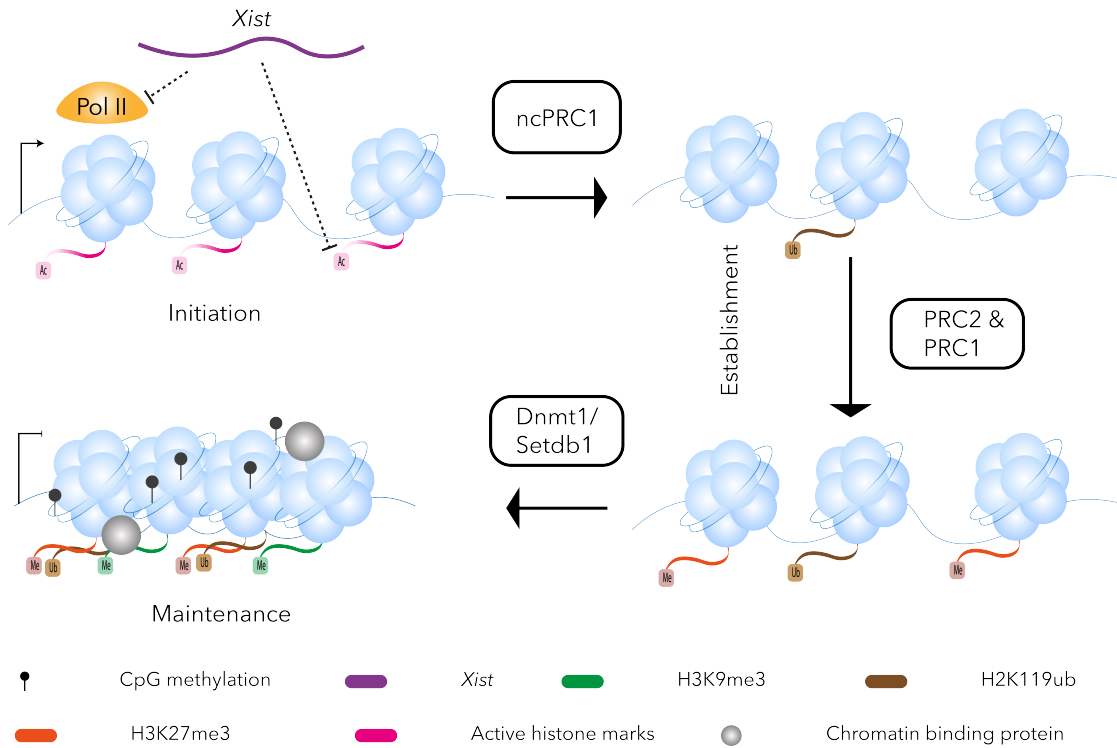
1.4.1.1 Mechanism of XCI

XCI occurs in early development of mammals via progressive accumulation of epigenetic marks to transcriptionally silence one of the two X chromosomes. Initiation of X inactivation is mediated by a X-inactivation specific transcript, (*Xist*) long ncRNA [Brockdorff et al., 1991, Brockdorff, 2019]. As the cells differentiate *Xist* is monoallelically upregulated and exclusively coats the X chromosome that it is

synthesised from [Brockdorff, 2019]. This induces a whole set of mechanisms that follow to dramatically silence one of the X chromosomes in each cell for the life time of the organism (Figure 1.7). Targeted deletion of one of the *Xist* alleles in embryonic stem cells showed that *Xist* is indeed essential for XCI *in cis* [Penny et al., 1996]. Transgene studies of *Xist* localisation to autosomes also demonstrated that *Xist* could indeed inactivate the chromosome *Xist* that is tethered to *in cis* [Lee and Jaenisch, 1997]. This proved that *Xist* is not only essential but also sufficient for XCI. It is suggested that matrix protein hnRNP U is required for chromosomal localisation of *Xist* RNA to the inactive X (Xi) [Hasegawa et al., 2010]. But how exactly *Xist* is upregulated from only one X chromosome needs further clarification.

Tsix, located on the antisense strand to *Xist* is the negative regulator of *Xist* [Lee et al., 1999]. Prior to initiation of X inactivation both *Xist* and *Tsix* are expressed at low levels in pluripotent cells, and the pluripotency factors such as *Sox2*, *Rex1*, *Oct4* also negatively regulate *Xist* transcription [Navarro et al., 2008, Augui et al., 2011]. The early events of XCI are controlled by a small regulatory region on the X known as X-inactivation centre (XIC) [Augui et al., 2011]. This unique region contains genes coding for key regulators of X inactivation including *Rnf12* and long ncRNAs (*Jpx*, *Tsix*, *Ftx*, *Xist*) [Maclary et al., 2013]. Studies based on transgenes, deletions, and chromosomal translocations have demonstrated that the XIC is a region of several mega bases, and XIC alone is imperative and sufficient to regulate the initiation of XCI [Lee et al., 1996, Wutz et al., 2002, Augui et al., 2011].

Prior to the initiation events, X to autosome ratio is counted and a choice is made to determine which X chromosome to inactivate [Ohno, 1967, Penny et al., 1996]. To this date very little is known about the exact process of counting and choice. For example the activation of ubiquitin ligase *Rnf12* in early cleavage embryos degrades the pluripotency factor *Rex1* which results in upregulation of *Xist* [Jonkers et al., 2009, Barakat et al., 2011, Gontan et al., 2012]. As *Rnf12* is expressed exclusively from the X chromosome for a very small window of time, the dosage of *Rnf12* can possibly help in counting number of X chromosomes. However, for the random form of X inactivation *Rnf12* is expendable as deletion of *Rnf12* still leads to X

FIGURE 1.6: **Kinetics of X-inactivation**

Schematic representation of kinetics of XCI mechanism. Upregulation of long ncRNA *Xist* initiates XCI by excluding RNA polymerase II and *active histone marks*. *Xist* is able to set about PRC cascade starting with non-canonical PRC1 which lays the H2AK119ub mark. PRC2 mark H3K27me3 induces canonical PRC1. Subsequently, gene repression is maintained by multiple enzymes and protein complexes including Dnmt1 (addition of methylation mark on chromatin), Setdb1 (adds repressive histone mark H3K9me3) and chromatin binding proteins.

inactivation [Jonkers et al., 2009, Shin et al., 2014]. This suggests that there are other factors involved in the process of counting.

After initiation, *Xist* recruits multiple silencing complexes that lead to the establishment of X inactivation. First *Xist* recruits RNA binding protein, SPEN/SHARP which leads to RNA Polymerase II exclusion and loss of active histone marks [McHugh et al., 2015, Minajigi et al., 2015, McHugh et al., 2015, Chu et al., 2015]. SPEN is recruited to active gene promoters and enhancers and disengages once X-linked genes become silenced [Dossin et al., 2020]. Upregulation of *Xist* also sets about the recruitment of PRC complexes to further silence the Xi. How exactly this process occurs has not always been clear. Originally it was suggested that the A repeat of *Xist* recruits PRC2, and then the PRC2 mark H3K27me3 facilitates PRC1 binding [Cao et al., 2002, Zhao et al., 2008]. But recently it was established

that it is the B-repeat of *Xist* along with *Hnrpk* which recruits non-canonical PRC1 (ncPRC1), which then replaces the active histone marks with H2AK119ub [Almeida et al., 2017, Pintacuda et al., 2017]. This then enables the recruitment of PRC2 and subsequently canonical PRC1 complex [Brockdorff, 2017]. Knock-out studies of ncPRC1 showed female specific embryo lethality [Almeida et al., 2017]. Removal of PRC1 components Ring1A and Ring1B abolished the recruitment of PRC2 complex by *Xist* RNA [Almeida et al., 2017]. Female embryos and cells lacking PRC2 component EED showed complete absence of PRC2 mediated methylation [Kalantry et al., 2006]. This highlights the importance of the sequential recruitment of PRC complexes to mediate establishment of the XCI.

The establishment phase of X inactivation also leads to the recruitment of multiple chromatin binding proteins such as *Smchd1* that leads further silencing of Xi [Jansz et al., 2018b]. Additionally, a member of a repressive protein group known as the HUSH complex adds H3K9me3 along repetitive regions of the X to further establish the Xi [Keniry et al., 2016]. Disruption to the establishment phase of XCI when H2K9me3 is removed suggests that Setdb1 mediated silencing plays an important role in the establishment of silencing possibly via recruiting other chromatin binding proteins or altering chromatin structure.

Finally the maintenance phase of the XCI is demarcated by late replication timing of Xi, accumulation of histone variant macroH2A and importantly accumulation of DNA methylation via *Dnmt1* enzyme [Hansen et al., 1996, Costanzi and Pehrson, 1998, Sado et al., 2000]. Once the X chromosome is silenced, it is maintained through the cell cycle and is mitotically heritable in all somatic cells.

1.4.2 Conformation of Xi

The inactive X chromosome assumes a unique conformation in the 3D space associating with nuclear lamina. It is demonstrated that localisation of the Xi to the periphery along with its association to nuclear lamina might not influence initiation of XCI [Pollex and Heard, 2019]. The nuclear periphery in mammals is

known to be associated with gene repression, therefore inactive chromatin. It is proposed that lamina association leads to the unique chromosome conformation and inactive state of the X [Minajigi et al., 2015, Chen et al., 2016a]. Interestingly, Xi is also shown to be tethered to the nucleolus via *Xist* during the S phase of the cell cycle [Zhang et al., 2007].

Hi-C data shows that upregulation of *Xist* also leads to adaptation of a unique bipartite chromosome conformation of the X with mega domain formation [Nora et al., 2012, Giorgetti et al., 2016]. These megadomains are separated by the conserved Dxx4 locus, nucleolus-associated domains that can help to tether Xi to the nucleolus [Splinter et al., 2011, Rao et al., 2014, Minajigi et al., 2015, Deng et al., 2015]. The Xi exhibits very few TADs compared to the active X or the autosomes [Jegu et al., 2017]. Only the few regions that escape X inactivation form TAD like structures. Even though *Xist* is seemingly dispensable for maintenance of XCI it is necessary for the 3D organisation of the inactive X [Splinter et al., 2011]. This could be due to proteins such as Smchd1 which may tether to Xi via a *Xist* mediated pathway and contribute to its conformation [Jansz et al., 2018b,a, Wang et al., 2018, 2019, Gdula et al., 2019].

Together multiple layers of redundant and non-redundant epigenetic mechanisms establish and maintain the inactive state of one of the female X chromosomes stably for decades of life.

1.4.2.1 XCI in mouse development

XCI is widely studied in the mice, therefore it is vital to understand the X inactivation process in mouse development [Lyon, 1963, Blewitt and Whitelaw, 2013]. XCI during female mouse development shows both imprinted XCI (iXCI) and random XCI (rXCI) (Figure 1.7). During the gametogenesis of the oocyte, *Xist* carries a maternal imprint which is controlled via non-canonical H3K27me3 [Inoue et al., 2017a]. *Xist* is silenced in the early cleavage of the embryo. The embryo undergoes iXCI in the morula stage where the paternal copy is preferentially inactivated [Okamoto and Heard, 2006], because *Xist* is silenced by H3K27me3 on the maternal

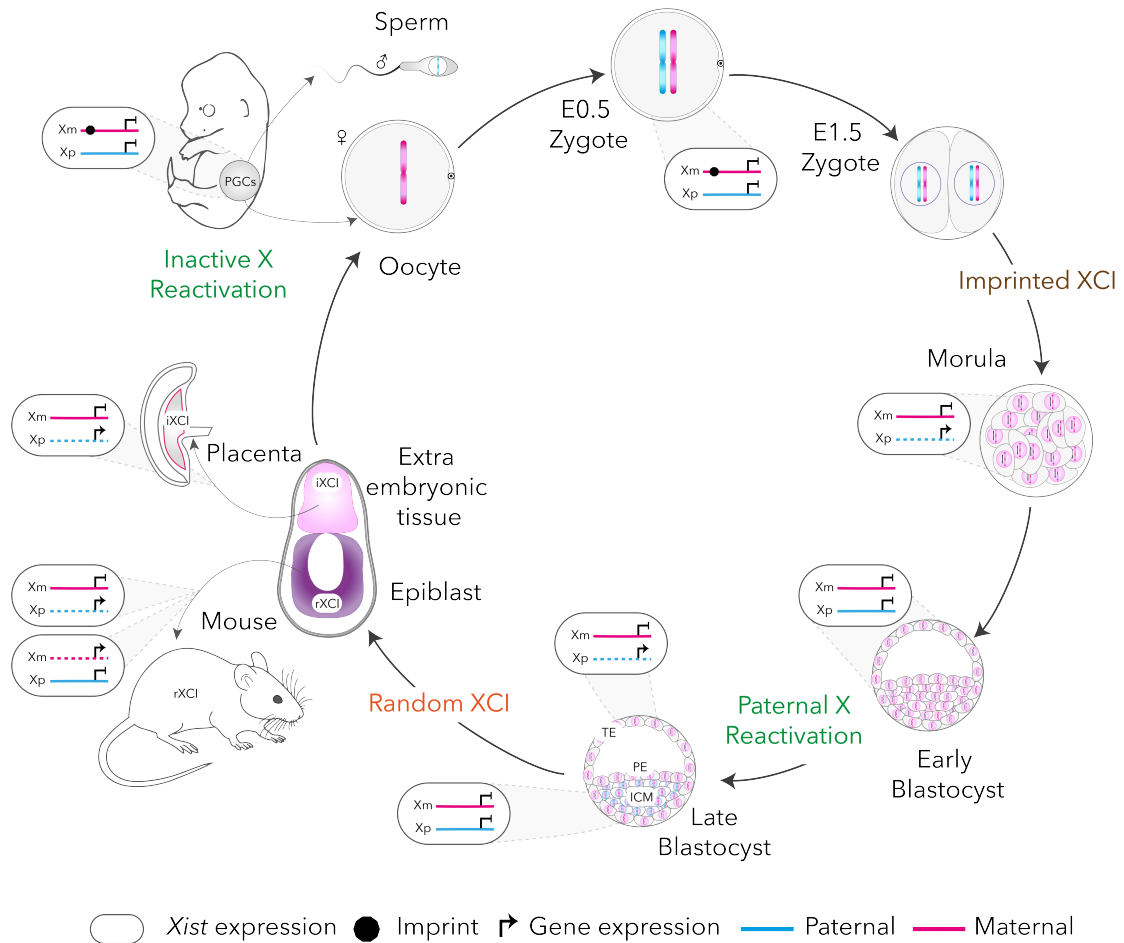


FIGURE 1.7: The ontogeny of XCI in mouse

X chromosome undergoes a cycle of activation and inactivation during mouse development. The maternal copy is shown in pink and the paternal copy in blue. *Xist* activity at different stages is shown inside the box. Both X are active in early embryo but in the morula only paternal X is active as paternal *Xist* is expressed. Both X are active in the inner cell mass (ICM) of late blastocyst which undergoes **random XCI (rXCI)** and give rise to **epiblast**, while trophectoderm (TE) and primitive endoderm (PE) which shows **imprinted XCI (iXCI)** gives rise to the extra embryonic tissue, after gastrulation. **Epiblast** forms the embryo and the **extra embryonic tissue** forms the placenta. Xi is **reactivated** in primordial germ cells (PGCs).

allele. This iXCI is maintained in the trophectoderm and the primitive endoderm in the blastocyst [Sado and Ferguson-Smith, 2005]. The extra embryonic tissue gives rise to the placenta. In the inner cell mass (ICM) of E3.5 blastocyst the paternal X is reactivated [Mak et al., 2004]. The cells that gives rise to epiblast will undergo rXCI. Once established rXCI is maintained in the somatic lineage of the female mice. However, in the germ line Xi is reactivated so that female germ cells have two active X chromosomes [Sado and Ferguson-Smith, 2005, Sugimoto

and Abe, 2007]. The unique occurrence of both iXCI and rXCI in the same life cycle makes the mouse an interesting genetic model to study XCI.

1.4.3 Genomic imprinting

Hallmark studies in 1984 described how parthenogenetic and androgenetic embryos generated by micro-manipulation manifest explicitly opposing phenotypes, demonstrating a strict requirement for both maternal and paternal genomes [Surani et al., 1984, McGrath and Solter, 1984]. Parthenogenetic embryos (two maternal genomes) showed characteristics of relatively normal embryo development until the blastocyst stage, but failure to develop extraembryonic tissue required to develop a functional placenta. While androgenetic embryos (two paternal genomes) showed embryonic lethality, they interestingly showed strong contribution to extra embryonic tissue, forming large placentae [Barton et al., 1984, Thomson and Solter, 1988]. The essence of this difference is the presence of parental specific markings on the genome laid during gamete development which are vital for differential expression responsible for normal development [Surani et al., 1990]. This phenomenon of monoallelic gene expression based on parent of origin is known as “genomic imprinting” [Ferguson-Smith, 2011].

Soon after the pioneering studies on isoparental gametes, another group was able to successfully produce a series of uniparental disomies (UPDs) using chromosomal translocations, which contained two copies of a particular gene from one parent but not the other [Cattanach and Kirk, 1985]. They identified 8 autosomal regions in which disomies caused a number of phenotypes in the progeny including embryonic and neonatal lethality, growth retardation, and behavioral abnormalities [Cattanach and Kirk, 1985, Cattanach and Beechey, 1990]. In the early 1990s another landmark study successfully identified the very first imprinted genes that were indeed differentially expressed between the parental alleles, *Igf2* (paternally expressed) and *Igf2r* and *H19* (maternally expressed) [Bartolomei et al., 1991,

Ferguson-Smith et al., 1991, DeChiara et al., 1991, Barlow et al., 1991]. Subsequent research identified that most imprinted genes are found in clusters [Ferguson-Smith, 2011]. Interestingly, imprinted expression could either be tissue specific or stage specific. And perhaps more importantly these imprinted genes mainly control development and neurological functions [Ferguson-Smith, 2011, Barlow and Bartolomei, 2014].

There are over 200 imprinted genes identified to date [Gigante et al., 2019]. Imprinted gene regions are rich in CGIs, with the presence of clustered direct repeats, and at least one long ncRNA expressed [Neumann et al., 1995, Paulsen et al., 2000]. Mechanistically imprinted gene clusters are regulated by imprinting control regions (ICRs). ICRs are commonly differentially methylated regions (DMRs) between genomes. DNA methylation was the first identified key regulator of genomic imprinting [Reik et al., 1987, Sapienza et al., 1987, Swain et al., 1987]. Based on their establishment, DMRs are classified into either primary/germline DMRs where DNA methylation is established in the gametogenesis or secondary/somatic DMRs in which the methylation is acquired after fertilisation [Barlow and Bartolomei, 2014]. Canonical ICRs can be effectively categorised to two different types: the ones that function as insulators (e.g *H19/Igf2* cluster) and those that serve as a promoter for a long ncRNA that regulates the imprinted cluster (e.g *Igf2r/Airn*, *Snrpn* and *Kcnqot1* cluster) (Figure 1.6). Truncating the mouse long ncRNAs regulating these clusters results in loss imprinted expression in a tissue specific manner [Sleutels et al., 2002, Mancini-Dinardo et al., 2006], demonstrating function of long ncRNAs at imprinted clusters.

More than three decades after the identification of DMRs, recent work points towards a novel non-canonical mediator for genomic imprinting; H3K27me3 [Inoue et al., 2017b, Hanna and Kelsey, 2017]. The non-canonical imprinted regulation appears to be transient where the imprinted loci gradually lose allelic H3K27me3 by blastocyst stage of the embryo development [Zheng et al., 2016]. As understanding of non-canonical imprint regulation is still in its infancy, many outstanding questions such as the exact mechanism of H3K27me3 imprint regulation, whether

it has a role in establishing secondary DMRs, or whether other PRCs and accessory proteins also play a role in imprint regulation, remain to be answered.

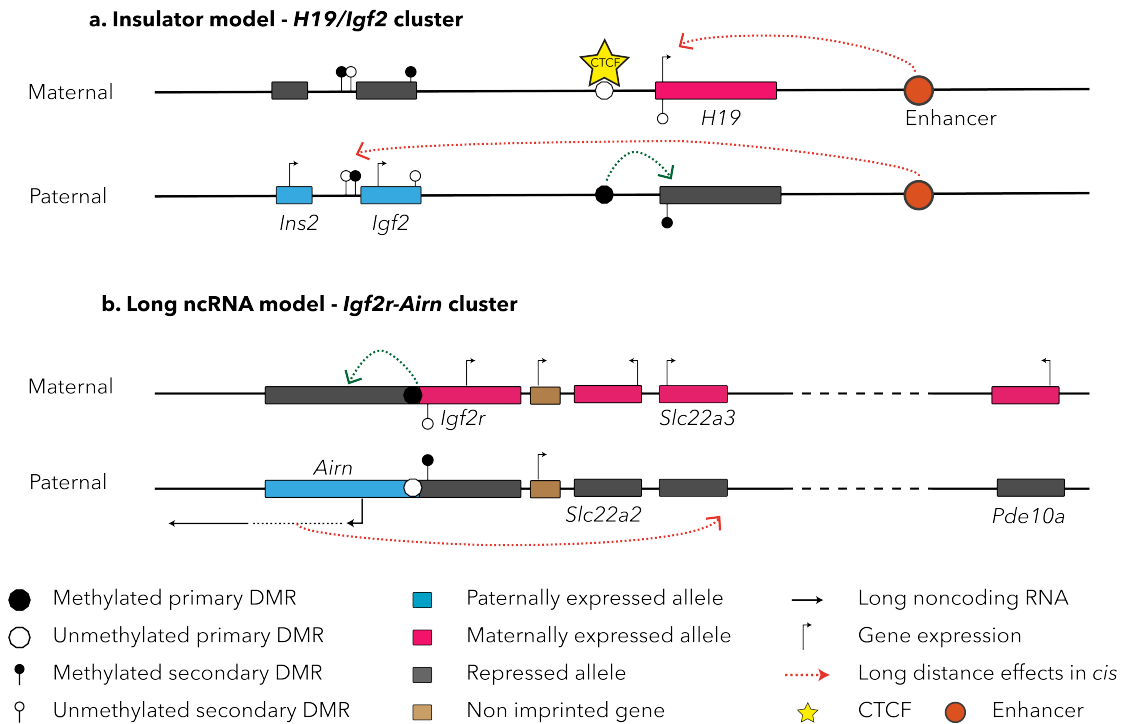


FIGURE 1.8: **Canonical imprinted gene regulation**

a. Insulator mediated regulation of *H19/Igf2* cluster in the endoderm. On the maternal strand unmethylated primary DMR binds the insulator protein CTCF which prevent the enhancer activity on *Igf2* and *Ins2*. Instead enhancer acts on *H19* long ncRNA. On the paternal strand insulator protein cannot bind to the primary DMR as it is methylated. *H19* is not expressed due to methylation. Therefore, the distal enhancer can now act on *Igf2* and *Ins2* genes. b. Long ncRNA mediated regulation *Igf2r/Airn* cluster in the placenta. The primary DMR is hypermethylated on the maternal strand, preventing transcription of *Airn* and allowing *Igf2r*, *Slc22a2* and *Slc22a3* to be transcribed. Conversely, in the paternal strand the DMR is unmethylated, allowing the expression of *Airn* long ncRNA and surrounding genes *Igf2r*, *Slc22a2*, *Slc22a3* and *Pde10a* are silenced.

1.4.3.1 Parental conflict theory

Several explanations were put forth to explain the evolution of imprinting. The most widely accepted explanation is the "parental conflict theory" [Moore and Mills, 2008]. The conflict theory was proposed to explain how gene regulation is controlled to favour the competing interests of either parent. Paternally expressed genes (e.g paternally expressed growth promoter gene *Igf2*) are primed to

extract more resources from the mother to benefit fetal growth and fitness. In contrast, maternally expressed genes (e.g. maternally expressed growth suppressor gene *Igf2r*) act to maximize the reproductive performance of the female to benefit future offspring by sparing the resources and suppressing embryonic growth. Therefore, the parental genes are essentially at a tug of war to allocate the limited maternal resources [Moore, 1991].

1.4.3.2 Imprinting disorders

The early landmark studies on uniparental inheritance of imprinted loci in mice showed phenotypic characteristics such as embryonic lethality, growth, development, and behavioural issues that indicated imprinting could be involved in human disease. Not so long after that the very first imprinting defects in humans, Prader-Willi syndrome (PWS) and Beckwith Wiedemann syndrome (BWS) were discovered [Nicholls et al., 1989, Ping et al., 1989, Koufos et al., 1989]. The genetic cause of these disorders is now understood (Table 1.1). Deregulation of imprinting can cause many congenital disorders affecting neurological and psychiatric function, metabolism, and in some cases increased cancer risk [Peters, 2014]. These congenital defects carry common clinical features such as growth defects, intellectual disability, hormonal malfunctions, and metabolic defects. For example, PWS is caused by loss of imprinted expression at *Snrpn* cluster genes; *Snrpn*, *MAGEL2*, *MKRN3* and *NDN* genes [Adam et al., 1993-2020]. Angelman syndrome (AS) is caused by loss of imprinted expression of *UBE3A* of the same cluster [Adam et al., 1993-2020].

Molecular defects underlying imprinting disorders are comprised of UPDs, aberrant methylation (epimutation) chromosomal imbalances, and genomic mutations in imprinted genes [Court et al., 2013, Monk et al., 2019]. UPD is a common occurrence in imprinting disorders, which occurs when the offspring inherits both copies for the respective imprinted loci from a single parent. Epimutations are changes that happens to DMRs either in the form of loss of methylation (LOM) or gain of methylation (GOM) without altering the underlying DNA sequence.

TABLE 1.1: Imprinting disorders

Imprinted disorder [OMIM]	Imprinted cluster/ gene	Human (mouse) Locus	Molecular defect(frequency)
Prader-Willi syndrome (PWS) [176270]	<i>SNRPN</i> cluster <i>SNRPN</i> , <i>MAGEL2</i> , <i>MKRN3</i> , <i>NDN</i> genes	15q11.2 (7qB5)	15q11–q13 pat deletion (75-80%) UPD(15)mat (20-25%) SNURF:TSS-DMR GOM (1% MLID: rare)
Angelman syndrome (AS) [105830]	<i>SNRPN</i> cluster <i>UBE3A</i> gene	15q11.2 (7qB5)	Mat deletion (70-75%) UPD(15)pat (3-7%) SNURF:TSS-DMR LOM (2- 3%, MLID: rare) <i>UBE3A</i> point mutations (10%)
Schaaf–Yang syndrome (SYS) [615547]	<i>SNRPN</i> cluster, <i>MAGEL2</i> gene	15q11.2 (7qB5)	<i>MAGEL2</i> point mutations (100%)
Beckwith Wiedemann syndrome (BWS) [130650]	<i>H19/IGF2</i> cluster, <i>CDKN1C</i> , <i>KCNQ1OT1</i> genes	11p15.4 (7qF5)	UPD(11p15)pat (20%) 11p15 CNVs (2–4%) H19–IGF2:IG:DMR GOM (5%) KCNQ1OT1:TSS-DMR LOM (50%, MLID: 25%) <i>CDKN1C</i> point mutations (5%)
Silver-Russell syndrome (SRS) [180860]	<i>H19/IGF2</i> cluster, <i>CDKN1C</i> , <i>KCNQ1OT1</i> genes	7(6qA3.3), 11p15(7qF5), 12q14, 8q12	UPD(7)mat (5–10%, MLID:rare) UPD(11p15)mat (rare) 11p15 CNVs (<1%) H19–IGF2:IG-DMR LOM (30–60%, MLID: 7–10%) <i>CDKN1C</i> , <i>IGF2</i> , <i>HMGA2</i> , <i>PLAG1</i> point mutations (rare)
Birk–Barel syndrome [612292]	<i>PEG-13</i> cluster, <i>KCNK9</i> gene	8q24.3 (15D3)	<i>KCNK9</i> point mutations (100%)
Transient neonatal diabetes mellitus (TNDM) [601410]	<i>ZAC</i> (<i>PLAGL1</i>) gene, <i>KCNK9</i> gene	6q24.2 (10qA2)	UPD(6)pat (41%) Pat duplications (29%) PLAGL1:TSS-DMR LOM (30%)

OMIM:Online Mendelian Inheritance in Man. UPD: uniparental disomy. mat: maternal, pat: paternal. GOM: Gain of methylation. LOM: Loss of methylation. CNV: Copy number variation. MLID: Multilocus methylation imprinting disturbances. Information for the table is collated from [Monk et al., 2019, Adam et al., 1993-2020]

Chromosomal imbalances occur in the form of deletions, translocations, and rearrangements that can result in copy number variations (CNV), causing repression or over-expression of imprinted loci completely. Point mutations occur in imprinted genes resulting in a number of imprinting disorders as listed in (Table 1.1). In most

imprinting disorders only one imprinted locus is affected, but there are a number of patients showing disturbed methylation at multiple imprinted loci, otherwise known as multilocus imprinting disturbances (MLID) [Monk et al., 2019].

1.5 Maternal effect proteins

Early studies by developmental biologists and geneticists revealed that the oocyte content plays a significant role in the onset of development. This effect was first described in 1898, when embryologists first noticed that the enucleated sea urchin embryo was able to undergo the first stages of development solely depending on cytoplasmic material of the oocyte [Ziegler, 1898]. Maternal effect genes were first described in fruit fly development [Redfield, 1926]. Maternal effect genes can be defined as “Genes whose products, RNA or protein, are produced or deposited in the oocyte or are present in the fertilized egg or embryo before expression of zygotic genes is initiated” [Marlow, 2010]. Subsequently, a wide array of maternal effect proteins and genes that are important in development including but not limited to, oocyte development, fertilisation, maternal to zygotic transition, and imprinted gene regulation were discovered.

1.5.1 Maternal effect proteins in imprinted gene regulation

Imprinted genes are protected during the second wave of [epigenetic reprogramming](#) after fertilisation event. Multiple studies have demonstrated that it is the maternal effect proteins that play this vital role of shielding imprinted genes in preimplantation embryos. The handful of maternal proteins that act to retain parent-of-origin specific marks include: Dnmt1o, Dnmt3L, Stella (DPPA3), Zfp57, Rlim, Trim28, Nlrp2, and Eed [Howell et al., 2001, Bourc’his et al., 2001, Payer et al., 2003, Li et al., 2008, Shin et al., 2010, Messerschmidt et al., 2012, Mahadevan et al., 2017, Inoue et al., 2017a]. All of these genes except *Eed* were known to regulate imprinting by maintaining canonical DNA methylation imprints. However, the recent work by Inoue *et al.* as [previously described](#) expanded our understanding on

imprinted gene regulation [Inoue et al., 2017b,a]. Therefore, Maternal effect genes lay down and protect canonical and non-canonical imprints during gametogenesis and preimplantation development. Failure to do so can result in irregular imprinted gene expression that can be detrimental in embryonic development. It can also lead to a subset of human disorders known as MLID (reviewed in [Sanchez-Delgado et al., 2016]). MLID are well characterized for maternal effect proteins ZFP57 and the NLRP family proteins [Docherty et al., 2015, Fontana et al., 2018]. Consequently, identification and characterization of Maternal effect genes are crucial to understand the early development process and to manipulate reproduction and development.

1.6 Structural Maintenance of Chromosome Hinge Domain containing 1 (Smchd1)

1.6.1 Discovery of the epigenetic regulator Smchd1

Smchd1 was first identified using a forward genetic screen to discover novel modifiers of murine metastable epialleles (*Mommes*) (Figure 1.9) [Blewitt et al., 2005]. This experiment was performed using a mouse strain containing a multicopy GFP transgene integrated into chromosome 1. The homozygous GFP transgene reporter is under the control of an α globin promoter and enhancer which show GFP expression only in 55% in erythrocytes [Preis et al., 2003]. When these mice were subjected to *N*-ethyl-*N*-nitrosourea (ENU) mutagenesis, offspring were assessed for any change in percentage of GFP expressing erythrocytes i.e the rate of variegation. An increase in GFP expression suggested loss of function mutation of a transcriptional repressor, whereas, a decrease suggested a loss of function mutation in an activator. Multiple strains were generated that showed dominant heritable mutations known as *Mommes*. Most of the *Mommes* identified in this forward genetic screen were epigenetic regulators that were known to play significant roles in gene expression control involved in development [Ashe et al., 2008, Daxinger et al.,

2013, Blewitt and Whitelaw, 2013]. These include the DNA methyltransferases (Dnmt1, Dnmt3b), histone deacetylase and methyltransferases (Hdac1, Setdb1), chromatin-remodelling factors (members of Smarcc family) and the transcriptional regulators (Klf1, Trim28) [Chong et al., 2007, Ashe et al., 2008, Whitelaw et al., 2010, Blewitt and Whitelaw, 2013, Youngson et al., 2013, Sorolla et al., 2015]. But most notably, this screen was able to identify novel players of epigenetic control such as Rlf, TASOR (D14Abb1e) and Smchd1 [Daxinger et al., 2013, Harten et al., 2014, Blewitt et al., 2008].

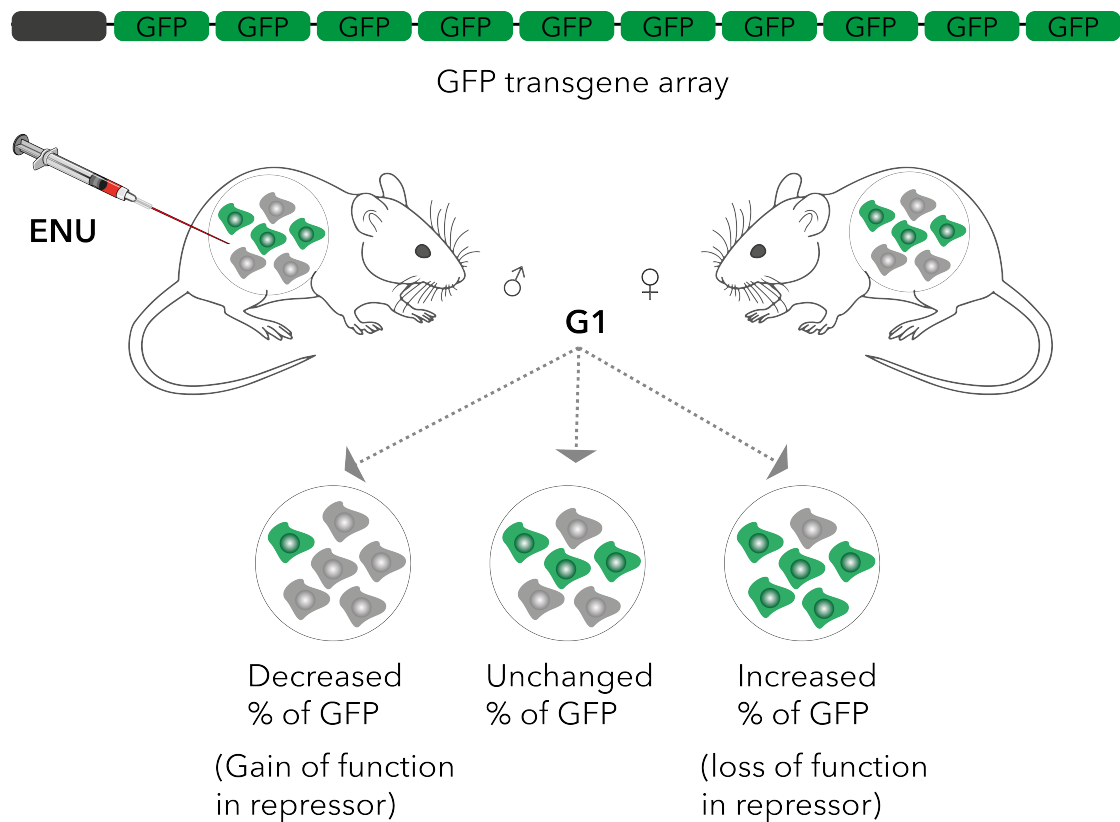


FIGURE 1.9: **ENU mutagenesis screen for modifiers of transgene silencing**

Mice containing multi copy **GFP transgene** expressing GFP in 55% of erythrocytes were subjected to ENU mutagenesis. Following random mutagenesis, generation 1 (G1) showed variegated GFP expression. Increased GFP expression suggests transcription repression and vice versa.

Smchd1 was initially known as modifier of murine metastable epialleles dominant 1 (MommeD1). ENU mutagenesis introduced a nonsense mutation that induced nonsense-mediated decay of the Smchd1 transcript [Blewitt et al., 2005, 2008]. This initial discovery of Smchd1 led to the discovery of its role in X inactivation,

autosomal gene silencing and its importance in development and disease which will be discussed in detail in this section.

1.6.2 Smchd1 in the XCI

When the *MommeD1* mutation was first discovered the progeny showed a notable absence of homozygous mutant females after mid gestation [Blewitt et al., 2005]. Upon further examination it was discovered that *MommeD1* homozygous mutants showed hypomethylation of X-linked genes, as well as failure in both rXCI and iXCI [Blewitt et al., 2008]. *Smchd1* was predominately found in the nucleus and was colocalised to the inactive X chromosome marked by H3K27me3 and other inactive X markers such as H2AK119ub in both mice and humans [Blewitt et al., 2008, Jansz et al., 2018b, Nozawa et al., 2013]. Together this suggested that *Smchd1* may play a role in XCI.

The ability for the female embryo to survive until mid gestation suggested the initiation of X inactivation occurred normally, suggesting *Smchd1* plays an important role in either establishment or maintenance of XCI [Blewitt et al., 2008]. For a while, it was not very clear whether *Smchd1* was important in establishment or the maintenance of XCI or both. Subsequent research attempted to shed a light on these events. *In vitro* experiments in fully differentiated cells (mouse embryonic fibroblasts - MEFs), carrying a GFP transgene in the Xi chromosome showed reactivation of the GFP allele upon *Smchd1* depletion [Keniry et al., 2016], suggesting failure to maintain X-inactivation. However, this experiment was based on transgene activation, which was below 1% and may not reflect endogenous gene behaviour. *In vivo*, *Smchd1* null mice displayed upregulation of a subset of X-linked genes [Blewitt et al., 2008, Gendrel et al., 2012, Mould et al., 2013]. Curiously, this was associated with hypomethylation of CGIs on the Xi [Blewitt et al., 2005, 2008, Gendrel et al., 2012]. As deposition of DNA methylation is the last phase of XCI, this indicates *Smchd1*'s role precedes DNA methylation events.

Recent studies suggest that although *Smchd1* may not directly interact with *Xist*, it localises to the inactive X in an *Xist* dependent manner [Jansz et al., 2018b, Wang et al., 2019]. The B repeat region of *Xist* plays an important role in *Smchd1* recruitment along with *HnrnpK* (the heterogenous nuclear ribonucleoprotein K), a direct interactor of *Xist* [Jansz et al., 2018b]. This was beautifully demonstrated in imaging experiments carried out in *HnrnpK* knockdown cells lines that showed loss of *Smchd1* localisation in the absence of HnrnpK. But no direct interactions between *Smchd1* and HnrnpK were reported, suggesting there could be other factors involved to tether *Smchd1* to the inactive X chromosome. Previously, it was shown that HnrnpK recruits non-canonical PRC1, which imparts H2AK119ub [Almeida et al., 2017, Pintacuda et al., 2017]. Consequently, removal of PRC1 catalytic subunits Ring1A and Ring1B resulted in the loss of *Smchd1* localisation to Xi. This effect was rescued using Ring1B over expression, but not by catalytically inactive Ring1B [Jansz et al., 2018b]. *Smchd1* is retained on the Xi after establishment of XCI suggesting that it could play a potential role in maintenance of the XCI. Therefore, it will be important to consider how it works and behaves at this point of time.

1.6.3 *Smchd1* in imprinted gene regulation

Just like XCI, multiple studies involving gene expression data demonstrated that *Smchd1* plays a significant role in clustered, monoallelic gene silencing [Gendrel et al., 2013, Mould et al., 2013, Chen et al., 2015, Jansz et al., 2018a]. These include clustered protocadherins (*Pcdh*), (*Hox*) genes and imprinted genes, demonstrating *Smchd1* as a key player in global gene regulation in both male and female mammals.

Loss of *Smchd1* results in loss of imprinting at *Snrpn* and *Igf2r* clusters [Gendrel et al., 2013, Mould et al., 2013]. Both these clusters are canonically regulated via DMRs as discussed above. While, the data in these studies mainly comprises of microarray data limited to very few SNPs, they provided insights into methylation patterns in the absence of *Smchd1*. They demonstrated that even though the

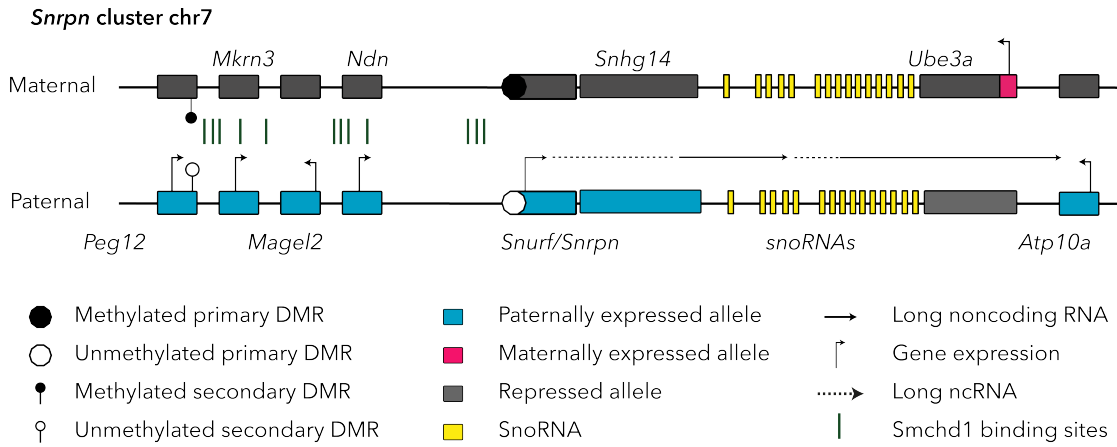


FIGURE 1.10: **Smchd1** regulates genes at **Snrpn** cluster

Snrpn cluster genes are mainly paternally expressed except *Ube3a* which is maternally expressed. *Snrpn* cluster is regulated by a gDMR at the *Snrpn* maternal allele, a sDMR at *Peg12* and a Long ncRNA. *Smchd1* binds only at one end of *Snrpn* cluster which contain *Peg12*, *Mkrn3*, *Magel2* and *Ndn* genes.

secondary DMR was hypomethylated the primary DMR was not affected by loss of *Smchd1* [Mould et al., 2013, Gendrel et al., 2013]. These data suggested *Smchd1* may play a role downstream of methylation of germline ICRs. Additionally, ChIP-Seq data showed that *Smchd1* directly binds to these imprinted clusters [Chen et al., 2015, Jansz et al., 2018a].

Interestingly, data thus far suggests that *Smchd1* acts only on one side of the *Snrpn* cluster involving *Peg12*, *Mkrn3*, *Magel2* and *Ndn* [Gendrel et al., 2013, Mould et al., 2013] (Figure 1.10), and only at *Slc22a3* in the *Igf2r* cluster [Mould et al., 2013]. The canonical imprinting regulation does not fully explain this partial mechanism of action imparted by *Smchd1*. Therefore it will be interesting to explore whether *Smchd1* plays a similar role as suggested for XCI and *Hox* gene regulation via chromatin organisation at imprinted clusters as well [Jansz et al., 2018a]. It is important to note that the genes regulated by *Smchd1*: *Peg12*, *Mkrn3*, *Magel2* and *Ndn* are implicated in imprinting disorders, Prader-Willi syndrome and Schaaf-Yang syndrome, highlighting the importance of elucidating *Smchd1*'s function to better understand these imprinted diseases.

1.6.4 Smchd1 in higher order chromatin structure

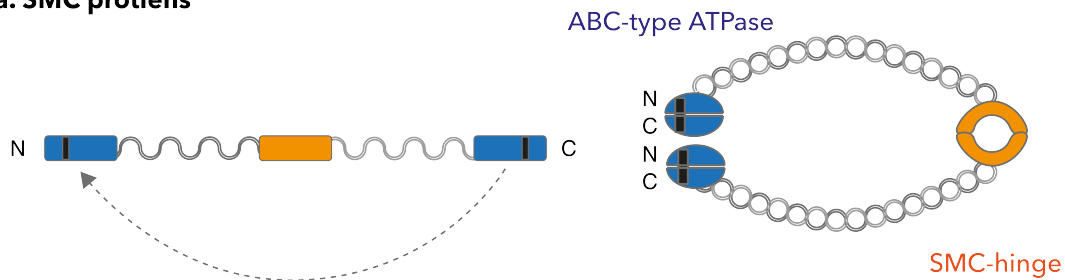
There was an eruption of data by multiple groups on Smchd1's role in higher order chromatin conformation demonstrating how Smchd1 exists in 3D genome space and how it is exerting its function at its binding sites [Jansz et al., 2018a, Wang et al., 2018, 2019, Gdula et al., 2019]. It is suggested that Smchd1 can inhibit enhancer promoter interactions mediated by chromatin loops, by encouraging long range interactions between genomic regions that are repetitive in nature [Wang et al., 2018, Jansz et al., 2018a]. Jansz *et al.* made use of HiC data generated using neural stem cells (NSCs) where Smchd1 was deleted *in vitro* after silencing. What this study was also able to show was how the loss of Smchd1 alters chromatin at the inactive X, without causing reactivation of the Xi, further establishing that while Smchd1 may play significant role in maintaining chromatin conformation it is less essential to maintain XCI. Another study also proposed that Smchd1 may mediate interactions between two chromatin compartments, labelled S1 and S2, in the Xi and form a compartment-less architecture in the absence of Smchd1 [Wang et al., 2018, 2019]. This further solidifies that Smchd1 plays a crucial role in establishing XCI and forms an epigenetic memory via establishing and maintaining chromatin architecture. Jansz *et al.* also characterised a new role for Smchd1 in silencing of Hox gene clusters. It will be interesting to resolve precisely the role Smchd1 plays in chromatin architecture at these autosomal loci in fine resolution.

Recently, using high resolution imaging “the loop extrusion model” for cohesin and condensin, [structurally similar](#) chromatin modifiers to Smchd1, was validated [Ganji et al., 2018, Davidson et al., 2019]. Loop extrusion process describes condensin's ability to feed DNA through its central port using ATPase activity. It will be fascinating to look at Smchd1 for similar capability *in vivo* when altering chromatin architecture to exert its function.

1.6.5 Smchd1 protein structure

Smchd1 is a large chromatin binding protein of 230 kDa that is evolutionary conserved amongst vertebrates. Smchd1 contains a N-terminal GHKL-type ATPase domain, a C-terminal SMC (structural maintenance of chromosomes) hinge domain and a large uncharacterised region in the middle (Figure 1.11) [Chen et al., 2016b,c, Pedersen et al., 2019, Chen et al., 2020]. Well known SMC family proteins include cohesin and condensin which are known to tether DNA to modulate chromatin conformation and regulate chromosome dynamics through cell cycle (reviewed in [van Ruiten and Rowland, 2018]). Structurally, SMC family proteins heterodimerise and facilitate DNA binding [Lammens et al., 2004]. While the hinge domain of Smchd1 binds to DNA, it homodimerises [Chen et al., 2015, 2016c]. The recently unraveled hinge domain structure of Smchd1 not only beautifully demonstrated this function but also gave finer details on which residues are important to elucidate its function and how it differs from canonical SMC protein hinge domains [Chen et al., 2020].

a. SMC proteins



b. Smchd1

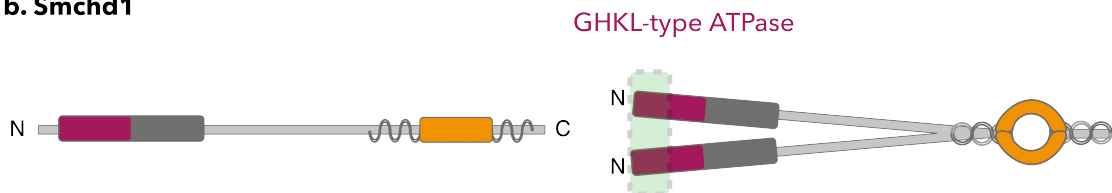


FIGURE 1.11: **Smchd1 is Structurally similar to SMC family proteins**

a. SMC family proteins contain **ABC type ATPase** domain and **hinge domain** and forms heterodimers. b. Smchd1 contains a **hinge domain** similar to SMC family that heterodimerise but also form **homodimers** at **GHKL type ATPase** domain.

The N terminal of Smchd1 contains a GHKL (Gyrase, Hsp90, histidine kinase, MutL) ATPases. Members of GHKL family proteins consist of an ATP-binding

fold required for its activity [Dutta and Inouye, 2000]. A recently presented crystal structure of the ATPase domain of Smchd1 shows that not only does it contain this ATP binding fold but it is stabilised by an N-terminal ubiquitin-like fold [Pedersen et al., 2019]. How exactly this fold plays a role in Smchd1 binding and dynamics is yet to be understood. It is also important to note that dimerisation of the ATPase domain was proposed using catalytically dead E147A mutant [Pedersen et al., 2019]. Therefore whether the wild type ATPase domain of Smchd1 dimerises and whether this contributes to the dynamics of the protein is not yet apparent. In fact, we are yet to elucidate the structure of the large uncharacterised middle region of Smchd1. The only data currently available for the full length structure is negative stain electron microscopy images showing Smchd1 to form dumbbell-like rods [Brideau et al., 2015]. Whether this structure is true in cells is yet to be validated. It is important to understand whether the mechanism of action proposed for the purified protein replicates *in vivo*.

1.6.6 Smchd1 mutations are implicated in disease

Thus far it is evident that Smchd1 plays a critical role in development. Clinically, SMCHD1 is known to be important in two very distinctive disorders in humans. Heterozygous loss of function mutations of Smchd1 cause a form of muscular dystrophy known as Facioscapulohumeral Muscular Dystrophy (FSHD) while controversially gain of function mutations cause a rare congenital disorder, Bosma Arhinia and Microphthalmia Syndrome (BAMS) [Lemmers et al., 2015, Gordon et al., 2017, Shaw et al., 2017].

FSHD is the third most common form of muscular dystrophy, which affects approximately 1 in 8000 people worldwide [Deenen et al., 2014]. FSHD is generally late onset and shows characteristic symptoms of weakening and atrophy of upper body muscles including the face and neck that eventually progresses to the rest of the skeletal muscles [Padberg et al., 1991]. Based on the underlying genetic mutation FSHD is classified into two different types: FSHD1 and FSHD2.

FSHD1 is the more prevalent form resulting in 95% of cases reported and FSHD2 patients account for the remaining 5% of cases reported [Hamel and Tawil, 2018]. Genetically, the relaxation of $D4Z4$ tandem repeat unit on chromosome 4q35 is the main causative of FSHD [Nguyen et al., 2017]. $D4Z4$ regulates the expression of $DUX4$ mRNA. Healthy individuals carry 11 to 100 $D4Z4$ repeats, capable of synthesising $DUX4$ mRNA, which in turn synthesise DUX4 in the skeletal muscle. Whereas patients showing FSHD1 have 1-10 such repeat arrays. This is commonly known as “contraction” of the $D4Z4$ sequence [Sacconi et al., 2015]. Given that $D4Z4$ is a repeat array, the region is heterochromatic in nature due to repeat-induced gene silencing. But when it undergoes “contraction” it leads to “chromosome relaxation” due to the loss of epigenetic signature, in the form of hypomethylation and loss of repressive epigenetic marks [Nguyen et al., 2017]. FSHD patients also shows permissive polyadenylation (poly-A) signal (4qA) of the $DUX4$ transcript. Poly-A tail stabilises the DUX4 transcript and this ectopic expression results in myotoxicity in the FSHD patients.

On the other hand FSHD2 patients exhibit digenic ontogeny of the disease. Approximately 80% of FSHD2 patients show heterozygous loss of function mutations in SMCHD1 [Sacconi et al., 2013, 2015]. The variety of SMCHD1 loss of function mutations including nonsense mutations, frameshift mutations, and deletions are observed in FSHD patients [Gurzau et al., 2018]. These mutations are also seen through out the length of SMCHD1 protein but are least prevalent in the DNA-binding hinge domain [Jansz et al., 2017, Gurzau et al., 2018]. Conspicuously, patients showing both SMCHD1 mutations and FSHD1 characteristic “contracted” $D4Z4$ arrays show more severe phenotype [Sacconi et al., 2013], highlighting the importance of SMCHD1 for all patients.

Considering the rare nature of BAMS onset, it is far less studied compared to FSHD. Currently less than 100 BAMS patients are reported world wide [Shaw et al., 2017, Gordon et al., 2017]. BAMS patients show characteristic absence of nose and reduction in eye size, along with other malformations [Gordon et al., 2017]. Unlike FSHD, BAMS mutations are all missense mutations that are restricted to ATPase domain of the Smchd1 structure [Jansz et al., 2017, Gurzau

[et al., 2018](#)]. This stipulates that hindering of the ATPase activity could possibly be a therapeutic target for BAMS.

It is evident that Smchd1 plays a significant role in gene expression control in early development and disease onset. It is vital to understand the exact molecular function of Smchd1, its spatio-temporal dynamics and structure to tackle issues that arise when things go awry in disease. Even though now we know that Smchd1 plays a significant role in development, this knowledge comes at the back of malfunction seen late in development of differentiated cells. We are yet to carry out studies on early developmental stages to better understand how exactly Smchd1 imparts its function in epigenetic silencing. Much of our understanding on Smchd1 also comes from studies carried out *in vitro* or in proteins at a specific point in time. Therapeutic intervention could also benefit from understanding the molecular dynamics of Smchd1 in its native environment.

1.7 Aims

This project focused on the epigenetic modifier Smchd1, its function as a maternal effect protein and unraveling its structure and dynamics using advanced microscopy.

1.7.1 Aim 1: To evaluate whether Smchd1 is a maternal effect protein that has a role in genomic imprinting

I aimed to determine whether Smchd1 is a maternal effect protein and if it is, a. whether Smchd1 has a role in genomic imprinting at autosomal imprinted genes b. to evaluate the function of maternal Smchd1 in imprinted XCI.

1.7.2 Aim 2: To unravel the Smchd1 macro molecular structure and dynamics in the nucleus

In this aim I sought to explore how Smchd1 behaves in its unperturbed state, in the nucleus of the cell. I aimed to understand how Smchd1 behaves in the cells by using various advanced microscopy techniques.

Chapter 2

Methods

2.1 Mouse strains

2.1.1 Ethics approvals

All mice were bred and maintained in house at The Walter and Eliza Hall Institute of Medical Research (WEHI) using procedures approved by in-house ethics approval committee (approval numbers 2014.026, 2018.004).

2.1.2 MMTV-Cre *Smchd1*^{-/^{fl}} line

The MMTV-Cre *Smchd1*^{-/^{fl}} line was generated by backcrossing MMTV-Cre transgene line A [Wagner et al., 1997] onto the C57BL/6 background from the FVB/N background for more than 10 generations. These mice contain a combination of *Smchd1* deleted allele (*Smchd1*⁻) in trans to the *Smchd1* floxed (*Smchd1*^{fl}) allele [de Greef et al., 2018].

2.1.3 Zp3-Cre *Smchd1*^{fl} line

The Zp3-Cre *Smchd1*^{fl} line was generated and then was backcrossed onto the *Smchd1*^{fl} allele [Lewandoski et al., 1997].

2.1.4 CAST/EiJ line

The CAST/EiJ strain used to achieve polymorphism necessary for RNA-seq analysis was purchased from the Jackson laboratories.

2.1.5 *Smchd1*^{GFP} line

The mice containing *Smchd1*^{GFP} allele were generated as previously described [Jansz et al., 2018b] by backcrossing to C57BL/6 for more than 10 generations, and maintained as a homozygous breeding line.

2.1.6 Matings

Embryos and extra embryonic tissue were collected via natural timed matings or superovulations. For all post implantation embryos, natural timed matings were used and the preimplantation embryos were obtained by superovulation as previously described [Keniry et al., 2016]. Vaginal plug detection was counted as embryonic day 0.5 (E0.5).

2.1.7 Genotyping

Genotyping for the *Smchd1* alleles and sex chromosomes was carried out as previously described [Jansz et al., 2018b] (primer sequences in Table 2.1). The Cre transgenes were genotyped using a general Cre PCR as previously described [Leong et al., 2013]. In brief, tail tips and embryonic tissue were digested for at least 2 hours - overnight at 55°C in DNA lysis buffer with 200 µg ml⁻¹ proteinase K

(Roche) in a shaker. Proteinase K is inhibited by heating samples to 85°C for at least 20 minutes. The Genotype was determined by PCR. The PCR reaction contained 1 µl of digested tissue, 1X GoTaq Green Mix (Promega) and 0.5 µM per forward and reverse primers. PCR cycle conditions were 94°C for 5 minutes; 35 cycles of 94°C for 30 seconds, 60°C for 30 seconds and 72°C for 1 minute; 72°C for 10 minutes.

2.2 Derivation and culture of cell lines

2.2.1 Mouse embryonic fibroblasts (MEFs)

Embryos were harvested at E14.5 in sterile phosphate buffered saline (PBS). Using sterile equipment in an sterile environment, the internal organs and head were removed and the remaining embryo body was placed in sterile PBS. The remaining embryo body was vigorously pipetted up and down to break down tissue and create a single cell suspension using 1 ml of culture medium DMEM (Gibco) with 10% (v/v) FBS (Life Technologies)). Each embryo suspension was diluted with a further 9 ml of culture medium in 10 cm tissue culture plate and cultured at 37°C in a humidified atmosphere with 5% v/v CO₂ and O₂. Cells were passaged every 2-3 days using 0.5% trypsin-EDTA (Life Technologies) to detach adherent cells and replated at approximately 3000 cells cm⁻² seeding density. Primary MEFs were generally not cultured beyond 5 passages. MEFs were frozen in FBS with 10% v/v DMSO.

2.2.2 Neuronal Stem Cells (NSCs)

NSCs were derived by Mrs Kelsey Breslin as previously described [[Chen et al., 2015](#), [Jansz et al., 2018b](#)]. I cultured NSCs in NeuroCult NSC Basal Medium (Mouse) (StemCell Technologies) containing NeuroCult Proliferation Supplement (Mouse) (StemCell Technologies), 0.2% heparin solution (StemCell Technologies),

20 ng ml⁻¹ recombinant human EGF (Peprotech), and 20 ng ml⁻¹ recombinant human basic FGF (Peprotech). The cells were seeded into tissue culture treated plates coated with 10 ng ml⁻¹ laminin (Sigma-Aldrich) added to the medium, at a density of cells 40 000 cells cm⁻². Cells were cultured at 37°C in a humidified atmosphere with 5% v/v CO₂ incubator and passaged every 2 days using Accutase (Sigma-Aldrich) to detach the cells from the plates. Primary cells were maintained for a maximum of 20 passages.

2.3 Retrovirus production and transduction

2.3.1 Retrovirus production

293T cells were cultured in DMEM (Life Technologies) supplemented with 10% FBS and grown at 37°C in a humidified atmosphere with 10% (v/v) CO₂. Retrovirus was produced using a calcium phosphate mediated transient transfection of 293T cells. 293T cells at 80% confluency were transfected with VSVg or CAG-Eco envelope vector, MD1-gag-pol structural vector and shRNA retroviral construct at a ratio of 24:8:1. Plasmid DNA was added to a 250 mM CaCl₂ solution and precipitated in 1 volume of 2X HBS solution. This mix was then added drop-wise onto 293T cells. The media was changed 16 hours post-transfection and supernatant harvested 48 hours or 72 hours post-transfection by passing through 0.45 µm filter to remove residual 293T cells and concentrated using Poly ethylene glycol (PEG) [Jansz et al., 2018b].

2.3.2 Concentration of retrovirus with PEG

Retroviral supernatant from viral preparations that were for VSVg coated virus was added to 1/4 volume of 20% PEG solution pH 7.2 (20% Poly(ethylene glycol), 410 nM NaCl) and inverted every 15 minutes over 90 minutes in ice. Followed by incubation at 4°C for 16 hours. The following day, the solution was centrifuged

at 4500 rpm for 45 minutes at 4°C and the supernatant discarded. The pellet containing retrovirus was resuspended in 1/100 original volume of unconcentrated retroviral supernatant before PEG addition, resuspended retrovirus was aliquoted into micro vials at 10 µl volume and snap frozen in dry ice before storing at -80°C [Ludwig et al., 2018].

2.3.3 Transduction of MEFs

MEFs were seeded 16-24 hours prior to transduction. At 50% MEF confluence, PEG concentrated viral supernatant or 1/10 dilution of neat unconcentrated ecotropic coat retrovirus was added in media containing 4 µg ml⁻¹ polybrene (Sigma-Aldrich). Media changed 24 hours later. Cells were selected with 5 µg ml⁻¹ of puromycin (Sigma-Aldrich) or hygromycin 3 µg ml⁻¹ - 5 µg ml⁻¹ (Roche) as required for the cell type (primary or immortalised).

2.3.4 Transduction of NSCs

NSCs were seeded at a density of 40,000 cells cm⁻² 8-16 hours prior to transduction. PEG concentrated viral supernatant was added in media containing 4 µg ml⁻¹ polybrene (Sigma-Aldrich). After 24 hours, cells were selected with 1 µg ml⁻¹ of puromycin (Sigma-Aldrich).

2.4 Microscopy

2.4.1 Immunofluorescence studies on mouse ovary sections

Mice 6-14 weeks old were selected and ovaries were harvested. Harvested ovaries were fixed in 10% formalin overnight. The following day washed with 70% ethanol and delivered to Dr Jessica Stringer at Monash University for sectioning

and immunofluorescence as described in Wanigasuriya *et al.* [Wanigasuriya *et al.*, 2020]. Scoring of the sections was also carried out by Dr Jessica Stringer.

2.4.2 Native GFP fluorescence in preimplantation embryos

To assess the native GFP fluorescence embryos were collected at E1.5 (2-cell), E2.0(4-cell), E2.25(8-cell), E2.75(16-cell), E3.5 (blastocyst) stages as described previously [Borensztein *et al.*, 2017]. Appropriate controls were carried out and imaged as described in Wanigasuriya *et al.* [Wanigasuriya *et al.*, 2020]. At least 3 embryos per each experiment was imaged using AxioObserver (Zeiss) microscope at 20x magnification and scored using FIJI software [Schindelin *et al.*, 2012].

2.4.3 Immunofluorescence in preimplantation embryos

Test (textitSmchd1^{GFP/GFP} females X *Smchd1*^{+/+} male and *Smchd1*^{+/+} females X *Smchd1*^{GFP/GFP}), negative control (*Smchd1*^{+/+} X *Smchd1*^{+/+}) and positive control (*Smchd1*^{GFP/GFP} females X *Smchd1*^{GFP/GFP} male) embryos were harvested at E2.25(8-cell), E2.75(12 to 16-cell), E3.5 (24 to 32-cell) stage. Immunofluorescence was carried out as described in Wanigasuriya *et al.* [Wanigasuriya *et al.*, 2020]. At least 2 embryos per each timepoint were imaged using LSM 880 (Zeiss) microscope with Airyscan processing. Analysis was carried out using FIJI software [Schindelin *et al.*, 2012].

2.4.4 Micro-CT

E14.5 concepti were harvested. The yolk sac and embryo were detached from the placenta and the placenta placed in 4% paraformaldehyde (PFA) (Sigma) for 24h at 4°C. Placenta were washed in 100% ethanol for 5minutes followed by two washes in 70% ethanol and stored in 70% ethanol until imaging. Placental volumes were measured by scanning each placenta in a Bruker SkyScan 1276 at 25µm voxel size, at 40kV and 200uA. Scanning, reconstruction and image processing

was performed by Dr Lachlan Whitehead. Reconstruction was performed using the NRecon software. Image stacks were processed with a custom FIJI macro which performed filtering and thresholding to isolate the placenta from the sample holder, and volumes were measured.

2.4.5 Allele specific RNA-FISH on preimplantation embryos

2.4.5.1 Denhardt's treatment of coverslips

Denhardt's treatment was carried out on 18 x 18 coverslips prior to the FISH experiment. Briefly, coverslips were placed in glass staining trays (approx. 20 per tray), washed in 25 M HCl in room temperature for 20 minutes and rinsed in ddH₂O. The coverslips were then dehydrated in 70%, 95% and 100% ethanol, for 5 minutes each. The coverslips were incubated for 1-3 h in a solution of 1x Denhardt's (50x Denhardt solution Cat # 750018 life technologies) at 65°C in the hybridisation oven. Coverslips were dipped in distilled water once and incubated in fixative (3 part 95% ethanol : 1 part glacial acetic acid) for 20 minutes at room temperature. Coverslips were air dried for at least 12 h. Acetylation of coverslips was carried out by incubating the coverslips for 10 minutes at room temperature in a solution of 0.1 M triethanolamine (pH 8.0) and acetic anhydride. Coverslips were dehydrated by dipping for 5 minutes each in solutions of 50% ethanol, 70% ethanol, 95% ethanol, 100% ethanol. Coverslips air dried for at least 24 hours before using.

2.4.5.2 Probe preparation

Allele-specific *Xist* RNA FISH was performed as previously described [[Harris et al., 2019](#)]. Allele-specific *Xist* RNA FISH probes were generated as described [[Levesque et al., 2013](#)]. Briefly, a set of short oligonucleotide probes (5 probes for each *Xist* allele) were designed to uniquely detect either the Bl6 or the Cast alleles of *Xist* exon 7. Dr Quentin Gouil assisted with probe design. Each probe contained

single nucleotide polymorphism (SNP) located at the fifth base pair position from the 5' end that differs between the Bl6 and Cast. The 3' end of each oligonucleotide probe was fluorescently tagged using Quasar dyes (Biosearch technologies). Bl6-specific oligos were labelled with Quasar 570 and Cast-specific oligos labelled with Quasar 670. In addition to labeled SNP-overlapping oligonucleotides, a panel of 5 'mask' oligonucleotides were also synthesized (IDT).

2.4.5.3 Preparation of strand-specific *Xist* guide RNA probe

Exon 7 of *Xist* RNA was selected as the strand-specific *Xist* guide probe. Exon 7 specific primers were designed (IDT) with T3 and T7 promoter overhangs. Exon 7 was amplified from 50 ng of an *Xist* cDNA clone [Wutz and Jaenisch, 2000]. Briefly, the PCR reaction contained cDNA, 5x Phusion HF reaction buffer (Cat # 13058S), 1 µl Phusion Taq, 10 mM dNTP, and 10 µM per forward and reverse primers. PCR cycle conditions were 98°C for 2 minutes; 30 cycles of 98°C for 30 seconds, 58°C for 30 seconds and 72°C for 30 seconds; 72°C for 4 minutes. PCR product was isolated using QIAquick gel extraction kit (Qiagen) according to manufacturer's instructions. Strand-specific *Xist* RNA probe was labelled with Fluorescein-12-UTP (Roche, Cat # 11427857910) and ethanol precipitated as previously described [Hinten et al., 2016]. Probe was re-suspended in hybridisation buffer containing 10% dextran sulfate, 2X saline-sodium citrate (SSC) and 10% formamide.

2.4.5.4 Preparation of pre-implantation embryos

E2.75 embryos were collected and the zona pellucida removed by keeping in acid tyrode's solution (Sigma) for 2 minutes. Embryos were placed in the middle of Denhardt's treated coverslips in 1x PBS 6 mg/ml BSA using finely pulled Pasteur pipette. Excess 1x PBS 6 mg/ml BSA was aspirated and embryos let dry for 20–30 minutes. Embryos were fixed and permeabilised with 50 µl of 1% PFA in 1x PBS with 0.05% Tergitol for 5 minutes. Embryos were rinsed with three changes of 70%

ethanol then dehydrated through an ethanol series (85%, 95%, 100%) 2 minutes each at room temperature. Samples were then airdried for 5-10 minutes.

2.4.5.5 Allele specific RNA-FISH

The precipitated guide RNA probe was mixed with the Bl6 and Cast detection probes, to a final concentration of 5 nM per allele-specific oligo, and 10 nM mask probe, yielding a 1:1 mask:detection oligonucleotide ratio. Coverslips were hybridised to the combined probe overnight in a humid chamber at 37°C. After overnight hybridisation, samples were washed twice in 2x SSC with 10% formamide at 37°C for 30 minutes, followed by one wash in 2x SSC for five minutes at room temperature. A 1/250,000 dilution of DAPI (Invitrogen, Cat # D21490) was added to the second 2X SSC with 10% formamide wash. Coverslips were then mounted on slides in Vectashield (Vector Labs, Cat # H-1000). Stained samples were imaged immediately using a LSM 880 (Zeiss) microscope.

2.4.6 Fluorescence recovery after photobleaching (FRAP)

NSCs or immortalised MEFs were seeded on Ibidi 1.5 µm 8 well glass bottom slides and incubated at 37°C in a humidified atmosphere with 5% v/v CO₂ overnight. Media was changed to fluorobrite DMEM (Life Technologies Cat # A1896701) prior to imaging. Imaging was carried out using a LSM 880 (Zeiss) microscope, 5 images were taken pre-bleach, and 1 every minutes post-bleach. The Bleach was performed at the region of interest using 405 and 488 laser at 60%. Analysis was performed by Dr Quentin Gouil.

2.4.7 Mitotic shake-off

Immortalised MEF scultured in T150 tissue culture treated flasks until confluent, and split 1 in 10 to new T150 tissue culture treated flasks. Following day when flasks are sub confluent (cells in mid-log phase), gently shaken by tapping and

washed off the loosely attached mitotic cells. Wash media containing cells enriched for mitosis cytopun at 300g for 5 min. fixed using 3%PFA added DAPI for 10minutes, washed 3x times and mounted used vectasheild H-1000. Images were acquired using the LSM 880 (Zeiss) microscope. Image analysis was carried out using FIJI software [Schindelin et al., 2012].

2.4.8 Lattice LightSheet microscopy

Cells were plated onto coverslips and incubated at 37°C in a humidified atmosphere with 5% v/v CO₂ overnight. Media was changed to L-15 Medium (Leibovitz) with 10%FBS for imaging on the lattice lightsheet microscope.

For all Lattice LightSheet microscopy imaging a custom home-built system was used, constructed as outlined in (Chen *et al.*, 2014) as per licensed plans kindly provided by Janelia Farm Research campus [Chen et al., 2014]. Excitation light from 488 nm and 561 nm diode lasers (MPB Communications) were focused to the back aperture of a 28.6X 0.7 NA excitation objective (Special optics) via an annular ring of 0.44 inner NA and 0.55 outer NA providing a light sheet 10 µm in length. Fluorescence emission was collected via a 25X 1.1 NA water dipping objective (Nikon) and detected by either one or two sCMOS cameras (Hamamatsu Orca Flash 4.0 v2. Imaging of Smchd1-GFP and H2B-mScarlet was performed using the dual-channel custom home-built LLSM. Emitted fluorescence was collected via a 561 nm dichroic before passing through a 405/488/561/633 nm multi-band filter (Chroma) and 525/50 nm filter. A liberally coated fluorescence bead slide was imaged to provide reference for registration of the two channels. Images were captured with a frame time of 20 ms and a full stack of 168 images over two channels was recorded every 2 minutes. Registration was performed using LLSpy and an Affine transformation. For imaging Mitotracker Green in combination with PKH26 525/50 nm and 405/488/561/633 multi-band filter sets were used respectively. All data was acquired in an imaging chamber (Okolabs) set to 37°C and 5% humidified CO₂. For all experiments Point Spread Functions (PSFs) were measured using 200 nm Tetraspeck beads on the surface of a 5 mm coverslip. Data

was deskewed and deconvolved using LLSpy, a Python interface for processing of LLSM data. Deconvolution was performed using a Richardson-Lucy algorithm using the PSFs generated for each excitation wavelength.

Data is presented as a Maximum Intensity Projection of a 3D stack. Noise was minimised for the Smchd1-GFP channel using a median filter of radius 2.

2.4.9 Direct stochastic optical reconstruction microscopy (dSTORM) sample preparation

MEFs were cultured at 40,000 cells per well overnight in the glass bottom chambers of 8-well plates (Sarstedt). Cells were then fixed with PFA (2% w/v) in PBS for 10 minutes at room temperature. The majority of the culture medium was aspirated using a glass pipette placed carefully into a corner of each chamber. Simultaneously, 500 ml fixative was added down the side of the chamber using a second glass pipette followed by PBS washing steps to remove excess fixative.

Important notes: This PFA was freshly prepared from or thawed from frozen 10% PFA, with filtering after thawing. PFA pH was adjusted to pH 7.4 by adding 2x Phosphate buffered saline (PBS) and then the equivalent milli-Q water to make an isotonic solution. This is again pH adjusted using 1 M KOH or HCl to achieve a pH of 7.4. PBS washes were done carefully by positioning the pipette tip in the same corner of each chamber to minimize perturbation damage. 500 μ l of PBS was added concurrently to the removal of the fixative solution. This PBS wash was aspirated after approximately 30 seconds and replaced with fresh PBS that was removed after another 60 seconds. Three more PBS washes were administered lasting approximately 5 minutes each. Free-Aldehyde fluorescence was minimised with the addition of freshly prepared 0.1% (w/v) NaBH₄ in PBS for 15 minutes, followed by two washing steps of PBS.

Samples were permeabilised with TritonX-100 (0.1% v/v in PBS) followed by 30 s wash in PBS followed by 3x washing steps with PBS. Samples were blocked using bovine serum albumin (BSA) (0.1% v/v in PBS).

Immunostaining was then performed, using primary antibody diluted in blocking buffer (0.1% BSA/PBS) over night (approx. 16 h) followed by three washes of 5 minutes in PBS. Following this, the samples were incubated for 1 hour with the secondary antibody diluted in blocking buffer (0.1% BSA/PBS). Samples were washed 3 times with PBS. 1:500 Tetraspeck beads (in PBS) were added and incubated for 10 minutes. Post fixation with 2% PFA was performed for 10 minutes. Samples were washed with PBS for 10 minutes prior to storage in dark chamber in PBS until imaging. Samples were imaged using Nikon nSTORM and analysed by Dr Michael Mlodzianozki.

2.5 Genomic assays

2.5.1 Embryos and placentae collection for bulk RNA-seq and DNA-seq

Embryos were dissected at E9.5 and placentae were dissected at E14.5 and snap-frozen in dry ice. RNA and DNA was extracted from these snap frozen samples using Qiagen All prep kit (Qiagen, Cat # 80204), or Zymo Quick DNA/RNA miniprep plus kit (Zymo research, Catalog # D7003) according to manufacturers' instructions. DNA and RNA quantity was calculated using nanodrop (Denovix DS-11 spectrophotometer).

2.5.2 Library preparation - bulk RNA-seq

For RNA-seq libraries were prepared using less than 500 ng total RNA. Libraries were prepared according to manufacturers' instructions using TrueSeq V1 or V2 RNA kits (Illumina). cDNA libraries were quantified using D1000 tape on the TapeStation (4200, Agilent Technologies) for sequencing using Illumina Nextseq500 platform, with 80-bp, paired-end reads. Obtained libraries were analysed by Dr Quentin Gouil as described in Wanigasuriya *et al.* [Wanigasuriya *et al.*, 2020].

2.5.3 Library preparation - bulk reduced representation bisulfite sequencing (RRBS)

RRBS was performed using extracted DNA. First DNA was cleaned using Zymo research DNA Clean & concentrator-5 kit and quantified using Qubit dsDNA assay kit (ThermoFisher Scientific Q32853). 100 ng was used to prepare RRBS libraries as described in Wanigasuriya *et al.* [Wanigasuriya *et al.*, 2020]. Analysis was also carried out by Dr Quentin Gouil according to Wanigasuriya *et al.* [Wanigasuriya *et al.*, 2020].

2.5.4 Single-preimplantation embryo sequencing - M&T-seq

E2.75 embryos were collected by flushing uteri of female mice and the collected embryos were sorted into 96 well plates containing preprepared mixture of 2.5 μ l of RLT Plus (Qiagen) lysis buffer with 1 U μ l⁻¹ SUPERase-IN (Ambion). These plates were snap frozen and sent to University of Newcastle. Library prep and sequencing were carried out in collaboration with Dr Heather Lee and Ellise Roper from University of Newcastle as described in Wanigasuriya *et al.* [Wanigasuriya *et al.*, 2020]. M&T-seq analysis was carried out by Dr Quentin Gouil as also described in Wanigasuriya *et al.* [Wanigasuriya *et al.*, 2020].

TABLE 2.1: Primers for genotyping

Primer	Sequence	concentration
Zfy F	GAC TAG ACA TGT CTT AAC ATC TGT CC	10 μ M
Zfy R	CCT ATT GCA TGG ACA GCA GCT TAT G	10 μ M
OTC-F	GTT CTT TCG TTT TCC CCT CTC	10 μ M
OTC-R	GGC ATT ATC TAA GGA GGA GCA TC	10 μ M
SmchD1 target F1	GCC TGC CTT GCT TCT ATG TC	10 μ M
SmchD1 target R1	GCC CCA TGA GAT TCT GAA AG	10 μ M
SmchD1 target R2	GAA TTC AGG GTC AGC TTG C	10 μ M
Cre F (WA 693)	CTGACCGTACACCAAAATTTGCCTG	10 μ M
Cre R (WA 694)	GATAATCGCGAACATCTTCAGGTTC	10 μ M
Smchd1 cKO F2	TCAGGTGGTCTCGAGCCC	10 μ M
Smchd1 cKO F4	CCA TGA GAA GCA ATG TGG GA	10 μ M
Smchd1 cKO R1	GGACAGCCAAAGTGACACAG	10 μ M

TABLE 2.2: Primary and secondary antibodies

Antibody	Use	Host	Manufacturer	Catalogue No.
Smchd1 (ATPase domain)	1:100	Rat	In house. Source - PMID: 30127357	#5 or Clone 5H4
Smchd1 (ATPase domain)	1:100	Rat	In house. Source - PMID: 30127357	#8 or Clone 2B8
C-Kit	1:500	Goat	Novus	af1356
GFP	1:500	Rabbit	ThermoFisher Scientific	A-11055
AF488	1:500	Rabbit	Life Technologies	A28175
AF488	1:500	Goat	Life Technologies	A11055
DyLight 550	1:500	Rat	Invitrogen	SA5-10027
AF568	1:500	rat	Life Technologies	A11077
AF568	1:500	rabbit	Life Technologies	A21207
AF647	1:500	rat	Life Technologies	A21248
AF647	1:500	rabbit	Life Technologies	A21244

TABLE 2.3: T3/T7 *Xist* primers for single stranded FISH probe template

Label	Primer sequence	Sequence with T7 or T3 promoter
mXist_ex7-1a	cttgtgtgtctaattcgttggtg	ggatcctaatacgactcactatagggtgtgtgtgtctaattcgttggtg
mXist_ex7-1b	cagttcctcttcttgggtgtg	agcgcgcaattaaccctcactaaagggcagttcctcttcttgggtgtg
mXist_ex7-2a	gctctgaagtgaacgcccaagtag	ggatcctaatacgactcactatagggtctgaagtgaacgcccaagtag
mXist_ex7-2b	cagaggaagaggaaggcacgaaaaa	agcgcgcaattaaccctcactaaagggcagaggaagaggaaggcacgaaaaa
mXist_ex7-3a	agggaagttaaatgacacaggatg	ggatcctaatacgactcactataggaggggaagttaaatgacacaggatg
mXist_ex7-3b	ttccatgtgtttctcttctgcta	agcgcgcaattaaccctcactaaagggtccatgtgtttctcttctgcta
mXist_ex7-4a	tgtgtttctctgcttttggtgagg	ggatcctaatacgactcactatagggtgtgtttctctgcttttggtgagg
mXist_ex7-4b	tcccggctttatagaactgtagac	agcgcgcaattaaccctcactaaagggtcccggctttatagaactgtagac
mXist_ex7-6a	accattttcttttcttactgtt	ggatcctaatacgactcactataggaccattttcttttcttactgtt
mXist_ex7-6b	gcttgtttgccccaggataat	agcgcgcaattaaccctcactaaagggtgtttgtttgccccaggataat
mXist_ex7-7a	accaattttgctgtgtcgtaa	ggatcctaatacgactcactataggaccattttgctgtgtcgtaa
mXist_ex7-7b	caactacccacccactcaac	agcgcgcaattaaccctcactaaaggcacaactacccacccactcaac

TABLE 2.4: *Xist* allele specific FISH probes

C57Bl6	Cast	Mask	SNPid
ATCACGCTGAAGACCCAGTTTTCTG	ATCATGCTGAAGACCCAGTTTTCTG	CAGAAAACGGGTCTT	rs29080486
ATGCTGGGAGAACTGCTGTTGTGATG	ATGCCGGGAGAACTGCTGTTGTGATG	CATCACAAACAGCAGTT	rs228261199
GCTCGGTGGATGAGTTTGAAAGAAAGTAC	GCTCAGTGGATGAGTTTGAAAGAAAGTAC	GTACTTTCTTTCAAACCTCA	rs237964379
ATTGGTTGCTTTTATCCAGTCCGCTG	ATTGATTGCTTTTATCCAGTCCGCTG	CAGCGGACTGGATAAA	rs265848601
CGGAAAGAAAAGTGCTAAGCTTCCTTCTC	CGGAGAGAAAAGTGCTAAGCTTCCTTCTC	GAGAAGGAAGCTTAGCACT	rs218242058

Chapter 3

Smchd1 is a maternal effect gene required for genomic imprinting

3.1 Abstract

The role for zygotic *Smchd1* is well established in genomic imprinting at the canonical imprinted clusters, *Snrpn* and *Igf2r* [Gendrel et al., 2013, Mould et al., 2013]. Here, I explored *Smchd1*'s function as a maternal effect gene and its involvement in genomic imprinting. I found that *Smchd1* is a novel maternal effect gene. When *Smchd1* was maternally deleted, relaxation of imprinting was observed at several non-canonical imprinted genes. I also found that this effect could be seen in preimplantation embryos, and it persists until later in gestation for some genes. Furthermore, by more comprehensively testing the role of zygotic *Smchd1* than had occurred previously I revealed a role for *Smchd1* at a new canonical imprinted cluster, *Kcnq1*. Interestingly, *Smchd1*-sensitive non-canonical and canonical imprinted genes are known to rely on polycomb mediated methylation for their primary and secondary imprints, respectively. These data suggest that *Smchd1* may operate via a polycomb-mediated mechanism. I propose that maternal *Smchd1* establishes imprinted expression downstream of polycomb group proteins, then zygotic *Smchd1* maintains this imprinted expression later in development.

3.2 Introduction

Classically, imprinted expression is enabled by differentially methylated regions (DMRs) established in the germline (primary DMRs) and maintained throughout development with the help from somatic (secondary) DMRs [Barlow and Bartolomei, 2014, Ferguson-Smith, 2011]. This DNA methylation dependent imprinted expression control is known as “canonical imprinting”. But not all imprinted gene expression can be explained by the canonical DNA methylation mechanism. Recently, maternally deposited H3K27me3 was described to mediate allele specific expression of some lone imprinted genes including *Sfmbt2*, *Phf17* (also known as *Jade1*), *Smoc1*, and long non coding RNA *Xist* [Inoue et al., 2017b,a]. This “non-canonical” polycomb-mediated imprinted gene expression control was observed to be transient, only existing in the early embryo. For some of the non-canonical imprinted genes a secondary DMR is established later in development. Little is known about what is required for these non-canonical imprints to bring about imprinted expression, or how they are maintained for the short window of time in which they exist.

Maternal effect proteins play an important role in enabling imprints to resist epigenetic reprogramming during preimplantation. These maternally derived proteins found in the oocyte control imprinted expression during early embryo development. Most of the known maternal effect genes that act on imprinted gene regulation (i.e. *Trim28*, *Zfp57*, *Stella*, *Rlim*, *Nlrp2*, *Dnmt3L*, and *Dnmt1o*) regulate canonically imprinted genes [Alexander et al., 2015, Hiura et al., 2006, Howell et al., 2001, Li et al., 2008, Messerschmidt et al., 2012, Payer et al., 2003, Shin et al., 2010]. While only *Eed* and the relevant demethylase *Kdm6b* are thus far described to partake in non-canonical imprinted gene expression regulation [Inoue et al., 2017b,a].

Zygotic *Smchd1* plays a crucial role in X chromosome inactivation [Blewitt et al., 2008, Gendrel et al., 2012] and silencing of clustered gene loci on autosomes, including genes at imprinted clusters, *Snrpn* and *Igf2r-Airn* [Chen et al., 2015, Gendrel et al., 2013, Jansz et al., 2018a, Mould et al., 2013]. Removal of zygotic *Smchd1* causes loss of imprinting of certain genes in these clusters and hypomethylation of

the secondary DMR at the *Snrpn* cluster [Mould et al., 2013]. Although we know much about zygotic Smchd1 we are yet to establish a role for the maternal form of Smchd1. Because of Smchd1's known role at these imprinted clusters [Chen et al., 2015, Gendrel et al., 2013, Jansz et al., 2018a, Mould et al., 2013] and also its expression in the oocyte [Midic et al., 2018] I set out to look at Smchd1 as a maternal effect gene. All genomic analyses described here were carried out in collaboration with Dr Quentin Gouil. Here we only looked at the male embryos and extra embryonic tissue in order to avoid any confounding effects due to X inactivation in females. Data from females will be presented in Chapter 4.

3.3 Results

3.3.1 Maternal Smchd1 expression in the early embryo

First, in order to confirm whether there is maternal Smchd1 present in the early embryo I carried out immunofluorescence on our Smchd1-GFP fusion knock-in mouse [Jansz et al., 2018b], as GFP acts as a marker of Smchd1. I used Smchd1-GFP dams crossed with wild-type males so that GFP marked maternal-specific expression. Embryos from wild-type dams and wild-type sires were used as negative controls. Controls were essential as the native GFP expression was low, and high autofluorescence was observed in preimplantation embryos. I found that maternal Smchd1 was present in the 8-cell stage embryo. Maternal Smchd1 colocalised with the nuclear DAPI signal (Figure 3.1). This is concordant with zygotic Smchd1 data [Blewitt et al., 2008], suggesting that maternal Smchd1 may play a functional role similar to zygotic Smchd1 at this developmental window.

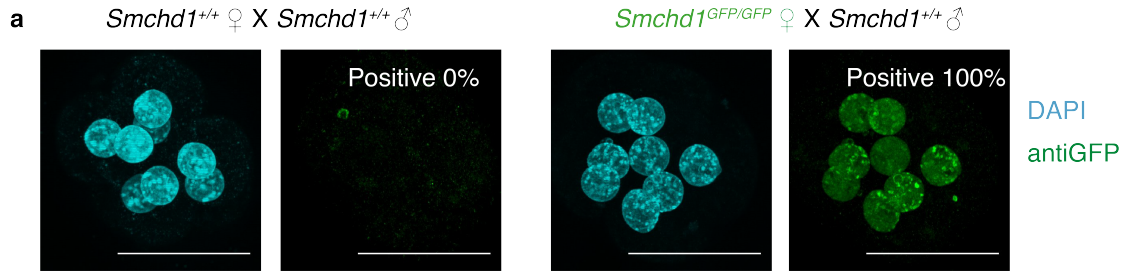


FIGURE 3.1: Maternal *Smchd1* is detected in preimplantation embryos.

Detection of maternal *Smchd1*-GFP in the nuclei of E2.5 (8-cell) embryos, with *Smchd1*^{+/+} used as negative controls and *Smchd1*^{GFP/+}, where GFP allele comes from the dam was used as the test (n = 1–2). Nuclei marked with DAPI (cyan). % refers to number of embryos with positive GFP signal. Scale bar: 50 μ m.

3.3.2 Timing of maternal *Smchd1* deletion during oogenesis in our maternal null models

To understand whether maternal *Smchd1* is present during oocyte development I collected control ovaries from mice with oocyte specific deletion of *Smchd1*. We used two different transgenes to delete *Smchd1* in the oocyte: MMTV-Cre and Zp3-Cre (Figure 3.2a). We collaborated with Prof. Karla Hutt and Dr Jessica Stringer from Monash BDI to carry out immunofluorescence on ovary sections from the two different Cre backgrounds along with control mice. We found that maternal *Smchd1* is present throughout oogenesis in both Zp3-cre and MMTV-Cre control ovary samples (Figure 3.2b panel 1). In the MMTV-Cre background *Smchd1* was gradually deleted from the early secondary to antral follicle stages of oocyte development, while in Zp3-Cre, *Smchd1* was deleted from the primary follicle stage, earlier than MMTV-Cre (Figure 3.2b panel 2 and 3).

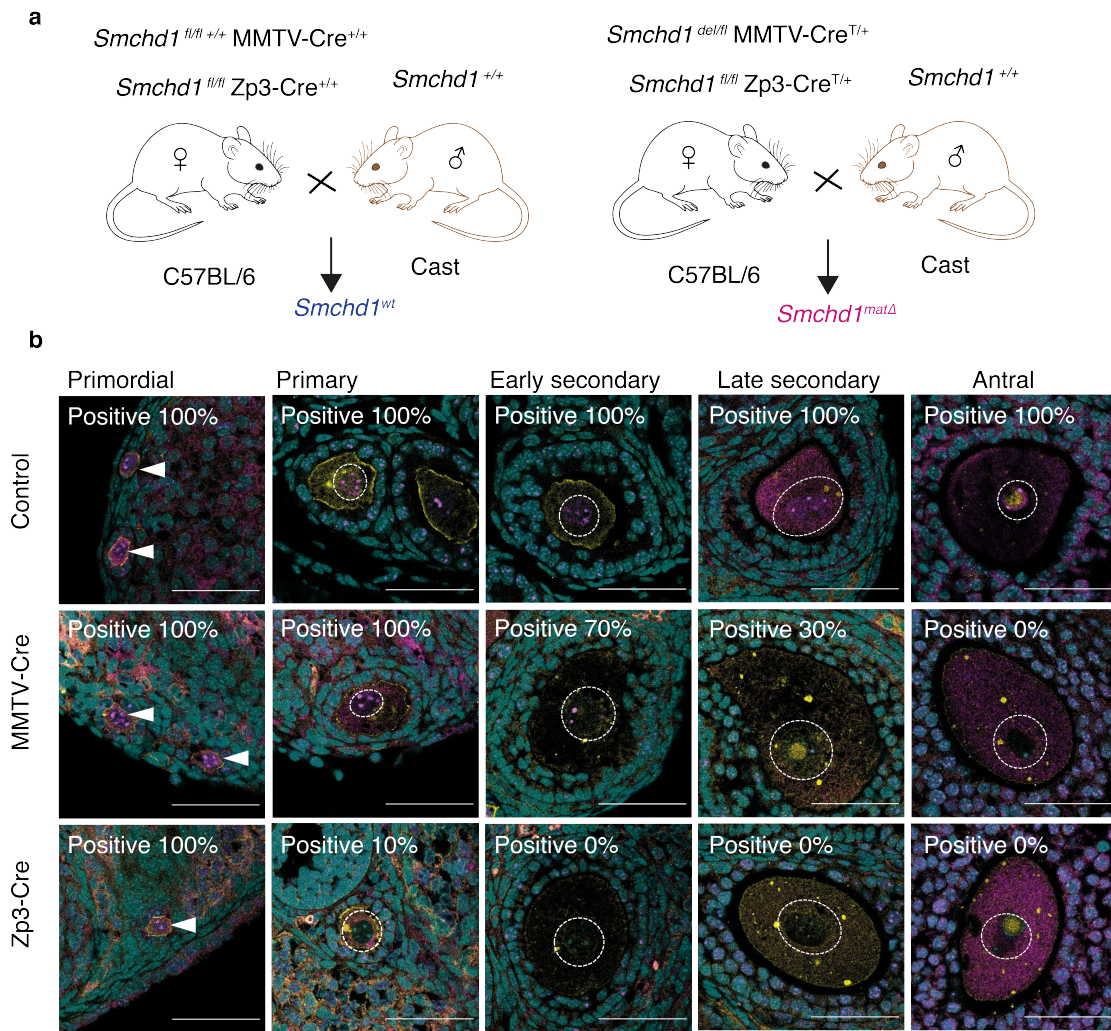


FIGURE 3.2: Maternal *Smchd1* is deleted at different stages of oocyte development by MMTV-Cre and Zp3-Cre.

a. Schematic for maternal deletion of *Smchd1*. **b.** Deletion of *Smchd1* in oocyte development with MMTV-Cre and Zp3-Cre. (In *Smchd1*^{T/+} T stands for the transgene). Arrowheads indicate primordial follicle oocyte nuclei, white dotted lines surround primary–antral follicle oocytes. *Smchd1*, c-Kit, DAPI. n = 15–27 sections for 2 ovaries per cohort. A total of 5–20 follicles were observed for primordial to late secondary stages and 2–3 antral follicles for each genotype. Scale bar: 50 μ m.

3.3.3 Maternal deletion of *Smchd1* depletes *Smchd1* until the 16-cell embryonic stage

Next I wanted to understand how long the embryo is under the control of maternal *Smchd1* before zygotic *Smchd1* is activated. Therefore, I performed imaging on a series of preimplantation embryos at different developmental stages from E1.5

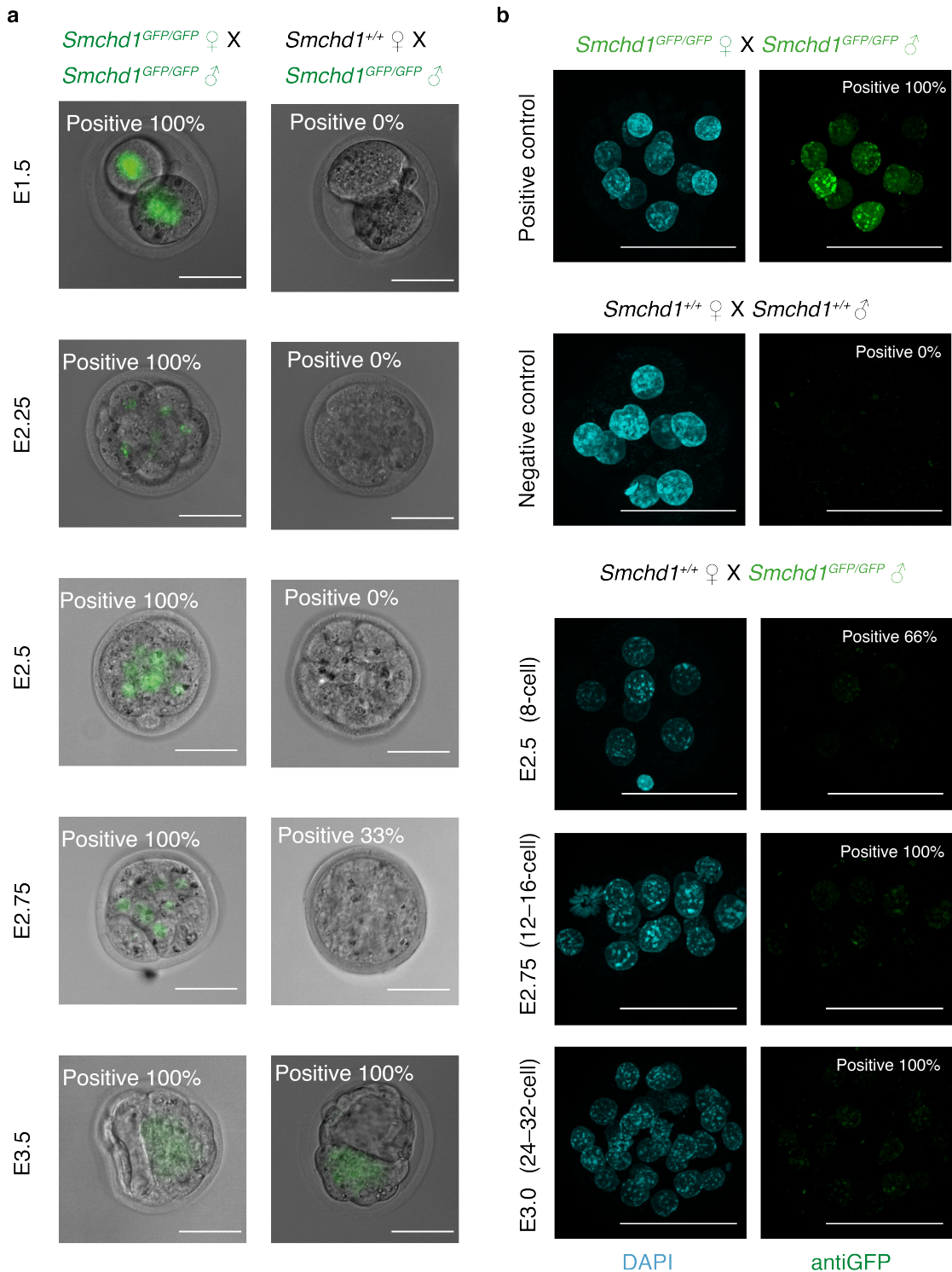


FIGURE 3.3: Zygotic *Smchd1* is not detectable until around the 16-cell embryonic stage of embryonic development.

a. Detection of maternal *Smchd1*-GFP in the nuclei of E2.5 (8-cell) embryos, with *Smchd1*^{+/+} used as negative controls (n = 2–3). Nuclei marked with DAPI. **b.** Immunofluorescence detection of maternal *Smchd1*-GFP using an anti-GFP antibody in the nuclei of 8-cell embryos, with *Smchd1*^{+/+} used as negative controls and *Smchd1*^{GFP/GFP} embryos used as positive controls. Zygotic *Smchd1* was detectable in some cells from the 8-cell stage onwards but comparatively had less signal intensity to the GFP control when embryos were imaged with identical settings, in experiments performed on the same day. n = 2 embryos. % refers to number of embryos with positive GFP signal. Scale bar: 50 μ m.

-E3.5. These embryos were obtained from our *Smchd1*-GFP fusion knock-in mouse [Jansz et al., 2018b]. Test embryos contained a GFP allele (*Smchd1*^{GFP/+}) inherited from their father. So, if the GFP expression was detected, that is due to the activation of zygotic *Smchd1* expression. *Smchd1*^{+/+} were used as negative controls. All paired test and control embryos were collected on the same day to allow for standardised imaging settings. Biological replicates were performed on separate days. By looking at the native fluorescence signal I found that paternally encoded *Smchd1*-GFP protein was not detectable until E2.75, when I estimate there would be 16 cells in the embryo (Figure 3.3a). Unfortunately using native fluorescence I could not incorporate a nuclear stain so was unable to definitively count cells. The timing of zygotic *Smchd1* expression was further confirmed using immunofluorescence against GFP (Figure 3.3b) where very faint nuclear *Smchd1* signal was detected after the 8 cell stage of embryo development, however this expression is much lower than maternal *Smchd1* at this time (Figure 3.1). All controls were imaged at the same time under same imaging settings along with the test embryos. The fluorescence intensity of the data presented in the Figure 3.3b is normalised to control images taken on the same day. Zooming in to Figure 3.3b is required for better visibility of faint staining. Together these data suggest maternally derived *Smchd1* is the primary source of *Smchd1* protein until at least the 32-cell stage. Together with the above oogenesis data, these data define the period of maternal deletion of *Smchd1* for each of our Cre-mediated deletion models.

3.3.4 Maternal deletion of *Smchd1* does not cause pre-weaning lethality or altered postnatal weight in progeny

To determine the effect maternally derived *Smchd1* has on phenotype, I used oocyte specific *Smchd1* deletion driven either by MMTV-Cre or Zp3-Cre (Figure 3.1). We found no embryonic or pre-weaning lethality following maternal deletion of *Smchd1* (Figure 3.4a), neither did we find any effect on the weights of the progeny at postnatal day 2 from crosses for maternal deletion and crosses where a

deleted allele was inherited from the father compared to the control crosses (Figure 3.4b). This contradicts the previous work carried out by depleting *Smchd1* in the preimplantation embryos using siRNA knockdowns, where embryos resulted in reduced number of blastocyst formation, hatching and survival to term [Midic et al., 2018, Ruebel et al., 2019].

3.3.5 Maternal deletion of *Smchd1* does not have a striking effect on gene expression

To observe the effect of maternal *Smchd1* deletion on gene expression, I collected *Smchd1* maternally deleted (*Smchd1*^{matΔ}) male E9.5 embryos and the embryonic portion of E14.5 placentae and carried out allele-specific RNA-seq. Dr Quentin Gouil analysed the RNA-seq data. Each Cre model resulted in only modest fold-change differences in gene expression, with 295 and 30 significantly differentially expressed genes for the MMTV-Cre and Zp3-Cre samples, respectively (Figure 3.5a-c). Although MMTV-Cre was found to delete slightly later in oogenesis we found a larger number of differentially expressed genes in this comparison, likely due to the higher number of samples (n=6 rather than n=4). Only five genes were consistently differentially expressed between *Smchd1*^{matΔ} and wild type control placental samples in both Cre models. These data suggest that removal of maternal *Smchd1* does not have a pronounced effect on gene expression in the placenta.

In the embryo, we observed 228 upregulated and 223 downregulated genes (Figure 3.5a,c). Again, the vast majority had very small fold changes. Interestingly, a subset of known *Smchd1* target genes, the clustered protocadherins, were more notably upregulated in the *Smchd1* maternal null samples (Figure 3.5c), suggesting either an unexpected maternal effect on these genes or perhaps a heterozygous effect. These data overall suggest the maternal *Smchd1* does not play a dominant role in regulating gene expression, and are consistent with no observable effect of removal of maternal *Smchd1* on viability or progeny weights (Figure 3.4)

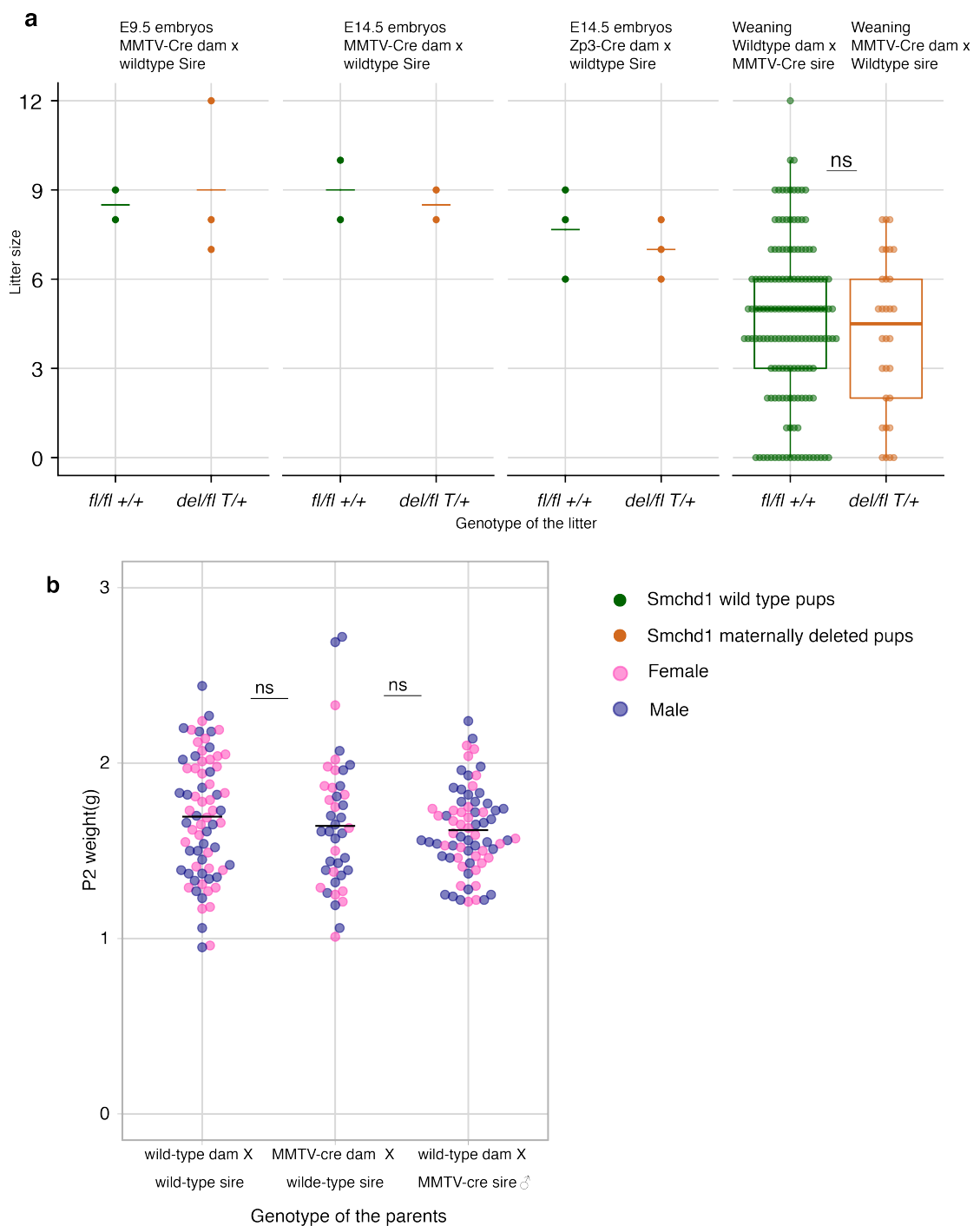


FIGURE 3.4: Maternal deletion of *Smchd1* does not result in pre-weaning lethality or altered postnatal weight

a. Litter size data for *Smchd1* maternal null crosses and control crosses at embryonic days 9.5, 14.5 and at weaning. Each individual litter is shown, with orange denoting the maternal null litters and green for control or reciprocal cross litters. For weaning $n > 20$. The p-value was calculated using Student's two-tailed T-test. **b.** Postnatal day 2 (P2) female and male pup weights for progeny generated from *Smchd1* maternal null, control and reciprocal crosses as shown. $n > 20$ pups for each cross. The P-value was calculated using student two-tailed T-test.

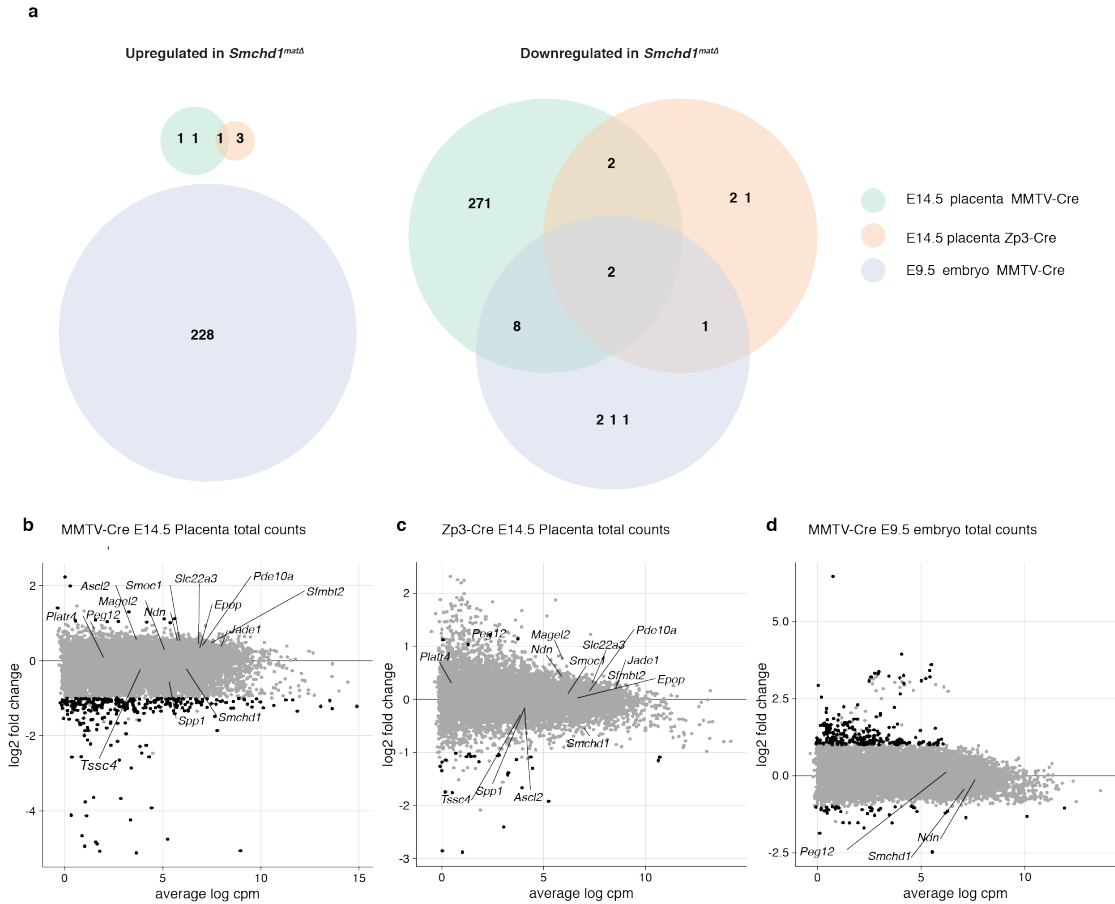


FIGURE 3.5: Maternal deletion of *Smchd1* does not have a major effect on gene expression in the embryo or the placenta

a. Venn diagram showing differentially expressed genes between E14.5 MMTV-Cre placentae, E14.5 Zp3-Cre placentae and E9.5 MMTV-Cre embryo. MA-plots of differential gene expression in **b.** MMTV-Cre and **c.** Zp3-Cre maternal deletion experiments, E14.5 placental samples, and **d.** MMTV-Cre E9.5 embryos. Genes below the 5% FDR and differentially expressed by at least 2-fold are plotted in black. RNA-seq: $n = 6$ MMTV-Cre, $n = 4$ Zp3-Cre, (wt); $n = 7$ MMTV-Cre, $n = 4$ Zp3-Cre, *Smchd1* maternal null (*matΔ*) E14.5 placental samples. $n = 3$ wt and $n = 4$ *matΔ* MMTV-Cre E9.5 embryo samples. *Smchd1* and genes with partial loss of imprinting are labelled.

3.3.6 No striking difference between *Smchd1* maternal deletion models - MMTV-Cre and Zp3-Cre

We next used allele-specific analysis of the RNA-seq to examine whether maternal *Smchd1* plays a role in autosomal imprinted gene expression. In order to do this we took advantage of the single nucleotide polymorphisms (SNPs) afforded by the fact all samples were F1 hybrids between C57BL/6 (maternal) and CAST/EiJ

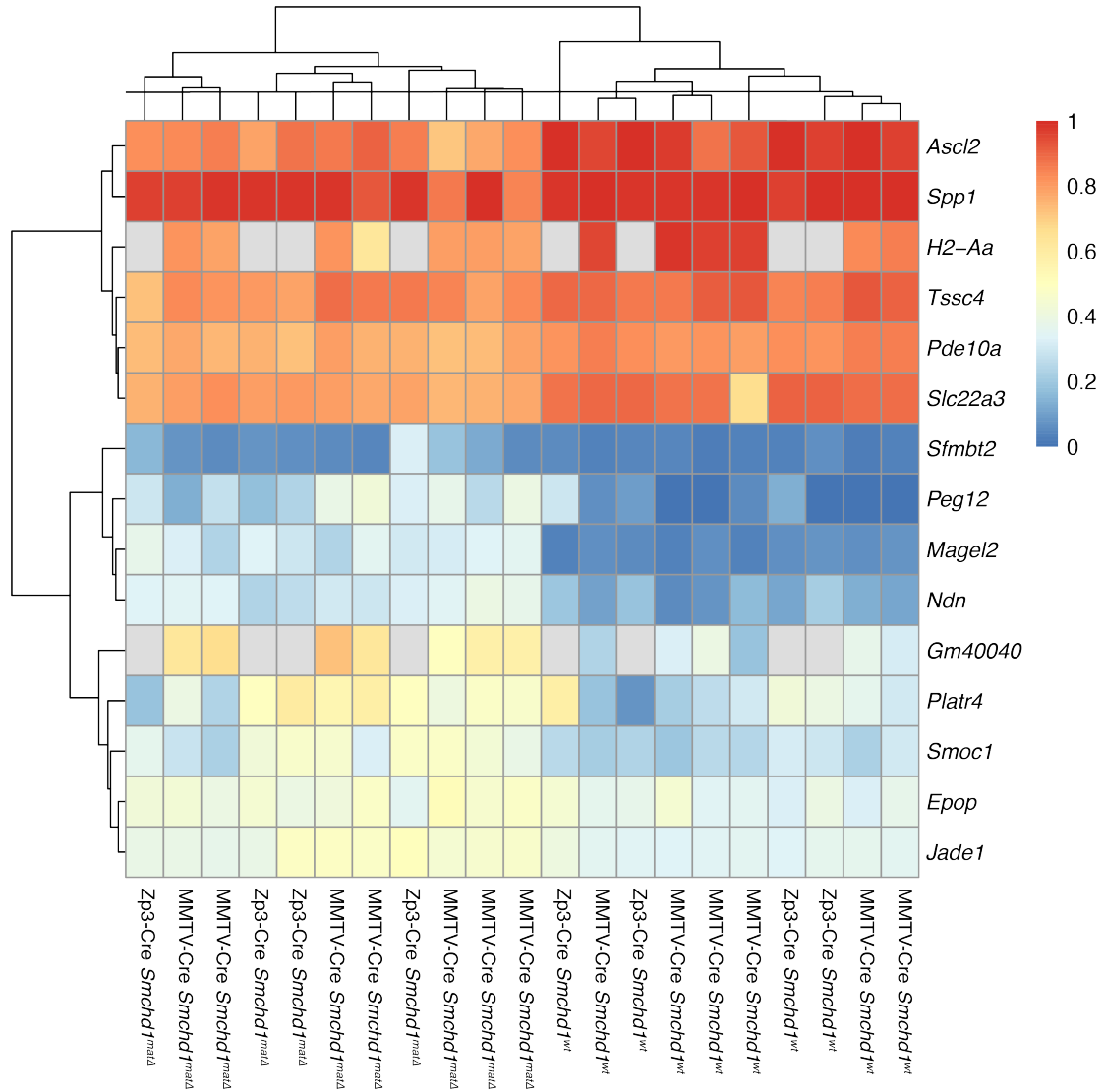


FIGURE 3.6: **MTTV-Cre and Zp3-Cre shows consistant effects on imprinted expression**

Heatmap of maternal allele expression proportion for genes that are differentially imprinted in at least one experimental set (MMTV-Cre and Zp3-Cre E14.5 placenta).

(paternal) (Figure 3.2b).

Reads were first mapped to an N-masked genome and then assigned to each genome based on available SNPs. This allowed us to compare imprinted gene expression between maternal null and control samples. The unsupervised hierarchical clustering data shows maternally deleted samples clustering together, demonstrating similar effects on imprinted expression with both Cre strains (Figure 3.6). These

data suggest that the stage of deletion during oogenesis does not influence outcome in terms of any effects on autosomal imprinting. Thus, we decided to pool data from both Zp3-Cre and MMTV-Cre strains for all statistical data analyses following this point. We imposed a 5% FDR as well as an absolute difference in the proportion of the RNA derived from the silenced allele greater than 5% to call differential imprinted expression.

3.3.7 Maternal *Smchd1* is not required to regulate the imprinted expression at known *Smchd1* targets in the *Snrpn* and *Igf2r* clusters

E14.5 *Smchd1*^{matΔ} placentae displayed partial loss of imprinted expression at previously known *Smchd1*-sensitive genes in the *Snrpn* cluster (*Magel2*, *Peg12* and *Ndn*, (Figure 3.7a) and the *Igf2r-Airn* cluster (*Slc22a3* (Figure 3.7b), shown as an increase in the proportion of the RNA derived from the silent allele as a proportion of the total expression, compared with controls. In addition, we observed that *Pde10a* displayed partial loss of imprinted expression in the *Smchd1*^{matΔ} samples, a new addition to the *Igf2r-Airn* cluster as well as the list of *Smchd1*-sensitive genes [Gendrel et al., 2013, Mould et al., 2013, Andergassen et al., 2019].

It is possible that the subtle relaxation of imprinting may arise due to the effect of the samples being heterozygous for *Smchd1*, rather than a maternal effect. In order to differentiate a maternal effect from haploinsufficiency for *Smchd1* I carried out allele specific RNA-seq in *Smchd1* heterozygous E14.5 placentae. I collected placentae that derived a null allele from the sire (labelled het in the Figures) together with litter-mate wildtype controls (labelled wt in Figures). These controls confirmed that the effect we observe at these known imprinted genes is indeed a heterozygous effect as they showed similar partial relaxation of imprinting similar to the maternal null samples (Figure 3.7a and b). Analysis of the E9.5 embryo data also showed an effect at *Snrpn* cluster genes, *Peg12* and *Ndn* (Figure 3.7c),

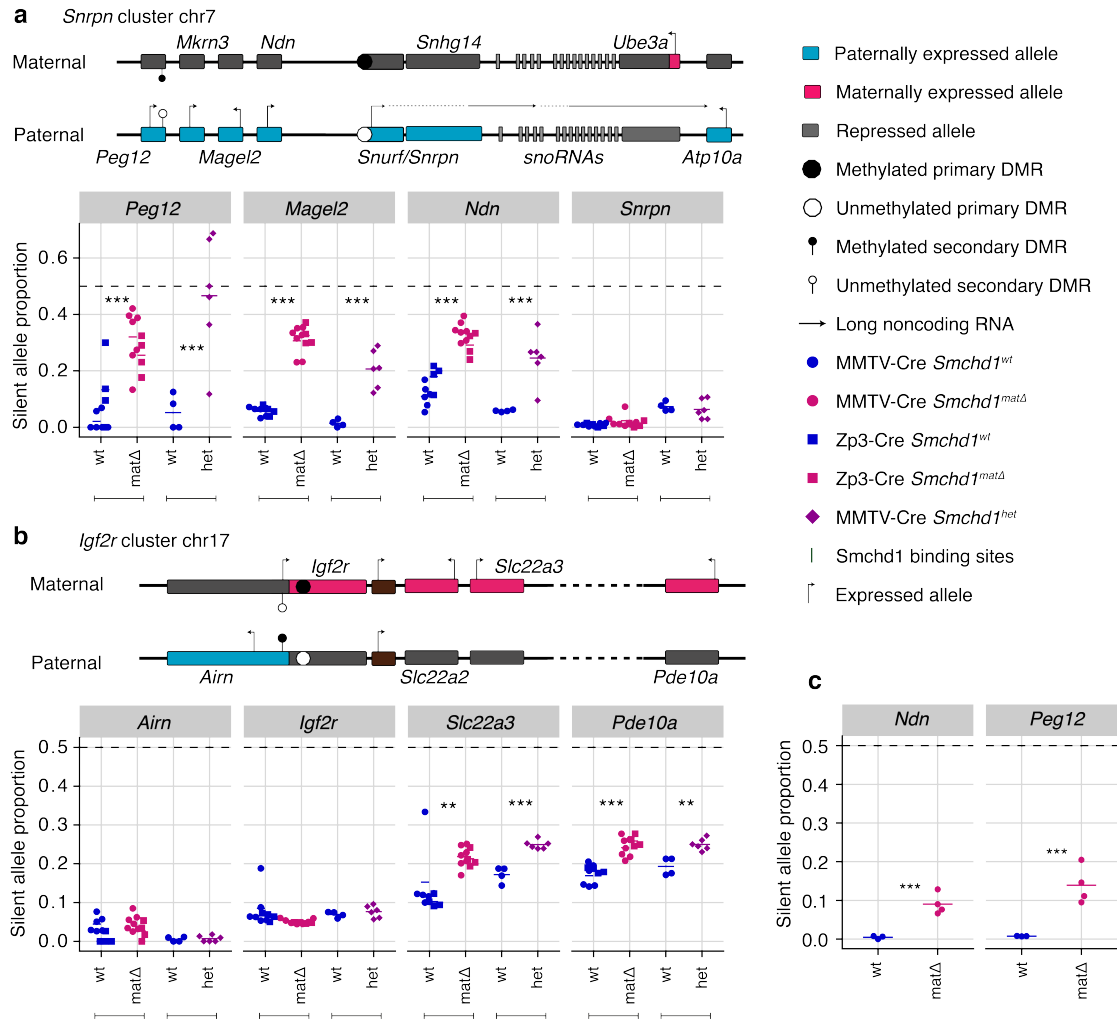


FIGURE 3.7: Maternal deletion of *Smchd1* does not affect imprinting at known *Smchd1*-sensitive clusters.

Expression of the silent allele as a proportion of total expression of the gene at the **a.** *Snrpn* cluster genes on chromosome 7 and **b.** *Igf2r* cluster genes on chromosome 17, in E14.5 *Smchd1* maternal null placentae compared with controls (*Smchd1*^{matΔ} and *Smchd1*^{wt}), or heterozygous placental samples where the deleted allele was inherited from the father, compared with controls (*Smchd1*^{het} and *Smchd1*^{wt}). **c.** Expression of the silent allele as a proportion of total expression of the gene at the *Snrpn* cluster genes *Ndn* and *Peg12*, in E9.5 *Smchd1* maternal null embryos. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, when the difference in mean silent allele proportions between genotypes is of at least 5%. RNA-seq: all replicates are shown as individual data points $n = 4-7$, except for the E9.5 embryo samples where $n = 3-4$ for maternal null (matΔ) and control (wt); MMTV-cre in circles, Zp3-cre in squares, and MMTV-Cre heterozygous (het) in diamonds.

that, based on the placental data, we assume is also likely to be a heterozygote effect. Compared to placenta, significantly fewer genes are imprinted in the embryo, therefore only placental data is shown from this point onwards.

While these data demonstrate that imprinting at the *Snrpn* and *Igf2r-Airn* clusters do not require maternal Smchd1, these experiments revealed a new Smchd1-sensitive gene in the *Igf2r-Airn* cluster. We also exposed a dose-dependent role for zygotic Smchd1 at these two imprinted clusters. These data are consistent with the fact that Smchd1 was first identified due to a heterozygous effect on transgene silencing, later extended also to a heterozygous effect on X chromosome inactivation [Blewitt et al., 2005, 2008]

3.3.8 Genes within the Kcnq1 cluster are also under the control of zygotic Smchd1

Next we looked at all other known canonical imprinted genes in our RNA-seq data from maternal null and heterozygous E14.5 placental samples that were not previously known to be Smchd1 dependent. Similar to the canonical clusters, again we observed partial relaxation of imprinting in both maternal null and heterozygous samples at the *Kcnq1* cluster as well as the imprinted singleton gene *Ascl2*, both on chromosome 7. However, for the gene *Tssc4* only the maternal null showed relaxation of imprinting (Figure 3.8). The expression difference observed for similar genotypes are due to strain effects (Zp3-cre has pure C57Bl6 background, MMTV-cre has FVB strain background) and sample numbers. These data highlight a previously unknown role for zygotic Smchd1 at the *Kcnq1* canonical imprinted cluster. Interestingly the maternal effect at *Tssc4* suggests a role for both maternal and zygotic Smchd1 at clustered imprinted genes, though a minor role for maternal Smchd1 given we only found one canonical imprinted gene sensitive to relaxation of maternal Smchd1.

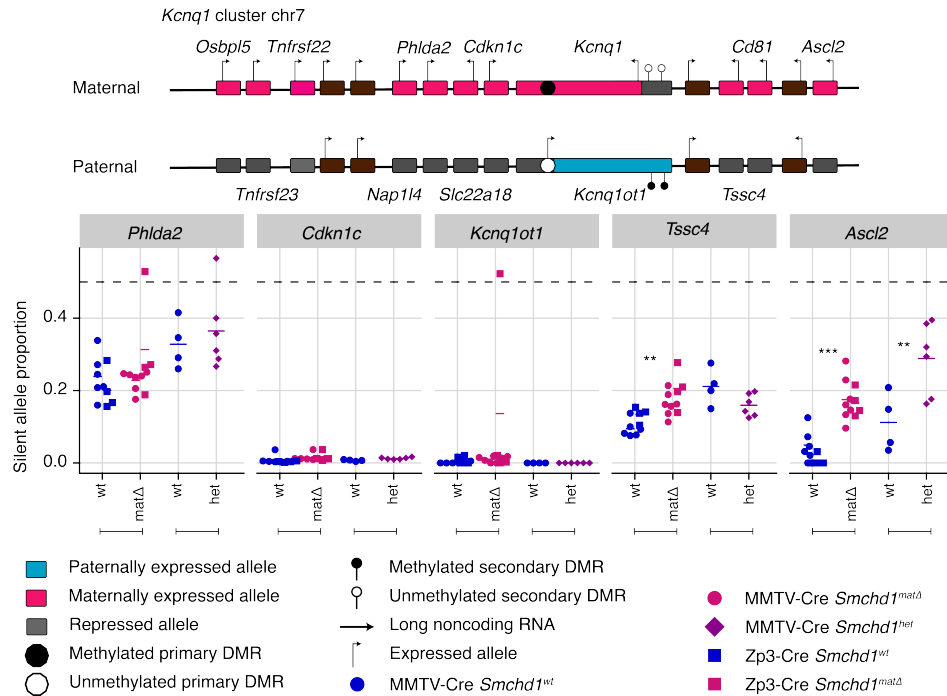


FIGURE 3.8: **Haploinsufficiency for *Smchd1* results in relaxation of imprinting at the *Kcnq1* imprinted cluster**

Expression of the silent allele as a proportion of total expression of the gene, obtained by allele-specific RNA-seq from the embryonic portion of the placenta of *Smchd1* maternal null conceptuses. Expression data for *Kcnq1* cluster genes on chromosome 7 in E14.5 *Smchd1* maternal null placentae compared with controls (*Smchd1*^{matΔ} and *Smchd1*^{wt}), or heterozygous placental samples where the deleted allele was inherited from the father, compared with controls (*Smchd1*^{het} and *Smchd1*^{wt}) * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, when the difference in silent allele proportions is of at least 5%. RNA-seq: all replicates are shown as individual data points $n = 4-7$ for Maternal null (matΔ) and control (wt); MMTV-cre in circles, Zp3-cre in squares, and MMTV-Cre heterozygous (het) in diamonds.

3.3.9 Maternal deletion of *Smchd1* does not affect DNA methylation at germline or secondary DMRs for canonical imprinted genes

While we did not observe a maternal effect at canonical imprinted genes in E14.5 placentae and E9.5 mouse embryo, the partial relaxation of imprinting meant we were interested in investigating the cause of this effect. To this end, Mrs Tamara Beck produced reduced representation bisulfite sequencing libraries from E14.5 maternal null and control placental samples. Dr Quentin Gouil analysed DNA methylation at differentially methylated regions (DMRs) in this data. We found

that deletion of maternal *Smchd1* had very little effect on global CpG island methylation. Only 33 and 42 CpG islands out of 13,840 showed appreciable hypo- and hyper-methylation, respectively (Figure 3.9a). We found that in spite of partial loss of imprinted expression in *Smchd1*^{matΔ} samples, DNA methylation at the *Snrpn* and *Igf2r-Airn* cluster germline DMR was maintained (Figure 3.9b). DNA methylation was also unchanged in the *Smchd1* maternal null samples at the secondary DMRs at the *Igf2r-Airn* and *Kcnq1* clusters (Figure 3.9c). The Blewitt lab and others previously showed that *Smchd1* null embryos show hypomethylation of the *Peg12* secondary DMR [Gendrel et al., 2013, Mould et al., 2013]; however we could not draw direct comparisons to these previous works as this region is not methylated in placentae (Figure 3.9c). Since we did not observe any significant effect on DNA methylation we can conclude that relaxation of imprinting occurs independently of a change in DNA methylation at these canonical imprinted genes, at least at this stage of development.

3.3.10 Novel imprinted genes under the control of maternal *Smchd1*

As we did not find a role for maternal *Smchd1* at canonical imprinted genes, we explored our allele specific RNA-seq data from maternal null and heterozygous E14.5 placental samples to look into non-canonical imprinted genes, those that are regulated via H3K27me3 germline imprints. Many of these genes only showed partial imprinting by this stage of development, as would be predicted by earlier work [Wang et al., 2011, Hanna et al., 2019]. We found that loss of maternal *Smchd1* does indeed result in relaxation of imprinted expression at some non-canonical imprinted genes (Figure 3.10a). Interestingly, heterozygous control samples showed no difference in gene expression, suggesting that these non-canonical imprinted genes are subject to control by maternal *Smchd1*. The non-canonical imprinted genes that showed relaxation of imprinting due to deletion of maternal *Smchd1*

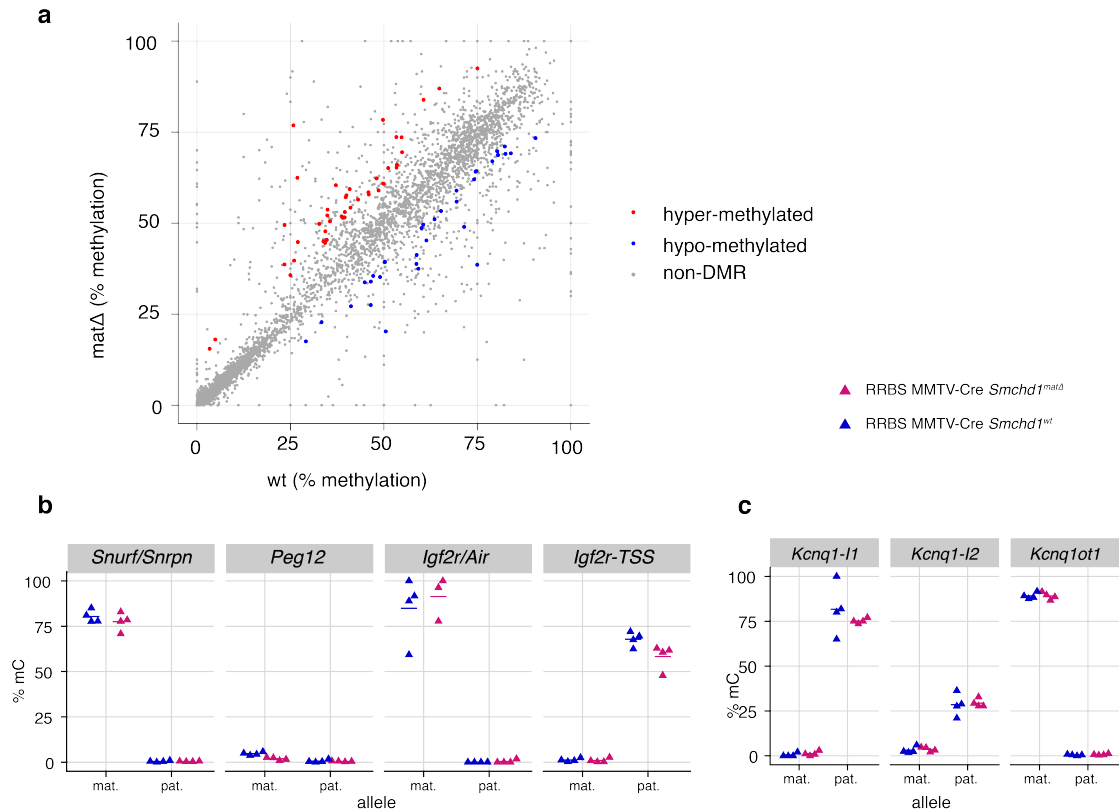


FIGURE 3.9: Maternal deletion of *Smchd1* does not affect DNA methylation at DMRs associated with canonical imprinted clusters sensitive to loss of *Smchd1*

a. Global CpG methylation, 33 hypomethylated CpG sites in blue circles and 42 hypermethylated CpG sites in red circles. Grey circles indicate CpGs that were not differentially methylated between the two genotypes **b-c.** Percentage methylation (% mC) for each parental allele at the **b.** germline DMRs and **c.** secondary DMRs at *Snrpn*, *Igf2r*, and *Kcnq1* clusters in *Smchd1* maternal null (*matΔ*) and wild-type (*wt*) placental samples. *Kcnq1-I1*: *Kcnq1-Intergenic1*; *Kcnq1-I2*: *Kcnq1-Intergenic2*. RRBS $n = 4$ MMTV-Cre for both *matΔ* and *wt* E14.5 placental samples.

included *Jade1* and *Platr4* that are co-located on chromosome 3, *Sfmbt2* on chromosome 2, *Smoc1* on chromosome 12, *Epop* on chromosome 11, and *Spp1* on chromosome 5 (Figure 3.10a). Also of note that the relaxation of imprinting observed was variable across samples as the event of imprint relaxation is usually stochastic between samples, as previously seen in literature [Alexander et al., 2015]. Similar to canonical imprinted genes here again we observed no hypomethylation at the *Sfmbt2* secondary DMR. But we did observe hypomethylation at the *Jade1* secondary DMR (Figure 3.10b). These data suggest that *Smchd1* is a maternal effect gene required for imprinted expression of at least six non-canonical imprinted

genes, that can exert its effect without altering the secondary DMR methylation.

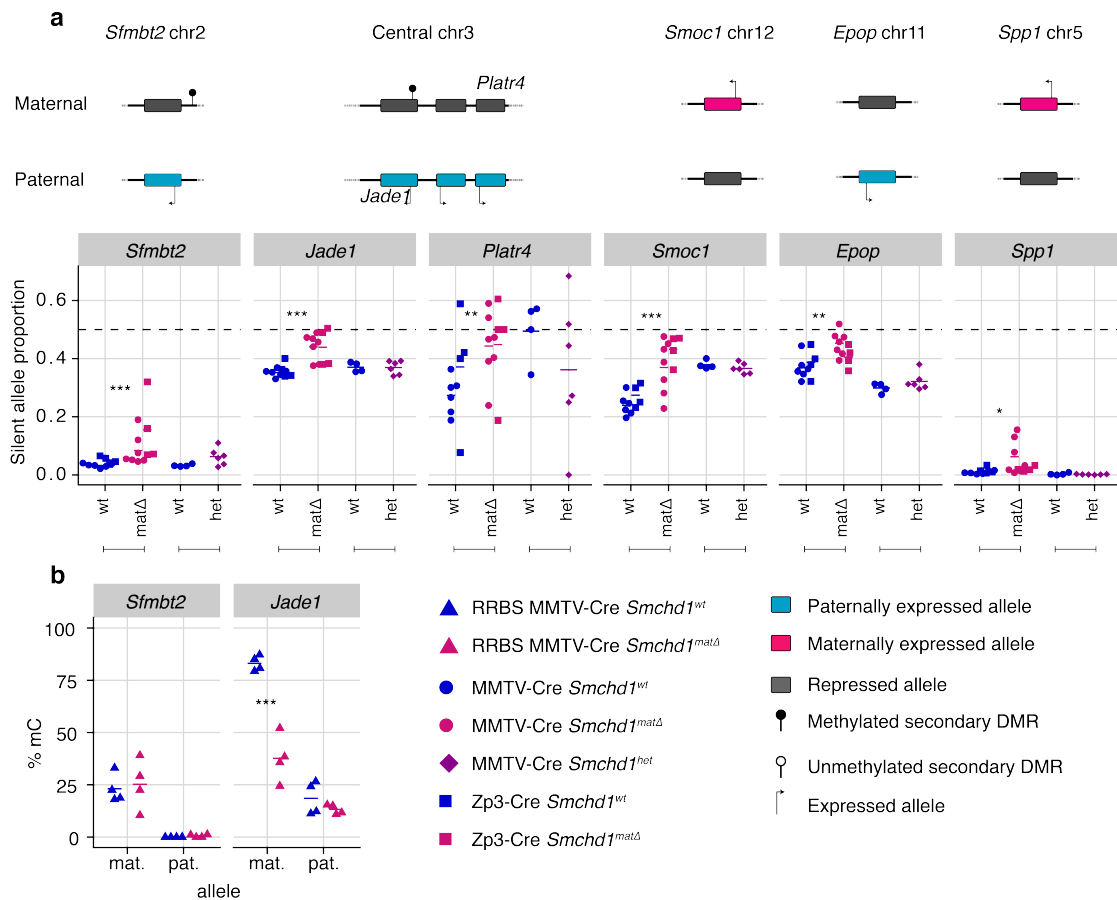


FIGURE 3.10: Maternal deletion of *Smchd1* results in relaxation of imprinting non-canonical imprinted genes

Expression of the silent allele as a proportion of total expression of the gene, obtained by allele-specific RNA-seq from the embryonic portion of the placenta of *Smchd1* maternal null and control, heterozygous and control conceptuses. **a.** at non-canonical imprinted genes: *Sfmbt2*, *Jade1*, *Platr4*, and *Smoc1*. **b.** Percentage methylation (% mC) for each parental allele at the DMRs for clusters and genes and samples shown in **a.** * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, when the difference in silent allele proportions is of at least 5%. RNA-seq: all replicates are shown as individual data points $n = 4-7$ for Maternal null (matΔ) and control (wt); MMTV-cre in circles, Zp3-cre in squares, and MMTV-Cre heterozygous (het) in diamonds. RRBS $n = 4$ MMTV-Cre for both matΔ and wt E14.5 placental samples.

3.3.11 Both maternal and zygotic *Smchd1* contribute to canonical imprinted gene expression

Previously, it was shown that either maternal or zygotic deletion of *Zfp57* and *Trim28* resulted in partially penetrant loss of imprinting, whereas deletion of both maternal and zygotic forms increased the penetrance; providing evidence that both the maternal and zygotic content of these proteins play a role in imprinted gene expression [Li et al., 2008, Alexander et al., 2015]. Therefore to further our understanding on the contribution of maternal versus zygotic *Smchd1* at imprinted genes, I produced wild-type, *Smchd1*^{matΔ}, zygotic-deleted (*Smchd1*^{zygΔ}) or maternal-and-zygotic-deleted for *Smchd1* (*Smchd1*^{matzygΔ}) E9.5 embryo and E14.5 embryonic portion of placentae samples (Figure 3.10a). In order to do this we produced an F1 line of mice from CAST/EiJ strain dams mated with C57BL/6 *Smchd1*^{Del/+} sires. These F1 mice were mated with C57BL/6 *Smchd1*^{Del/+} dams or *Smchd1*^{Del/fl}; MMTV-Cre^{T/+} dams. This specific breeding protocol resulted in progeny containing approximately half of the imprinted clusters that had CAST/EiJ allele in trans to C57BL/6. Therefore, only some genes contained the polymorphisms between parental alleles required for allele-specific analyses and some genes had no informative reads. This was a major limitation for our capability to fully analyse and interpret the acquired data.

Dr Quentin Gouil analysed the RNA-seq data. We did not identify any additional *Smchd1*-sensitive imprinted genes in *Smchd1*^{zygΔ} or *Smchd1*^{matzygΔ} samples compared with the maternal deletion alone (Figure 3.11). In general, we observed a stronger effect of zygotic compared with maternal *Smchd1* deletion or heterozygosity for *Smchd1* on imprinted expression. Canonical imprinted genes at the *Snrpn* cluster *Peg12*, *Magel2* and *Ndn* showed complete loss of imprinted expression in *Smchd1*^{zygΔ} and *Smchd1*^{matzygΔ} E9.5 embryos and E14.5 placentae (Figure 3.11b and Figure 3.12a), compared with only partial loss of imprinted expression in *Smchd1*^{matΔ} and heterozygous samples (Figure 3.11a). For the *Igf2r-Airn* cluster *Slc22a3* and *Pde10a* showed significantly higher loss of imprinted expression compared to *Smchd1*^{matΔ} and heterozygous samples in placentae (Figure 3.12b).

For the *Kcnq1* cluster we did not have sufficient samples with informative SNPs to see whether there is any effect due to *Smchd1*^{zygΔ} and *Smchd1*^{matzygΔ} deletions. Overall these data suggest that zygotic *Smchd1* plays the most important role at canonical imprinted genes, consistent with the heterozygous rather than maternal effect described above.

3.3.12 Both maternal and zygotic *Smchd1* contribute to non-canonical imprinted gene expression

At the non-canonical imprinted genes, in general we did not observe enhanced loss of imprinting from the zygotic null placentae with the exception of *Sfmbt2*. It is important to note that we wouldn't have the capacity to see more complete loss of imprinting in *Jade1*, *Platr4*, *Smoc1* and *Epop* as these genes are only partially imprinted in these samples. At *Sfmbt2* we did observe strong imprinting and significantly stronger loss of imprinting in the *Smchd1*^{zygΔ} and *Smchd1*^{matzygΔ} compared with *Smchd1*^{matΔ} placental samples (Figure 3.13). By contrast for *Spp1*, which also displayed strong imprinted expression, zygotic deletion of *Smchd1* did not result in more complete loss of imprinting than maternal deletion. Therefore, these data suggest that maternal *Smchd1* may have a more prominent role at some non-canonical imprinted genes, but at *Sfmbt2* zygotic *Smchd1* also plays an important role.

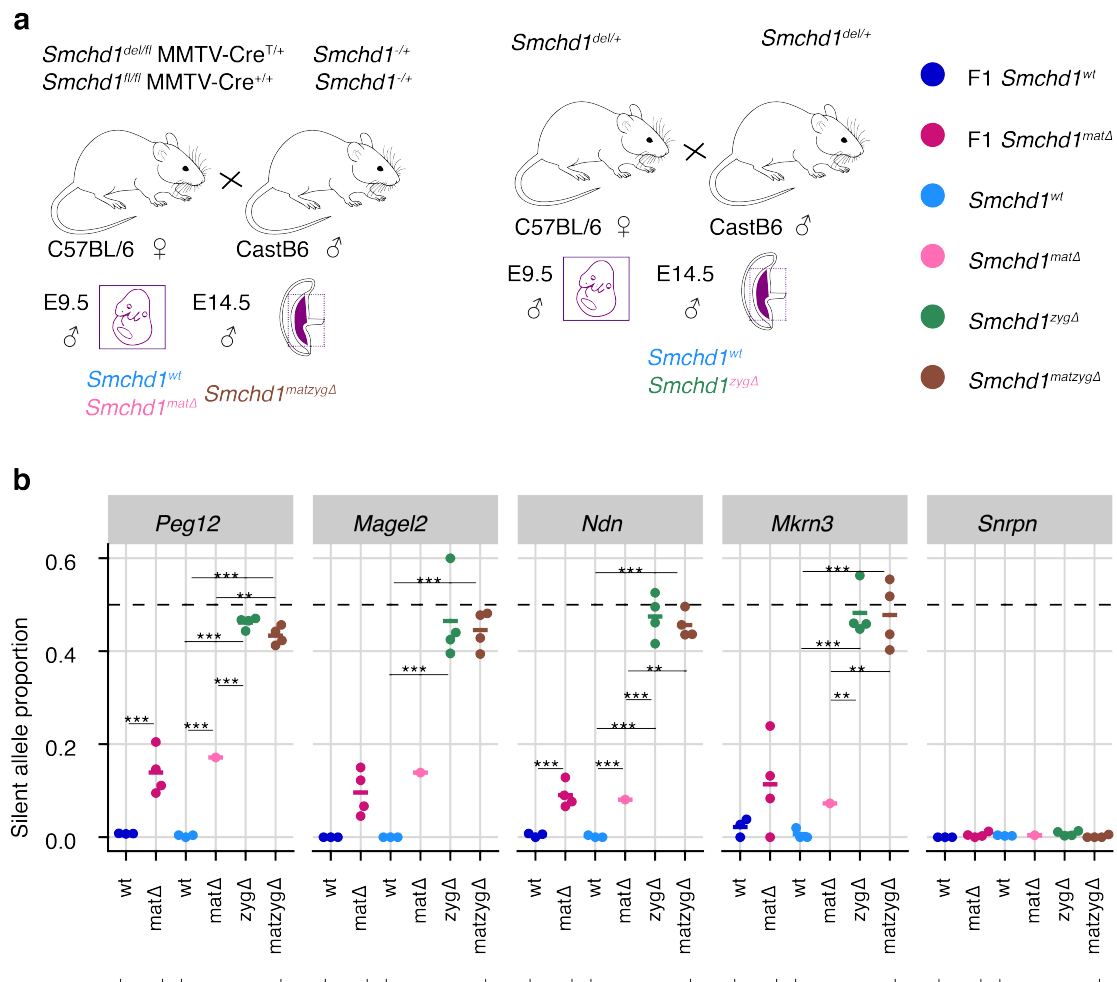


FIGURE 3.11: Zygotic deletion of *Smchd1* results in complete loss of imprinting at proximal genes in the *Snrpn* cluster in embryos.

a. Schematic for breeding scheme to generate oocyte (matΔ), zygotic (zygΔ), oocyte-and-zygotic deletion (matzygΔ) of *Smchd1*. Strain background is shown underneath each parent. Genotype is shown above each parent. MMTV-Cre *Smchd1* oocyte-deleted data (F1 wt, F1 matΔ) along with samples produced to compare *Smchd1* wild-type (wt), oocyte-deleted (matΔ), zygote-deleted (zygΔ), and oocyte-and-zygote-deleted (matzygΔ) genotypes. Samples from the E9.5 embryo and expression of the silent allele is shown as a proportion of total expression of the gene, obtained by allele-specific RNA-seq. **b.** *Snrpn* cluster genes. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, when the difference in silent allele proportions is at least 5%. RNA-seq for the zygotic deletion samples: n = 3 wt, n = 1 matΔ, n = 4 zygΔ, n = 4 matzygΔ for E9.5 MMTV-Cre embryo.

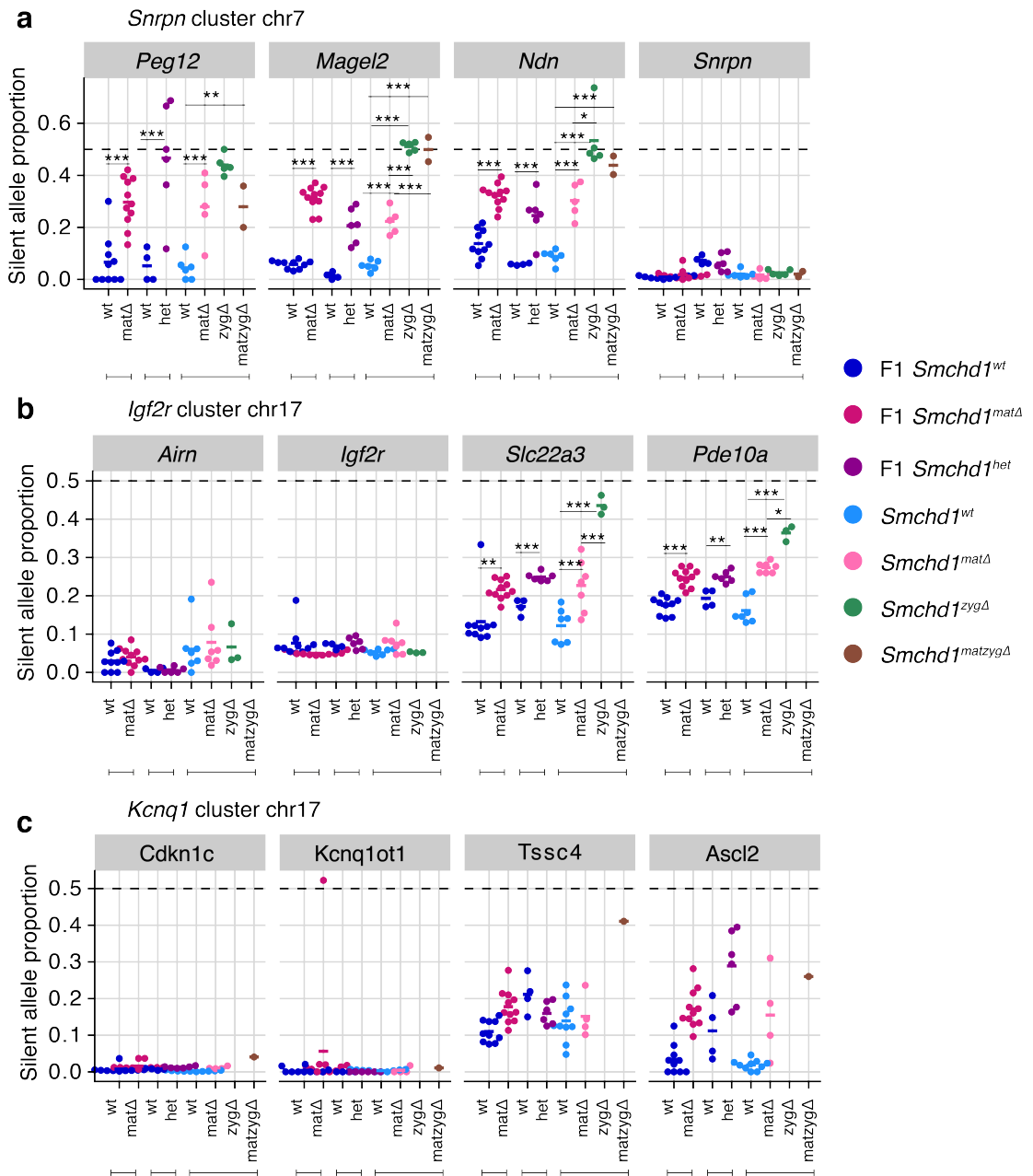


FIGURE 3.12: Zygotic deletion of *Smchd1* results in more severe loss of imprinting at canonical imprinted genes sensitive to heterozygosity for *Smchd1* in the placenta

MMTV-Cre *Smchd1* oocyte-deleted data (F1 wt, F1 *matΔ*) and heterozygous data (wt, *het*) from Figs. 3.6 and 3.9, along with samples produced to compare *Smchd1* wild-type (wt), oocyte-deleted (*matΔ*), zygote-deleted (*zygΔ*), and oocyte-and-zygote-deleted (*matzygΔ*) genotypes. Samples from the embryonic portion of the placenta and expression of the silent allele is shown as a proportion of total expression of the gene, obtained by allele-specific RNA-seq. **a.** *Snrpn* cluster genes. **b.** *Igf2r*-*Airn* cluster genes. **c.** *Sfmbt2*, *Jade1*, *Platr4*, *Smoc1*, *Epop* and *Spp1* genes. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, when the difference in silent allele proportions is at least 5%. RNA-seq for the zygotic deletion samples: $n = 13$ wt, $n = 7$ *matΔ*, $n = 8$ *zygΔ*, $n = 6$ *matzygΔ* for E14.5 MMTV-Cre placentae.

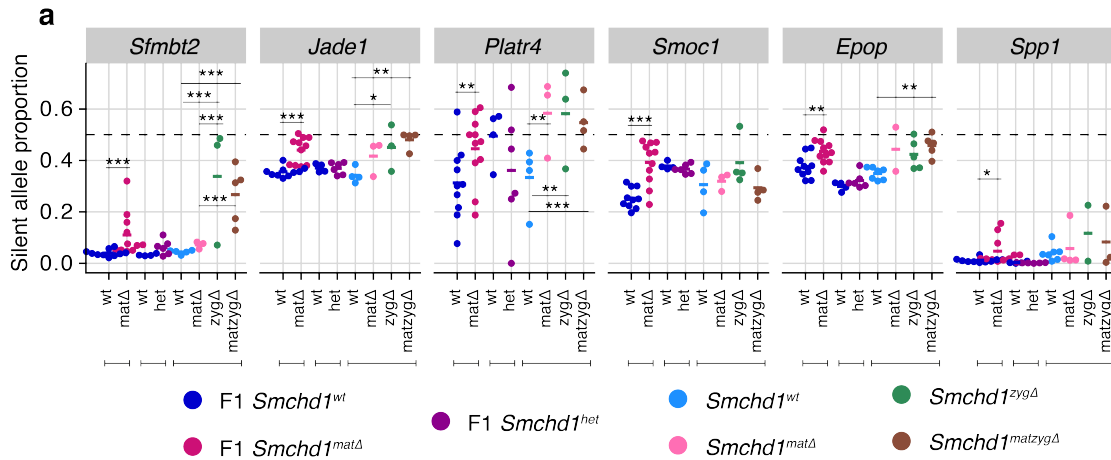


FIGURE 3.13: **Zygotic deletion of *Smchd1* has variable effects on non-canonical imprinted genes sensitive to maternal *Smchd1***

MMTV-Cre *Smchd1* oocyte-deleted data (F1 wt, F1 matΔ) and heterozygous data (wt), het) from Figs.3.6 and 3.9, along with samples produced to compare *Smchd1* wild-type (wt), oocyte-deleted (matΔ), zygote-deleted (zygΔ), and oocyte-and-zygote-deleted (matzygΔ) genotypes. Samples from the embryonic portion of the placenta and expression of the silent allele is shown as a proportion of total expression of the gene, obtained by allele-specific RNA-seq for *Sfmbt2*, *Jade1*, *Platr4*, *Smoc1*, *Epop* and *Spp1* genes. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, when the difference in silent allele proportions is at least 5%. RNA-seq for the zygotic deletion samples: $n = 13$ wt, $n = 7$ matΔ, $n = 8$ zygΔ, $n = 6$ matzygΔ for E14.5 MMTV-Cre placentae.

3.3.13 Maternal *Smchd1* sets up imprinted expression pre-implantation

Previous work demonstrates that non-canonical imprinted gene expression is strongest during preimplantation embryo development until gastrulation [Inoue et al., 2017b]. In the previous section I found that maternal *Smchd1* is present until the 16 cell (E2.75) stage of development (Figure 3.3). Therefore, in order to directly test the role of maternal *Smchd1* in imprinting of the non-canonical imprinted genes we carried out allele-specific methylome and transcriptome sequencing (M&T-seq) of *Smchd1*^{matΔ} and control 16-cell stage embryos. I collected the embryos and plated ready for treatment. Library preparation and sequencing was carried out at The University of Newcastle in collaboration with Dr Heather Lee and Ms Ellise Roper. Data were analysed by Dr Quentin Gouil.

We analysed the RNA-seq data from the 16 cell embryos. We did not find considerable differential expression genome-wide, however we found 89 genes showing imprinted expression at this stage (Appendix A). It is important to note that not all genes had informative SNPs or sufficient coverage for analysis, so these 89 genes likely represent a subset of all imprinted genes at this time. Four of these imprinted genes showed significant relaxation of imprinted expression in the *Smchd1*^{matΔ} samples (Figure 3.14 a and b). A gene is called differentially imprinted

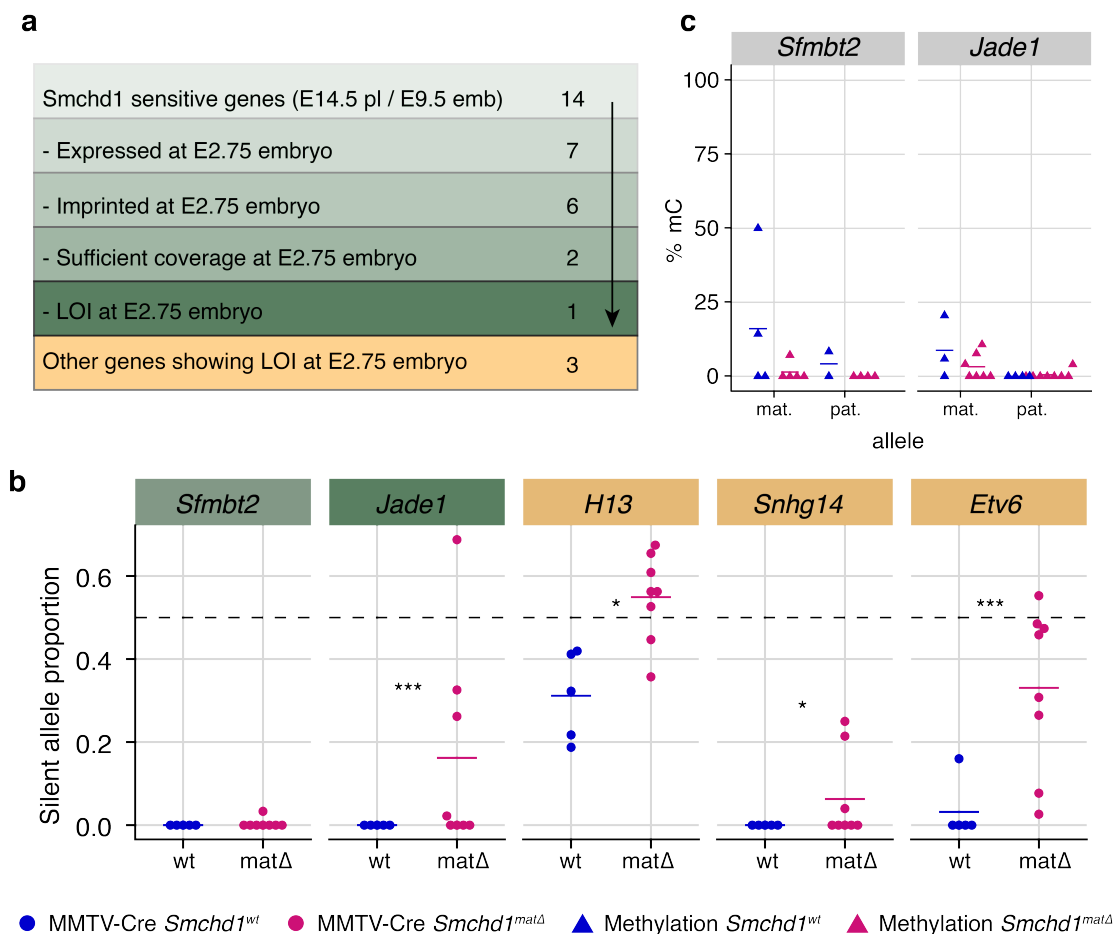


FIGURE 3.14: **Maternal *Smchd1* is required for imprinted expression of select non-canonical imprinted genes in the preimplantation embryo.**

a. Summarised analysis of *Smchd1*-sensitive imprinted genes from E14.5 placentae and E9.5 embryos in E2.75-embryo transcriptome sequencing. **b.** Expression of the silent allele as a proportion of total expression of the gene, obtained by allele-specific RNA-seq from whole 2.75-day *Smchd1* maternal null (*matΔ*) and wild-type (*wt*) embryos. **a.** Percentage methylation (% mC) on each parental allele for the DMRs in *Smchd1*-sensitive non-canonical imprinted clusters and genes in *Smchd1* maternal null (*matΔ*) compared with control (*wt*) E2.75 embryos. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, when the difference in silent allele proportions is at least 5%. $n = 5$ *wt* and $n = 8$ *matΔ* E2.75 embryos.

if the multiple-testing corrected p-value is below 0.05 and the absolute difference in silent allele proportion greater than 0.05). All four of these genes were non-canonical imprinted genes (*Jade1*, *H13*, *Snhg14*, and *Etv6*). Of these four genes, only one was identified as Smchd1-sensitive in the placenta (*Jade1*)[3.10 a](#)). The other non-canonical imprinted gene which was Smchd1 sensitive in the placenta and had sufficient coverage to be assessed in this data set was *Sfmbt2*. Interestingly, *Sfmbt2* did not show any change in imprinted expression at this stage. These data suggest Smchd1 is already necessary for imprinting of *Jade1* and 3 other non-canonical imprinted genes in the preimplantation embryo. In contrast, these data suggest that Smchd1 is required for maintenance of imprinted expression for *Sfmbt2* (Figure [3.14b](#)), which may explain why loss of zygotic Smchd1 has a more striking effect on *Sfmbt2* expression than other non-canonical imprinted genes. Potentially maternal Smchd1 perhaps establishes an epigenetic memory at *Sfmbt2* that is later required for its imprinted expression. We tested the effect of maternal Smchd1 on the secondary DMR methylation at *Sfmbt2* and *Jade16t*, and observed no significant hypomethylation (Figure [3.14c](#)), however little methylation was acquired at this timepoint. Since all four genes identified as Smchd1 sensitive in this period are non-canonical imprinted genes [[Inoue et al., 2017b](#)], these support our earlier data suggesting that maternal Smchd1 regulates non-canonical imprinted gene expression.

3.3.14 Maternal Smchd1 does not affect methylation at DMRs during the preimplantation period

Previously, I demonstrated that maternal Smchd1 was not required for canonical imprinted gene expression in E14.5 placentae. However, these samples were taken 11 days after the zygotic Smchd1 has been switched on. So, to confirm that maternal Smchd1 does not play a role at the canonical imprinted clusters we analysed the expression and DNA methylation at these clusters in the 16 cell embryo data. We did not observe loss of imprinting at any canonical imprinted genes. However, it is important to note that none of the imprinted genes where

zygotic *Smchd1* plays a role could be analysed in this dataset due to these genes not showing imprinted expression at this stage and also because of the low coverage in our M&T-seq data (Appendix A). Keeping this in mind we next analysed the DNA methylation data. When we analysed methylation at germline DMRs of the canonical imprinted genes where zygotic *Smchd1* plays a role, we found no methylation differences (Figure 3.15a). We also did not see any changes at secondary DMRs, though they are not well established at this stage (Figure 3.15 b). These data suggest that even at this stage of development, maternal *Smchd1* is not involved in controlling the canonical imprinted genes. No change in germline DMR methylation is consistent with the lack of *Smchd1* binding reported at the germline DMRs in embryonic cells [Chen et al., 2015, Jansz et al., 2018a]. It also aligns well with the relaxation of imprinting only observed at a subset of genes within the *Snrpn*, *Kcnq1* and *Igf2r-Airn* clusters, despite them sharing common DNA methylation-dependent imprinting control regions.

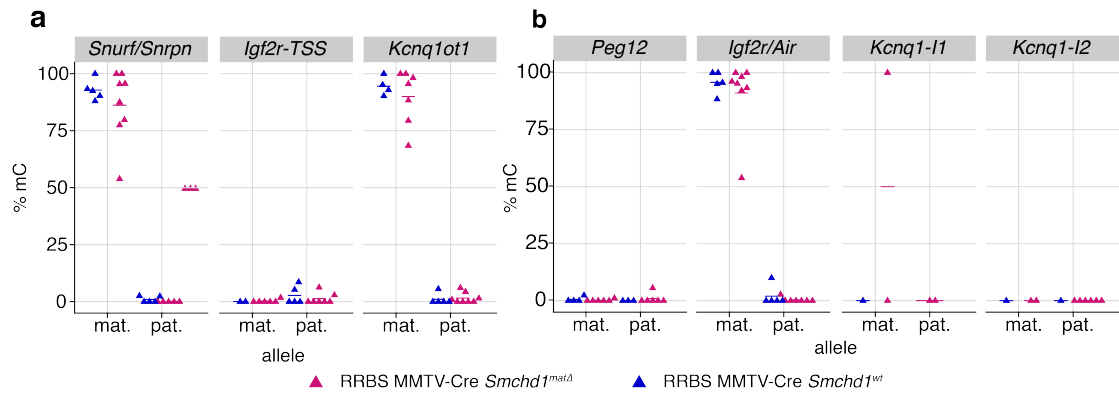


FIGURE 3.15: Maternal *Smchd1* does not affect DMR methylation at canonical imprinted genes in the preimplantation period

Percentage methylation (% mC) on each parental allele for the **a.** germline DMRs and **b.** secondary DMRs in *Smchd1*-sensitive canonical imprinted clusters and genes in *Smchd1* maternal null (*matΔ*) compared with control (*wt*) E2.75 embryos. *Kcnq1-I1*: *Kcnq1-Intergenic1*; *Kcnq1-I2*: *Kcnq1-Intergenic2*. $n = 5$ *wt* and $n = 8$ *matΔ* E2.75 embryos.

3.4 Discussion

Epigenetic regulation during early embryonic development is essential not only to produce a viable embryo but also for the precise control of gene expression. Epigenetic modifier Smchd1 plays a crucial role in development to silence the inactive X chromosome as well as other autosomal genes. Although data to date explore the function of zygotic Smchd1, we were yet to understand the function of maternally derived Smchd1. Here, I have presented strong data in support of Smchd1 being a maternal effect gene. Using various imaging techniques I have shown that maternal Smchd1 is present throughout oogenesis and persists until the 16-cell stage of embryo development after which zygotic Smchd1 is activated. This period from late oogenesis to the 16-cell stage is a long time period for a maternal effect protein to persist. The long presence of maternal Smchd1, until E2.75 (enough to be detected by microscopy as a native signal) along with its presence in the nucleus suggested that maternal Smchd1 had the potential to play an important role in gene regulation.

I decided to look at Smchd1's role on imprinted gene expression by allele-specific genomics in E14.5 male placentae, where imprinted genes are abundantly found [Andergassen et al., 2019], and in male preimplantation embryos, before zygotic Smchd1 activates. By analysing these data allele-specifically we discovered that Smchd1 sensitive genes can be categorised into two main classes: classic imprinted genes that are present in clusters and controlled by DNA methylation at the germline DMR (canonical imprinted genes), and rather recently discovered imprinted genes that are regulated by PRC mark H3K27me3 imprints (non-canonical imprinted genes). Here I found ten canonical imprinted genes, six of which were previously unknown, that are sensitive to zygotic Smchd1. For the first time I also found seven non-canonical/other imprinted genes regulated by maternal Smchd1 (Table 3.1).

By testing heterozygous samples along with the wild type samples we were able to discount any effect due to haploinsufficiency in our maternal Smchd1 deleted samples. This also enabled us to look at the effect of zygotic Smchd1 deletion

TABLE 3.1: **Smchd1** sensitive imprinted genes in mice

Imprinted gene	Imprinted cluster and Chromosome location	Type	Sensitive tissue	Smchd1 sensitivity	Haplo-insufficiency
<i>Magel2</i>	<i>Snrpn</i> , Ch.7	Canonical	E9.5 embryo, E14.5 placenta	Zygotic	Yes
<i>Peg12</i>	<i>Snrpn</i> , Ch.7	Canonical	E9.5 embryo, E14.5 placenta	Zygotic	Yes
<i>Ndn</i>	<i>Snrpn</i> , Ch.7	Canonical	E9.5 embryo, E14.5 placenta	Zygotic	Yes
<i>Mkrn3</i>	<i>Snrpn</i> , Ch.7	Canonical	E9.5 embryo	Zygotic	Not significant
<i>Snhg14</i>	<i>Snrpn</i> , Ch.7	Canonical	E2.75 embryo	Maternal	No
<i>Slc22a3</i>	<i>Airn/Igf2r</i> , Ch.17	Canonical	E14.5 placentae	Zygotic	Yes
<i>Pde10a</i>	<i>Airn/Igf2r</i> , Ch.17	Canonical	E14.5 placentae	Zygotic	Yes
<i>Tssc4</i>	<i>Kcnq1</i> , Ch.7	Canonical	E14.5 placentae	Maternal and Zygotic	No
<i>Ascl2</i>	<i>Kcnq1</i> , Ch.7	Canonical	E14.5 placentae	Zygotic	Yes
<i>Sfmbt2</i>	Lone, Ch.2	Non-canonical	E14.5 placentae	Maternal	No
<i>Jade1</i>	Lone, Ch.3	Non-canonical	E14.5 placentae and E2.75 embryo	Maternal	No
<i>Platr4</i>	Lone, Ch.3	Non-canonical	E14.5 placentae	Maternal	No
<i>Smoc1</i>	Lone, Ch.12	Non-canonical	E14.5 placentae	Maternal	No
<i>Epop</i>	Lone, Ch.11	Unknown	E14.5 placentae	Maternal	No
<i>Spp1</i>	Lone, Ch.5	Unknown	E14.5 placentae	Maternal	No
<i>H13</i>	Lone, Ch.2	Canonical	E2.75 embryo	Maternal	No
<i>Etv6</i>	Lone, Ch.5	Non-canonical	E2.75 embryo	Maternal	No

at each of the imprinted clusters. Previously, it was known that zygotic *Smchd1* acts on genes at canonical imprinted clusters *Snrpn* (*Magel2*, *Peg12* and *Ndn*) and *Igf2r-Airn* (*Slc22a3*) [Mould et al., 2013, Gendrel et al., 2013, Chen et al., 2015]. Here I was able to reveal that zygotic *Smchd1* also acts on *Pde10a* in the

Igf2r-Airn cluster and another canonical cluster *Kcnq1 (Ascl2)*. Together these data expand the number of canonical imprinted loci that are sensitive to *Smchd1* removal.

In addition to the canonical imprinted genes, I importantly discovered a role for maternal *Smchd1* in regulating a selection of non-canonical imprinted genes, including *Sfmbt2*, *Jade1*, *Platr4*, *Smoc1*, *Epop* and *Spp1* in E14.5 placentae and *H13*, *Snhg14*, *Etv6* and *Jade1* in 16-cell preimplantation embryos. This reveals for the first time that *Smchd1* is a maternal effect gene that regulates the expression of some non-canonical imprinted genes. These results are summarized in Table 3.1

The 16-cell RNA-seq data we obtained did not have the coverage to look at canonical imprinted genes sensitive to zygotic *Smchd1*. So, in future it will be interesting to look at these canonical imprinted genes either by carrying out M&T-seq on more embryos, or using another single cell analysis system such as RNA FISH in preimplantation embryos. Our data on zygotic *Smchd1* is based on complicated genetics that limited our ability to obtain enough embryos showing SNP variance at all relevant genes. More samples would give us a better understanding of these data. Another caveat of our experiments is that certain non-canonical imprinted genes such as *Jade1*, *Platr4*, *Smoc1* and *Epop* are only partially imprinted in E14.5 placentae. Therefore, a more extensive developmental timeline, potentially using purified cell populations from the placenta would not only give more precise data for these imprinted genes but also might reveal other *Smchd1* sensitive imprinted genes at different stages. Another tissue where imprinted genes are abundantly found is the brain. Therefore, future studies could also be focused on different tissues where tissue specific imprinting is observed.

It was recently discovered that non-canonical imprints are mediated by a H3K27me3 mechanism where the primary H3K27me3 imprint is lost post implantation and through an unidentified mechanism transferred to a DNA methylation imprint in the cases where imprinted expression persists [Hanna et al., 2019]. While all the non-canonical imprinted genes sensitive to maternal *Smchd1* contain primary H3K27me3 imprints, the non-canonical genes [Inoue et al., 2017b], (*Smoc1*, *Jade1*

and *Sfmbt2* also fall into the category that contains a secondary imprint post implantation. However, the canonical imprints at the *Kcnq1* and *Igf2r-Airn* clusters along with nearby *Pde10a* and the *Magel2* region of the *Snrpn* cluster are dependent on germline imprints. These sites gain secondary H3K27me3 in post-implantation tissues [Hanna et al., 2019]. This mechanism in the *Kcnq1* and *Igf2r-Airn* clusters is driven by long ncRNAs [Sleutels et al., 2002, Nagano et al., 2008, Redrup et al., 2009, Andergassen et al., 2019]. The *Kcnq1* cluster is the largest canonical imprinted cluster described to date. While the central *Kcnq1* cluster genes are maintained via germline DMR methylation, distal genes that are *Smchd1* sensitive, *Tssc4* and *Ascl2*, are maintained via histone modifications [Lewis et al., 2004, Umlauf et al., 2004]. Therefore, multiple mechanisms are at play at this cluster and could potentially explain the maternal effect observed at the *Tssc4* canonical imprinted gene. Cumulatively, either germline or secondary H3K27me3 dependency unites the *Smchd1*-sensitive imprinted genes. This hypothesis can be tested by measuring H3K27me3 imprints in preimplantation embryos using ChIP-seq or CUT&Run in future, or even in embryonic stem cells as it is apparent from earlier published work that the H3K27me3 imprints are retained in the inner cell mass [Zheng et al., 2016]

Here for the first time I have revealed that *Smchd1* is a maternal effect gene involved in genomic imprinting at H3K27me3 dependent imprinted genes. I also propose that *Smchd1* may function secondary to H3K27me3 imprints to form an epigenetic memory. The data presented in this chapter is based on male samples. But given that the inactive X regulator, *Xist* also has an imprint that is regulated by a non-canonical imprint H3K27me3 [Inoue et al., 2017a], it will be interesting to explore whether the *Xist* imprint is affected by maternal deletion of *Smchd1* and how it may affect the imprinted XCI. I'll be exploring this avenue in chapter four.

Chapter 4

Maternal Smchd1 is required for imprinted XCI

4.1 Abstract

Paternal specific expression of the long ncRNA *Xist* is crucial in achieving imprinted X chromosome inactivation (XCI). Imprinted XCI occurs in female marsupials and some tissues in rodents, such as mouse extra-embryonic cells. Recently it was discovered that a polycomb mark of H3K27me3 on the maternal X induces imprinted expression of *Xist* [Inoue et al., 2017a]. However, the exact mechanism of how this imprinted expression is achieved is yet to be understood. In the previous chapter I discussed that Smchd1 is a maternal effect gene contributing to autosomal imprinting where we hypothesise it acts secondarily to H3K27me3 imprints. Using allele specific RNA-seq and RNA-FISH, here I show that in the absence of maternal Smchd1, both X chromosomes in females and the only X in males express *Xist* and undergo silencing in pre-implantation embryos. Interestingly, when we analysed gene expression in the placenta at E14.5, normal imprinted *Xist* expression was restored. Despite this, at E14.5 the placenta showed failed iXCI. These data suggest that maternal Smchd1 is required for regulating imprinted expression of *Xist* pre-implantation, without which imprinted XCI cannot be established

normally. These data suggest that maternal *Smchd1* plays two roles in iXCI: one in establishing imprinted *Xist* expression, and the second in established long-lived chromosome-wide silencing.

4.2 Introduction

As a form of dosage compensation, either the maternal X (Xm) or paternal X (Xp) chromosome in female mammals undergoes X inactivation mediated by the expression of the long ncRNA *Xist* [Lyon, 1963, Brockdorff et al., 1991]. However, in marsupials and in the early embryonic development of mouse, an imprinted form of X chromosome inactivation (iXCI) occurs [Cooper et al., 1971]. During iXCI the Xp is exclusively inactivated in all relevant tissues. In the case of mice this includes extra embryonic tissue that gives rise to the placenta [Okamoto et al., 2011]. Evidence shows that a repressive imprint on the Xm laid down during oogenesis prevents the long ncRNA *Xist* from being expressed, causing parent of origin specific expression of *Xist* from the Xp [Tada et al., 2000]. The repressive imprint on the Xm is not mediated by a DNA methylation imprint like those that classically control imprinting [Chiba et al., 2008]. Therefore, how exactly iXCI is established has long been unclear.

Recently, it was discovered that this DNA methylation independent, parent of origin expression of *Xist* was established by maternally deposited H3K27me3 [Inoue et al., 2017a, Harris et al., 2019], similarly to the non-canonical autosomal imprinting discussed in chapter 3. Previously, I showed that maternal *Smchd1* regulates these non-canonical imprinted genes, where we hypothesise *Smchd1* acts secondary to H3K27me3. This discovery led me to explore whether maternal *Smchd1* plays a role in regulating the imprinted expression of *Xist* thereby affecting iXCI. I employed allele specific genomics along with imaging to understand the expression of *Xist* in both male and female mouse embryos and placentae. Note all

the genomic analysis was carried out by Dr Quentin Gouil (WEHI) and the micro-CT image processing was done by Dr Lachlan Whitehead (Centre for Dynamic Imaging, WEHI).

4.3 Results

4.3.1 Upregulation of expression from the paternal X chromosome in pre-implantation embryos in the absence of maternal *Smchd1*

Previously, I discovered that zygotic *Smchd1* activates after the E2.75 (16 cell) stage of development. At this stage of development only maternal *Smchd1* is at play. At this stage, all cells in the embryo show iXCI, where the Xp is exclusively silenced [Borensztein et al., 2017]. To understand the effect of maternal *Smchd1* on X chromosome expression, I decided to carry out allele-specific methylome and transcriptome sequencing of female *Smchd1*^{matΔ} and control(*Smchd1*^{wt}) E2.75 embryos. The oocyte specific *Smchd1* deletion was achieved using MMTV-cre transgenic mice. These mice were generated by back-crossing for more than 10 generations onto the C57BL/6 (BL6) background. These dams were crossed with sires from CAST/EiJ (Cast) strain giving rise to progeny containing Bl6 (maternal) and Cast (paternal) specific SNPs so that allele-specific genomic analyses could be achieved (Figure 4.1a). I optimised the timing of embryo sampling and collected the embryos for the experiment. The preparation of libraries and single embryo methylome and transcriptome sequencing (M&T-seq) in whole embryos were carried out at the University of Newcastle in collaboration with Dr Heather Lee and Ms Ellise Roper. Data were analysed by Dr Quentin Gouil.

To understand the gene expression differences between maternal null (matΔ) and wild type (wt), we took the total counts across each chromosomes in all female embryo data separated by alleles and calculated the log fold change between the two genotypes (Figure 4.1b). First, when considering the female embryos, there

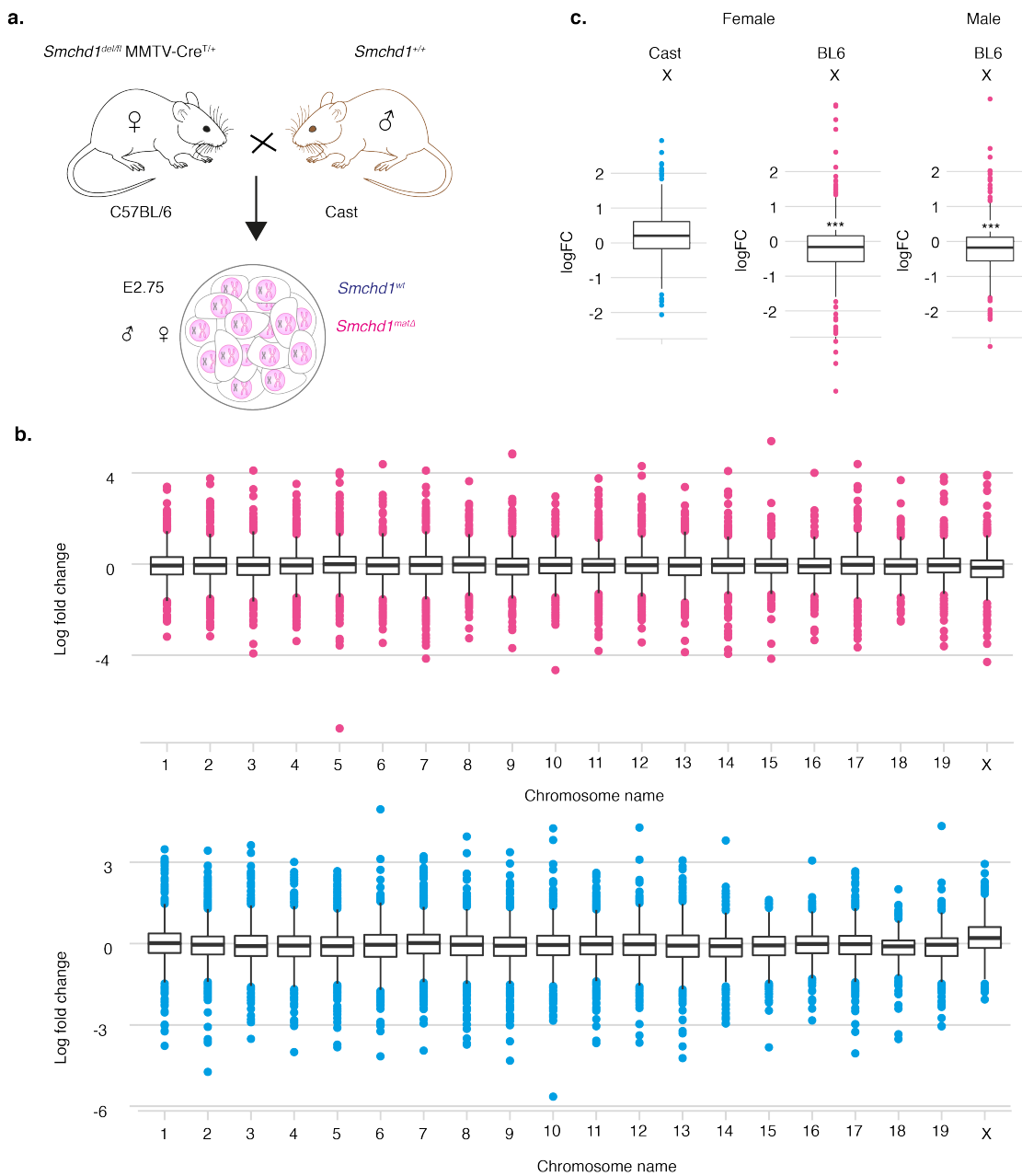


FIGURE 4.1: Maternal deletion of *Smchd1* causes aberrant expression from the paternal X chromosome in female E2.75 preimplantation embryos

a. Schematic for maternal deletion of *Smchd1*. **b.** Chromosome 1-19 and X chromosome is shown along the X axis and the Y axis shows log fold change ((matΔ vs wt) for each chromosome's gene expression data. **c.** Log fold change of gene expression from the X chromosome separated by paternal (Cast) and maternal (BL6) allele in female and male (BL6) E2.75 embryos. Log fold change is the difference between matΔ and wt based on transcriptome sequencing. For b and c, the box indicates the average gene expression across the chromosomes, the whiskers show the outliers. For *** $p < 0.001$ using student T-test. $n = 6$ wt and $n = 7$ matΔ E2.75 female embryos.

was a noticeable difference between the expression from the paternal (Cast) and maternal (Bl6) expression. Normally, the Cast X (Xp) is repressed due to iXCI. In our data we saw that when comparing the *Smchd1* maternal null and control samples, the Xp (Cast) was upregulated in the maternal null compared with control, while the Xm which is expressed at E2.75 was down-regulated (Figure 4.1c). The upregulation of Xp chromosome data suggest a failure in iXCI; however the lower expression of the Xm compared with controls was not consistent with simply failed iXCI. To investigate this further and as a control for the female data we looked at male pre-implantation embryos, where the maternal X (Xm, Bl6 - the only X present in male) is expressed. Here we saw a downregulation of Xm compared to the autosomes (Figure 4.1c), suggesting that there is some form of silencing happening of the active X chromosome that should be expressed in both males and females. This lead us to look at long ncRNA *Xist* expression in order to understand whether this irregularity in X chromosome expression is due to a failure in XCI at its very first stage.

4.3.2 Loss of maternal *Smchd1* results in expression of maternal *Xist* in both male and female E2.75 embryos

Based on the aberrant expression of both the paternal and maternal X chromosome in female *Smchd1* maternal null pre-implantation embryos, along with our prior knowledge of *Smchd1*'s role in regulating imprints, we decided to analyse the expression of *Xist* in our preimplantation embryo transcriptome data. In the female control embryos we observed *Xist* counts from the Xp (Cast), as expected as the Xp is silenced via an *Xist*-mediated mechanism in iXCI. However, we observed both Xm (Bl6) and Xp (Cast) *Xist* counts in female maternal null (mat Δ) E2.75 embryos, indicating *Xist* is inappropriately expressed from the maternal X chromosome (Figure 4.2 a) in the absence of maternal *Smchd1*. In the male control samples, as expected we did not observe expression of *Xist*. Interestingly, we observed Xm (Bl6) counts in male maternal null (mat Δ) E2.75 embryos indicating aberrant *Xist* expression in the male embryo (Figure 4.2 b). These data reveal a

failure to silence the imprinted maternal *Xist* in both males and females in the absence of maternal *Smchd1*.

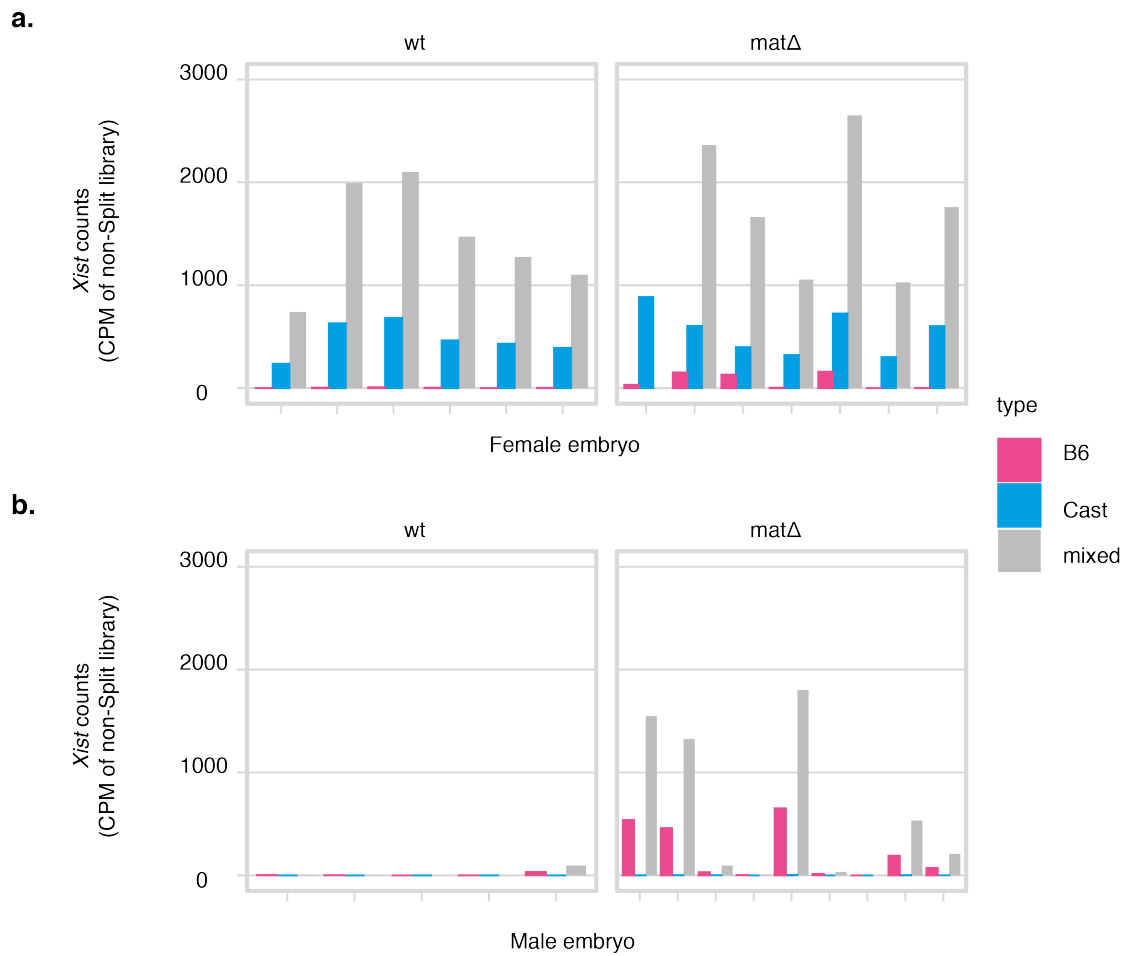


FIGURE 4.2: Maternal deletion of *Smchd1* results in *Xist* expression from the maternal allele in both male and female E2.75 preimplantation embryos

Xist Counts per million separated by maternal (BL6) allele, paternal allele (Cast), or those that do not have a SNP to define the allele (grey, mixed) in **a.** female and **b.** male E2.75 embryos *Smchd1* maternal null (*matΔ*) or wild type (*wt*) data. $n = 6$ *wt* and $n = 7$ *matΔ* E2.75 embryos. The reads are shown as a proportion of the total library size (counts per million, CPM), before being split into each allele.

4.3.3 Maternal *Smchd1* is required for faithful monoallelic *Xist* expression in preimplantation embryos

In the transcriptome sequencing data described above, reads were derived from all 16 cells of the preimplantation embryos. Such a readout gives no indication of

whether all cells in the embryo show similarly aberrant *Xist* expression patterns, how much *Xist* is expressed from each cell in the embryo, or importantly whether the expression of *Xist* in the female cells is truly biallelic. To accurately look at this we need to be able to observe what happens in each cell of E2.75 embryos using a single cell technology. To overcome the limitations of the single embryo transcriptome sequencing, I performed a new technology published by Harris *et al.* [Harris *et al.*, 2019], involving allele specific *Xist* RNA-FISH in which each allele is marked by a different probe conjugated to a fluorescent dye (Figure 4.3a) [Levesque *et al.*, 2013]. The female wild type (wt) embryos showed paternal *Xist* expression in all cells (Figure 4.3a), typical for iXCI at the 16-cell stage. But in the maternal null (mat Δ) female embryos maternal *Xist* was also expressed along with paternal *Xist*. Upon scoring of nuclei from 6 maternal null female embryos, I found that all embryos had at least 2 cells or more with *Xist* expressed in a biallelic fashion (Figure 4.3b-e). Further, and confirming the *Xist* expression from the Xm observed in male embryos (Figure 4.1b), I saw maternal *Xist* expression in mat Δ male embryos at around the same proportion as I observed biallelic expression in mat Δ female embryos. The male controls had a complete absence of *Xist* expression, as expected. Together these data show failed silencing of maternal *Xist* in the absence of maternal *Smchd1*, suggesting that maternal *Smchd1* is required in some way to bring about imprinted *Xist* expression.

Since DNA methylation plays a major role in epigenetic regulation we looked at DNA methylation in the X inactivation centre (XIC) in the preimplantation embryo data to understand if this may contribute to aberrant expression of *Xist*. When we looked at the methylation data from M&T-seq, we saw very little DNA methylation at the XIC in E2.75 embryos, and no significant differences between *Smchd1* maternal null and wild-type samples (Figure 4.4). However, it is interesting to note the suggestion of hypomethylation at the Xite region, one of the few areas with a small amount of methylation in the wild-type embryos.

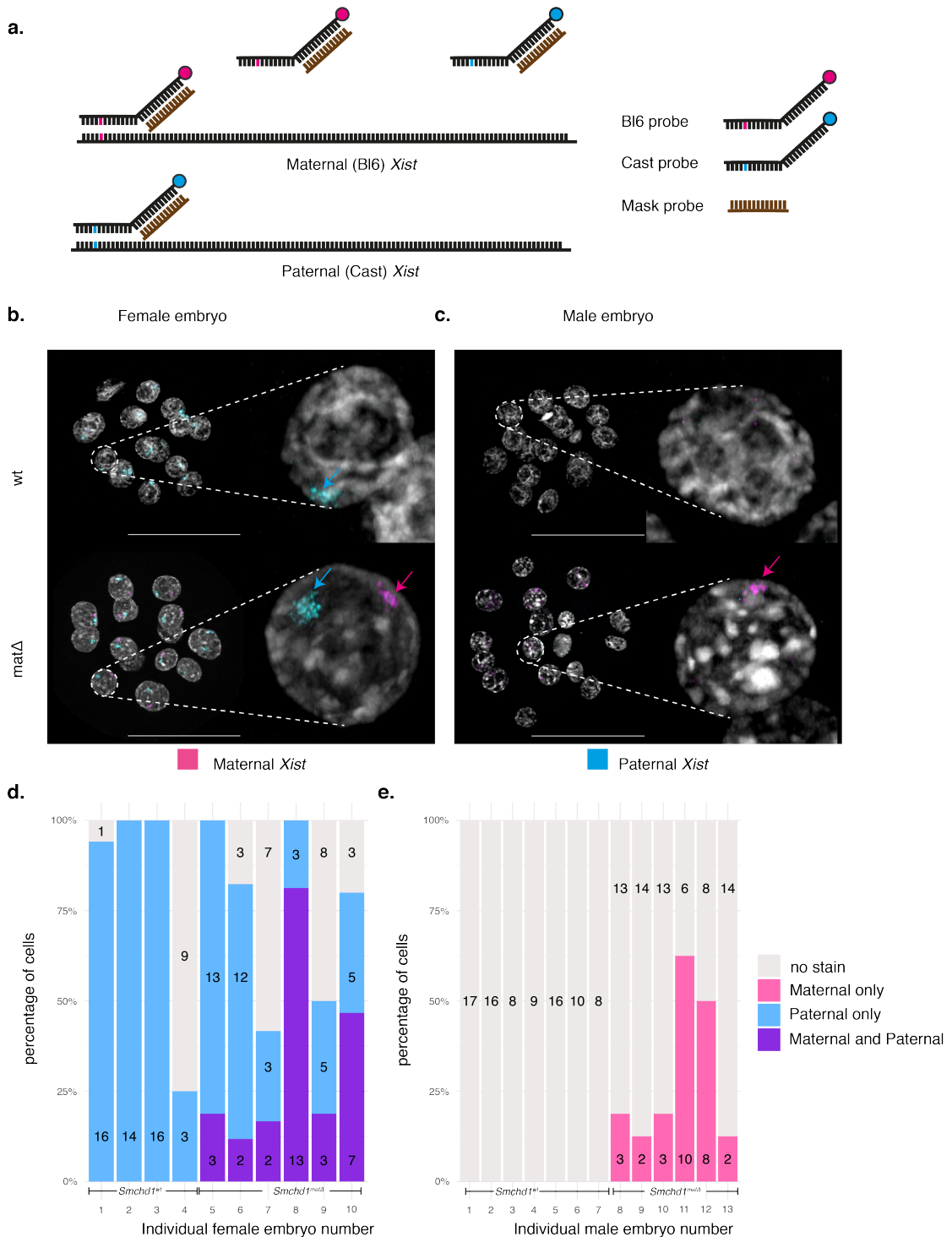


FIGURE 4.3: Maternal deletion of *Smchd1* results in biallelic expression of *Xist* from female embryos and *Xist* expression from male embryos

a. Schematic representation of the experiment - allele specific *Xist* RNA-FISH. Maternal (BL6) and paternal allele (Cast) are indicated by coloured arrows **b.** female and **c.** male wild type (wt) and maternal null (matΔ) E2.75 embryo images. **d.** Stacked bar graph showing percentage and number of cells in the embryo showing Xm, Xp or biallelic *Xist* expression for *Smchd1* **d.** wt and **e.** matΔ E2.75 embryos. Data representative of n = 4 wt female, n = 6 wt male and n = 6 female, n = 6 male matΔ E2.75 embryos. Scale bar: 50 μm.

4.3.4 Maternal deletion of *Smchd1* results in failed iXCI in the placenta

To test the role of maternal *Smchd1* in iXCI post implantation and to understand whether the effect we see pre-implantation is long lasting, we performed allele-specific bulk RNA-seq on the embryonic portion of E14.5 maternally deleted (*Smchd1*^{matΔ}) and control (*Smchd1*^{wt}) female placentae. In order to account for any effect due to haploinsufficiency I also looked at (*Smchd1*^{het}) vs (*Smchd1*^{wt}) female placentae, where the null allele was inherited from the father. Dr Quentin Gouil analysed the data. No appreciable difference was observed in expression from the active X in heterozygous or maternal null samples, compared with controls (Figure 4.4a and b, left panels). When considering expression from the paternal inactive X chromosome, we found only 31 of 397 detectable genes were upregulated (10% FDR) in heterozygous samples (Figure 4.5a, right panel), whereas 127 of 607 detectable genes (10% FDR) were upregulated in the maternal null samples (Figure

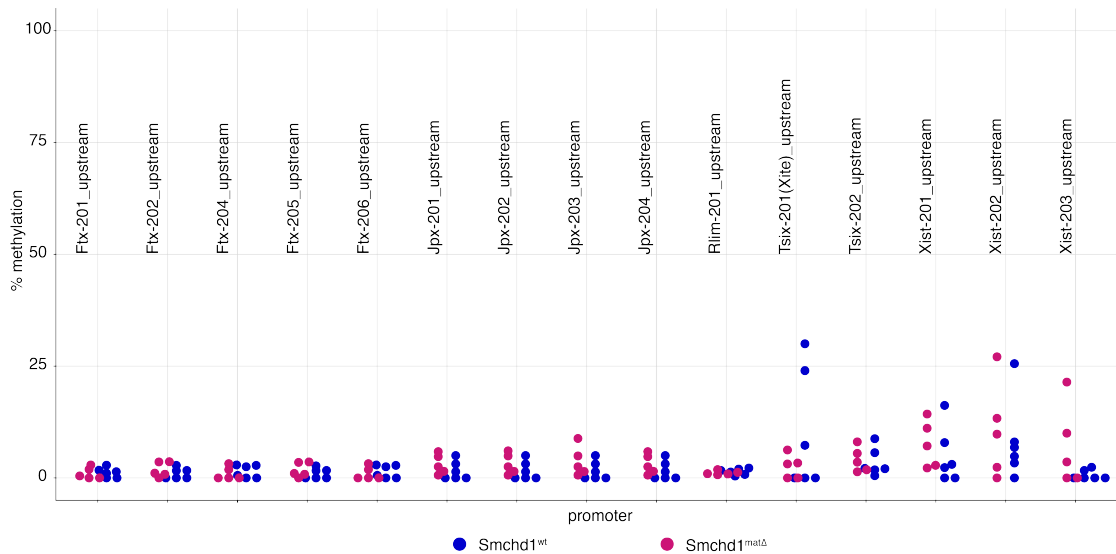


FIGURE 4.4: Maternal deletion of *Smchd1* does not affect methylation at XIC in female embryos

Percentage methylation at promoters of XIC gene transcripts. X axis shows ensemble transcript IDs for XIC genes. Promoter-centric analysis was carried out, -3kb to +1 kb for each transcript. Ftx201-206 = *Ftx*, Jpx201-204 = *Jpx*, Rlim201 = *Rlim*, Tsix201 = *Xite*, Tsix202 = *Tsix*, Xist201-203 = *Xist* gene promoter CGIs. n = 5 for Maternal null (*matΔ*) and n = 6 control (*wt*) in female E2.75 embryos.

4.5b, right panel). Out of the 127 genes, 113 were also detectable in the heterozygous samples and so we were able to compare the extent of relaxation of imprinting between the heterozygous and maternal null scenarios. Maternally deleted samples showed significantly greater loss of iXCI compared to the heterozygous samples, as indicated by a much stronger logFC in the maternal null samples as well as a greater number of genes that reach statistical significance for relaxation of imprinted expression (Figure 4.5c). These data show that the failure of XCI we see preimplantation persists for at least 11 days post zygotic *Smchd1* activation. However up-regulation of these genes does not have any apparent developmental consequences.

4.3.5 Maternal deletion of *Smchd1* does not affect *Xist* expression in E14.5 placentae

Next we wanted to understand whether the iXCI defect we observe in the placenta is due to misregulation of *Xist*. In (Figure 4.2a) we found biallelic expression of *Xist* in female embryos. It was not known whether this irregularity lasts longer than preimplantation embryo development or whether zygotic *Smchd1* or another mechanism rescues *Xist* expression later in development. Therefore, we looked at the *Xist* expression in the E14.5 allele specific RNA-seq data. We saw that *Xist* expression in oocyte null data is similar to wt data (Figure 4.5), revealing that *Xist* expression is recovered at some point between E2.75 embryo development and E14.5 placentae.

4.3.6 Maternal deletion of *Smchd1* results in a widespread effect on DNA methylation on the Xi in E14.5 placentae

To further investigate the significant up-regulation of expression from the imprinted Xi, we performed allele-specific RRBS on the same E14.5 maternally

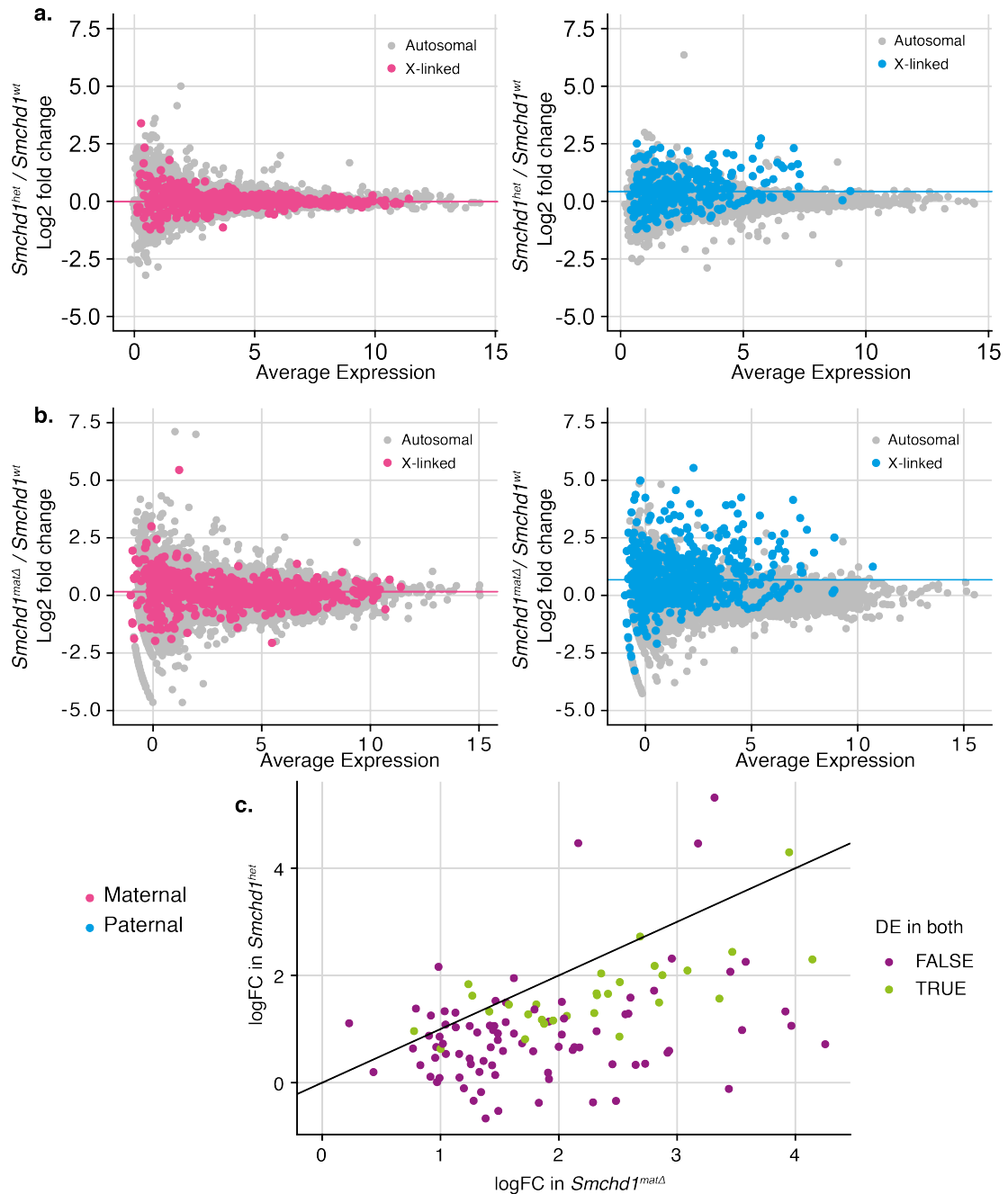


FIGURE 4.5: Maternal deletion of *Smchd1* results in failed silencing of the imprinted Xi in the placenta

Allele specific RNA-seq data for **a.** wild type (wt) vs *Smchd1* heterozygous (het) and **b.** wild type (wt) vs *Smchd1* maternal null (*matΔ*) embryonic portion of the E14.5 placenta. Data are shown split by parental allele. X-linked genes on the **paternal allele** and **maternal allele** are highlighted, all genes shown in grey. The y axis shows average gene expression and the x axis shows log fold change of gene expression between het or *matΔ* compared with wt controls. The left panels indicate the maternal alleles, right panels indicate the paternal alleles. **c.** Comparison of the log fold changes of the differentially expressed genes in the het and *matΔ* deleted samples. Genes that are differentially expressed in both are marked as **TRUE** and genes that are only expressed in one marked as **FALSE**. The diagonal line indicates same logFC under the two conditions. $n = 4$ MMTV-Cre for both *matΔ* and wt, $n = 6$ het and littermate $n = 4$ wt control E14.5 placentae.

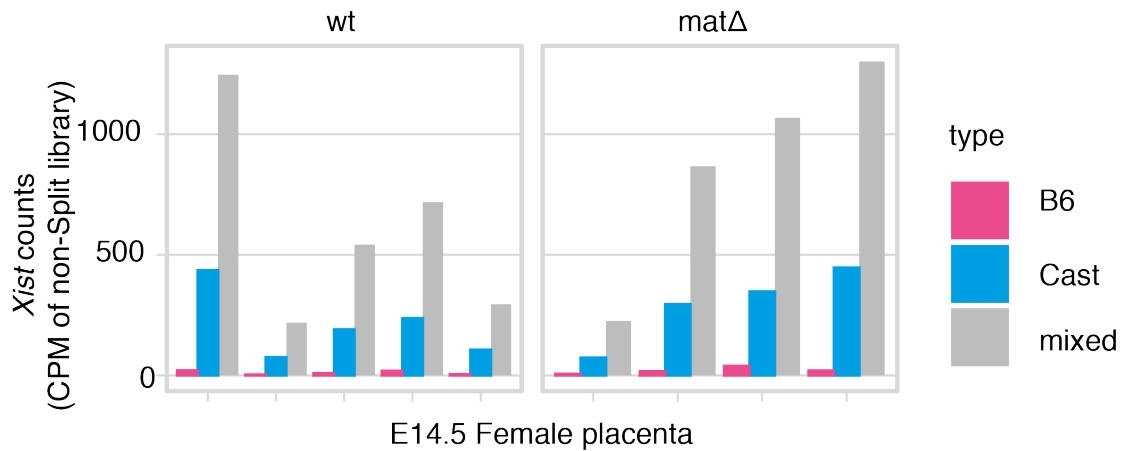


FIGURE 4.6: **Imprinted expression of *Xist* is unaffected in *Smchd1* maternal null placentae**

Total *Xist* counts separated by maternal (BL6) allele, paternal allele (Cast) or those without SNPs (mixed, grey) in female E14.5 placentae *Smchd1* maternal null (*matΔ*) or wild type (wt) data. The reads are shown as a proportion of the total library size (counts per million, CPM), before being split into each allele by SNP split. $n = 5$ wt and $n = 4$ *matΔ* Allele specific MMTV-Cre E14.5 female placental samples.

deleted and control female placentae. Ms Tamara Beck made libraries and Dr Quentin Gouil analysed the data. We observed that DNA methylation was significantly reduced at 139 of 185 CpG islands (1% FDR) on the inactive X chromosome (Figure 4.7). No significant change was observed on the active X, which was hypomethylated at CpG islands, as expected. This widespread effect on the Xi due to the loss of maternal *Smchd1* could be due to chromosome-wide effect of *Smchd1* or occur as a consequence of failed silencing. In either case, since promoter associated CpG islands on the Xi become methylated during normal development [Norris et al., 1991] following X inactivation [Grant et al., 1992], hypomethylation of the Xi in the absence of maternal *Smchd1* suggests a failure of XCI.

4.3.7 Maternal deletion of *Smchd1* does not affect methylation of CpG islands at XIC in placental samples

We next examined DNA methylation of genes specifically within the X inactivation centre (XIC), that houses *Xist* and all the DNA control elements for *Xist*. The CpG island at the *Xist* promoter was not differentially methylated (Figure 4.8).

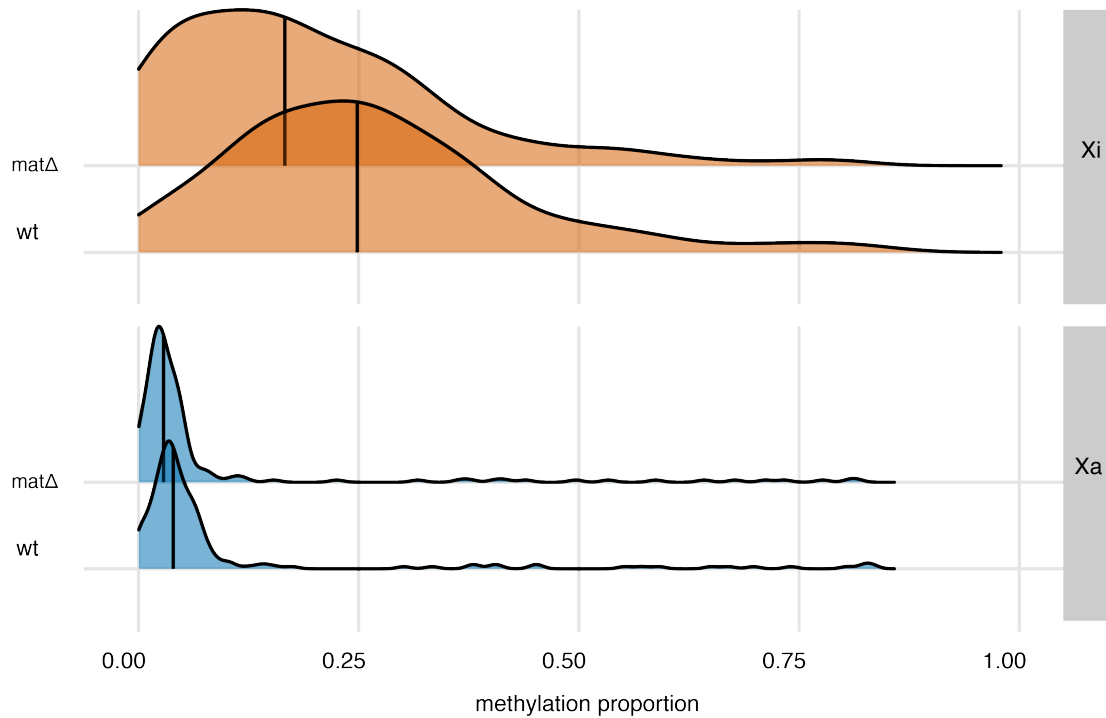


FIGURE 4.7: Maternal deletion of *Smchd1* shows significant effect on CpG island hypermethylation on the imprinted Xi

Allele specific methylation of CpG islands on the paternal inactive X (Xi) and maternal active X (Xa) in *Smchd1* wild-type (wt) and maternal null (*matΔ*) embryonic portion of E14.5 placental samples, as determined by RRBS. The proportion of methylation at each of 185 CpG islands on the X is shown in a ridgeline plot where the methylation density distribution is shown across the chromosome. Differential methylation at CpG islands was determined by fitting a logistic regression on methylated and unmethylated counts. RRBS; wt $n = 4$, *matΔ* $n = 3$ MMTV-Cre E14.5 female placental samples.

Interestingly, *Tsix*-CTCF and the *Tsix*-CGI sites were hypomethylated, for the Xa and Xi (Figure 4.8). *Tsix* is the antisense transcript to *Xist* and counter the *Xist* expression ([Lee et al., 1999]). Therefore it'll be interesting to dissect whether this hypomethylation of *Tsix* CpG islands may influence *Xist* expression.

Loss of another important maternal effect gene DNA cytosine methyltransferase10 (*Dnmt10*) resulted in hypomethylation of the inactive X including at the XIC in blastocysts, specifically at the *Xite* promoter, along with placental defects post implantation [McGraw et al., 2013]. Comparatively, we did not observe significant hypomethylation at the *Xite* promoter a day earlier in gestation, in morula, in the absence of maternal *Smchd1*. Given that we already observe overexpression of *Xist* at this stage, it would appear that maternal *Smchd1* acts upstream of DNA



FIGURE 4.8: Maternal deletion of *Smchd1* does not significantly affect methylation at XIC in E14.5 placentae

a. Percentage methylation at promoters of XIC gene transcripts. X axis shows ensemble transcript IDs for XIC genes. Promoter-centric analysis was carried out, -3kb to +1 kb for each transcript. Ftx201-206 = *Ftx*, Jpx201-204 = *Jpx*, Rlim201 = *Rlim*, Tsix201 = *Xite*, Tsix202 = *Tsix*, Xist201-203 = *Xist* gene promoter CGIs. $n = 4$ for Maternal null (*matΔ*) **b.** RRBS annotation tracks displaying **a.** (*Xist* promoter CGI) **b.** (*Tsix*-CTCF and the *Tsix*-CGI) of X inactivation centre in wt and maternal *Smchd1* (*matΔ*) deleted E14.5 female placentae. **c.** Coverage plot of X inactivation centre in the same samples. $n = 3$ control (*wt*) for RRBS of E14.5 female placentae.

methylation in this context, however this requires further work considering the degree of variability in methylation and the small decrease in methylation at *Xite* we found, despite it not reaching statistical significance.

4.3.8 Maternal deletion of *Smchd1* results in a female-specific placental defect

Previous work showed that maternal loss of PRC2 components resulted in a variety of distinct phenotypic effects such as altered offspring weights and placental volumes [Prokopuk et al., 2018]. We were interested to understand whether deletion of maternal *Smchd1* had similar phenotypic effects. In Chapter 3 I showed that deletion of maternal *Smchd1* did not result in a change in offspring weight. To test whether the deletion of maternal *Smchd1* has an effect on placental size, particularly in the females given the iXCI defect, I carried out Micro-CT on embryonic day 10.5 (E10.5) placentae. I collected and prepared samples and Dr Lachlan Whitehead acquired and processed the data. I chose this technique as it allowed me to compare the measurements directly as opposed to other methods for measuring volume, for example histology where sections of the placentae are imaged and then reconstructed via computer models. Since we knew that we did not observe effects in offspring weight at weaning, several weeks after the time when maternal *Smchd1* works, I chose to analyse E10.5, a timepoint closer to the period when maternal *Smchd1* acts. E10.5 is a simple stage for dissection and also a stage when we know loss of zygotic *Smchd1* results in defects in the placenta.

I compared wt and mat Δ placentae from both MMTV-Cre and Zp3-Cre oocyte specific deletions. Placental volume was measured from these 3D images. I observed that placentae from male and female control embryos were not significantly different in volume in either Cre model. By contrast, when considering mat Δ placentae, there was a significant difference between male and female embryos in both deletion models (Figure 4.9 a and b), where the placentae from female were

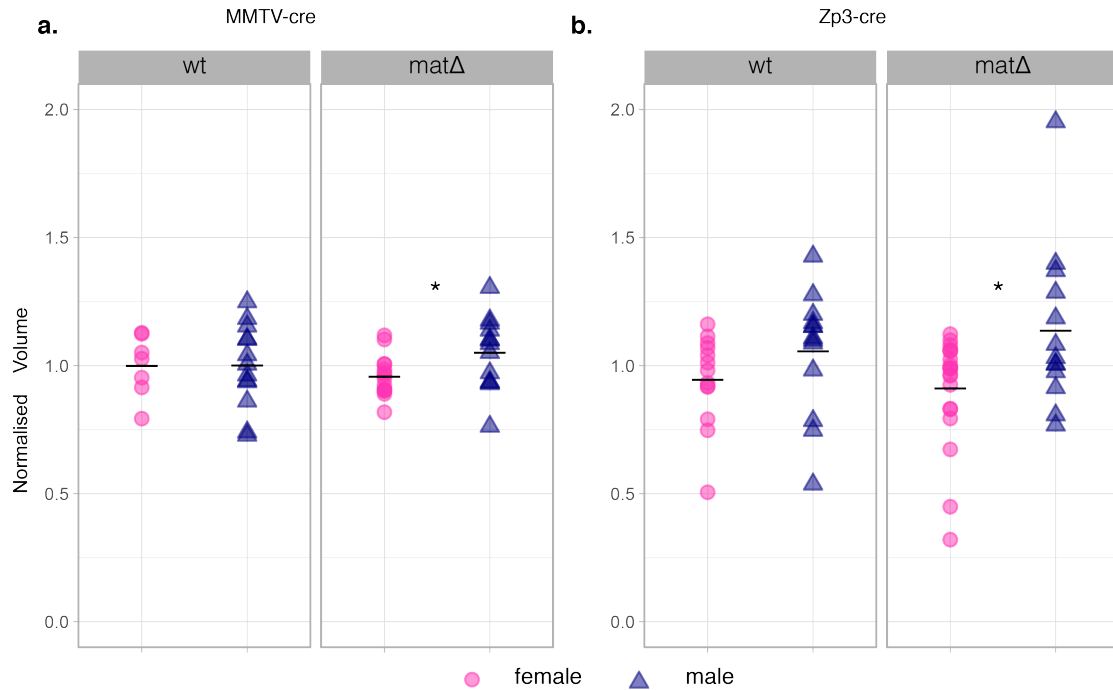


FIGURE 4.9: Maternal deletion of *Smchd1* results in smaller female placentae at E10.5

E10.5 placentae volume calculated by MicroCT in **a.** MMTV-cre and **b.** Zp3-cre wild type (wt) and maternal null (matΔ) placentae. Volume was normalised to the average volume of the litter. Statistical analysis was carried out by fitting a linear regression model to analyse the volume of wt or matΔ by litter. * $p < 0.05$

smaller and the placentae from male were larger. Given the iXCI defect that persists to E14.5, these data suggest there is a female-specific placental defect in the absence of maternal *Smchd1*.

4.3.9 Maternal deletion of *Smchd1* does not have a sex specific effect on the weight of the progeny

To understand whether the placental specific effect persists after birth and whether maternal deletion of *Smchd1* affects the weights of the progeny, we collected pups and weighed them at P2, as described in chapter 3 (Figure 3.4b). I carried out the analysis. While I didn't have enough litters to account for the litter size variance, I saw no difference in weight of the pups at P2 from maternal null versus control dams. Since we saw a small female-specific effect on placental size, I re-analysed

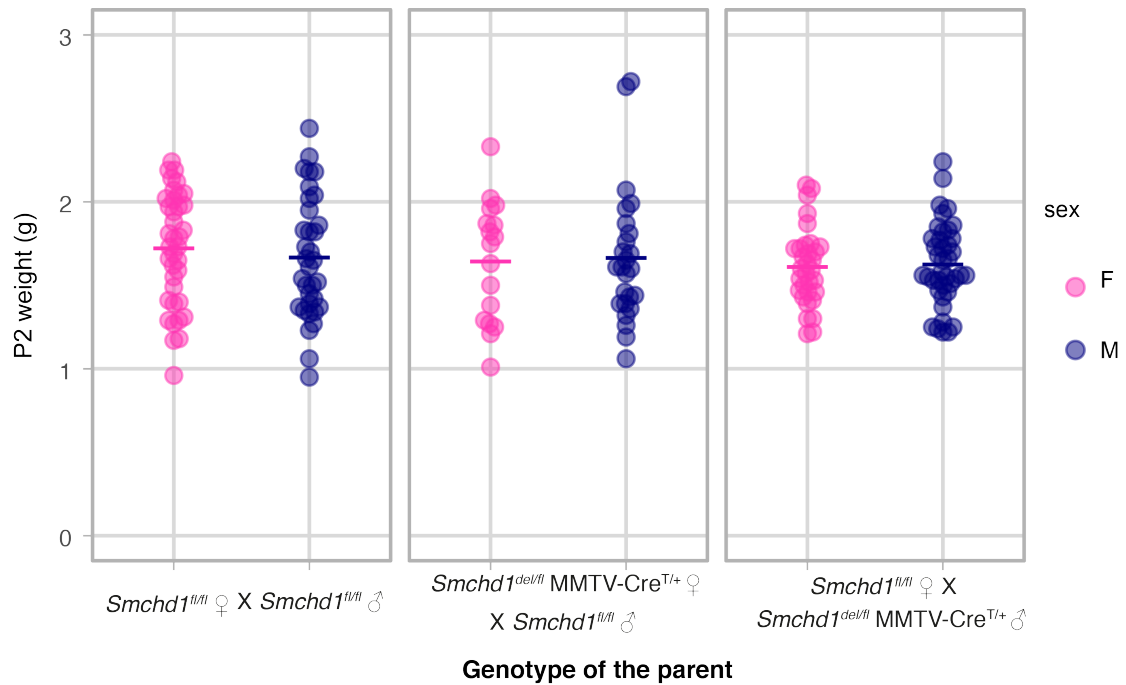


FIGURE 4.10: Maternal deletion of *Smchd1* does not have a sex specific effect on the progeny weights at P2

Embryo weights data at P2 female and male for progeny generated from *Smchd1*^{fl/fl} non-transgenic dams, or a reciprocal cross compared to controls. $n > 20$ pups for each cross. The p-value was calculated using student two-tailed T-test.

this data taking sex into account. I did not observe a sex specific effect on the weights at P2 (Figure 4.10), suggesting the placental size difference does not result in an effect on pup size postnatally.

4.4 Discussion

In mouse, extraembryonic tissues undergo iXCI where the paternal X chromosome is preferentially inactivated [Takagi and Sasaki, 1975]. iXCI is set up in early embryonic development as early as the four cell stage of development [Borensztein et al., 2017]. While the inner cell mass (ICM) undergoes the random form of XCI giving rise to the embryo proper, iXCI is maintained in the placenta, an organ that is crucial for embryonic development [Okamoto et al., 2011]. This parent of origin inactivation is catalysed by the long ncRNA *Xist* which contains an imprint on the Xm acquired during oogenesis [Tada et al., 2000]. Non-canonical imprinted expression control via the H3K27me3 polycomb mark regulates this mechanism by repressing *Xist* expression in cis on the Xm [Inoue et al., 2017a]. Extending this mechanism here, for the first time I found that *Smchd1* also modulates *Xist* expression during preimplantation embryo development. Expression of maternal *Xist* was accompanied by downregulation of the maternal X in *Smchd1* maternal null preimplantation embryos compared with controls. In addition, I found that similar to zygotic *Smchd1*, loss of maternal *Smchd1* also results in failed iXCI in both preimplantation and later stages of placental development.

I used allele specific genomics and imaging to tease apart the effect of maternal *Smchd1* in regulating *Xist* expression during preimplantation embryo development in both males and females. Allele specific RNA-FISH showed that in the absence of maternal *Smchd1* female embryos have biallelic *Xist* and males show *Xist* expression from their single (maternal) X chromosome. These effects were observed in all embryos analysed although not every cell in the embryo showed this phenomenon. This irregularity could be due to protocol issues with probe penetration, hybridization time, or the probe concentration. It could also be due to the fact that zygotic *Smchd1* is just beginning to be activated at the 16 cell stage, potentially resulting in restoration of appropriate *Xist* silencing on the Xm. It will therefore be interesting to carry out allele specific RNA-FISH in earlier embryos to check if the effect is more extreme earlier in development.

In our E14.5 placental data, we no longer observed expression of maternal *Xist*. A similar experiment was carried out by Harris *et al* where biallelic *Xist* expression in females and *Xist* expression in males is observed in the absence of maternal *Eed* in preimplantation embryos [Harris *et al.*, 2019]. Inoue *et al* observed that genomic imprinting of *Xist* is established during preimplantation development via a H3K27me3 mediated mechanism. Together the data from Inoue *et al* and Harris *et al* shows that in the absence of the H3K27me3 imprint on *Xist*, there is monoallelic *Xist* expression in the placenta [Inoue *et al.*, 2017a, Harris *et al.*, 2019], but this is due to random XCI in extra embryonic tissue rather than iXCI. This appears to occur due to some resolution of the biallelic expression of *Xist* early post-implantation. In our data we observe iXCI in the placenta in the absence of maternal *Smchd1* and we do not observe random XCI like these previous data. However, a similar resolution of biallelic expression of *Xist* may occur. We cannot discriminate whether the iXCI eventuates because those cells with expression of maternal *Xist* die and are lost, as may be expected given the large number of essential genes on the X chromosome, or whether zygotic *Smchd1* can restore appropriate *Xist* expression. Therefore, it would be interesting to obtain allele specific FISH data during gastrulation and placental development, to see how long Xm *Xist* expression persists. Overall, our data are consistent with maternal *Smchd1* being required to repress maternal *Xist*.

The second effect we saw in preimplantation embryos was failed inactivation of the paternal X. This effect persists post implantation where we see failure of iXCI in the absence of maternal *Smchd1*. Therefore, the effect exerted by maternal *Smchd1* has not been fully recovered by zygotic *Smchd1* or other factors by this stage. However this effect is subtle and has no developmental consequences apart from the slight differences in placental sizes. Dramatic effects on the progeny was previously observed in the event of two active X chromosomes resulting from *Xist* knockout [Kalantry *et al.*, 2009]. We observe hypomethylation of most CpG islands on the Xi, consistent with failed gene silencing. Interestingly though, absence of DNA methyltransferase 1 (*Dnmt1*) does not cause X-linked gene upregulation in the placenta, at least not in terms of a reporter allele [McGraw *et al.*, 2013]. Based

on these data and because maternal *Smchd1* exerts its effect preimplantation, we suggest that maternal *Smchd1* is likely to help establish iXCI. If this does not occur properly then iXCI has a persistent problem that cannot be recovered, and DNA hypomethylation simply represents failed silencing of the Xi.

We observed a phenotypic effect on E14.5 placentae, where female maternal null placentae appeared smaller than their male counterparts. It would be interesting to compare wild type versus *Smchd1* maternally deleted female and male placentae, however low number samples and the litter variance did not allow for this possibility to be explored at this point of time. It is of note that unlike the PRC2 component *Eed* maternal knockout mouse models we did not observe any embryonic lethality or effects on the weight of our maternal *Smchd1* null progeny [Inoue et al., 2017a, Prokopuk et al., 2018]. This could be because the effects that we observe are both more limited in scope than those seen for other factors and also more limited in effect size. This might be because other players in the non-canonical imprinted regulation pathway rescue the effect exerted by maternal *Smchd1*. It may also be because zygotic *Smchd1* is able to rescue the effect of loss of maternal *Smchd1* to some extent, for example in restoration of monoallelic *Xist* expression exclusively from the paternal X chromosome in iXCI.

Data on rXCI that involves zygotic *Smchd1* clearly show that *Smchd1* is recruited to the Xi through a PRC1- catalysed H2AK119ub pathway [Jansz et al., 2018b, Wang et al., 2019]. H2AK119ub also results in acquisition of H3K27me3 on the Xi [Pintacuda et al., 2017]. Although it is evident that the PRC2 mark H3K27me3 plays a significant role in non-canonical imprinting, the role of PRC1 catalysed H2AK119ub in non-canonical imprint regulation has not been formally tested and reported. It is likely that H2AK119ub also recruits *Smchd1* to non-canonical imprinted regions such as maternal *Xist*, similar to recruitment of *Smchd1* in rXCI.

As future directions, in addition to teasing out when maternal *Xist* expression is lost, as described above, it'll be interesting to analyse enrichment of polycomb marks H3K27me3 and H2AK119ub using ChIP-seq or CUT&RUN in relevant

tissues in the absence of maternal *Smchd1*. This would be most informative to test in preimplantation embryos while the non-canonical imprints are retained at all loci, and where we still see expression of maternal *Xist*. Previous studies also show that *Smchd1* modulates the higher order chromatin structure of the Xi following rXCI [[Jansz et al., 2018b](#), [Gdula et al., 2019](#), [Wang et al., 2018](#)], so it'll be also interesting to see whether maternal *Smchd1* operates via a similar mechanism to establish and maintain the non-canonical *Xist* imprint on Xm, which we could test by HiC or CaptureC methods.

Chapter 5

Macro-molecular structure and dynamics of Smchd1

5.1 Abstract

Thus far little is known about the structure and function of native Smchd1 in the nucleus. Here I used various imaging approaches to look into the molecular structure of native Smchd1 in the nucleus and how it behaves dynamically, in live cells. First, I performed fluorescence recovery after photo bleaching (FRAP) to understand the dynamics of Smchd1 binding to the inactive X chromosome. FRAP data shows that Smchd1 has a very slow turnover on the inactive X. Comparison of the dynamics between wild type Smchd1 and our unique gain of function mutant showed faster dynamics which may explain the gain of function observed in the mutant, and suggests the turnover rate is important for Smchd1 function. Using lattice light-sheet microscopy for the first time I show that Smchd1 unloads from the inactive X chromosome during mitosis and reloads once cell division is complete. Finally to understand the structure of Smchd1, I tagged both ends of Smchd1 and performed 3D-direct stochastic optical reconstruction microscopy (3D-dSTORM). I found that Smchd1 is 40 nm in length. Together these approaches of studying

both wild-type and mutant forms of native Smchd1 will further our understanding of the molecular mechanisms by which Smchd1 contributes to normal gene silencing and may shed a light on how things go wrong in disease.

5.2 Introduction

Up to now, studies on [Smchd1's structure](#) and function have mostly focused on characterizing the function and activity of its individual protein domains, in the context of recombinant protein. Smchd1 is a 2007 amino acid long protein containing an N-terminal catalytic GHKL ATPase domain and a C-terminal SMC hinge domain [[Chen et al., 2016b,c, 2020](#), [Pedersen et al., 2019](#)]. The middle region forming the junction between the N-terminal and C-terminal domains remains uncharacterised. It also remains to be determined how exactly the protein behaves in its full form, especially in a nuclear context. Therefore, I set out to use imaging to investigate Smchd1's structure and behaviour in its native environment, the nucleus, using several advanced microscopy methods. In particular, I was interested to ask three key questions: a) what are the dynamics of Smchd1 during interphase; b) what are the dynamics of Smchd1 during mitosis and c) what is the structure of Smchd1 in the nucleus. There was no prior knowledge in any of these areas at the start of my PhD. As such, I set out to establish a series of robust microscopy experiments that would allow us to look into the structure and function of the protein and set up a foundation to look at the impact and consequences the mutant forms have on Smchd1 dynamics and macro molecular structure.

5.2.1 Smchd1 gain of function mutant - MommeD43

For my studies, I was fortunate to have access to an unpublished mutant of Smchd1, called MommeD43. Smchd1 was first discovered using an [ENU mutagenesis screen](#) [[Blewitt et al., 2005](#)]. The original mutation was a null mutation called,

MommeD1, standing for modifier or murine metastable epialleles dominant mutation 1. Interestingly, the MommeD43 mutation, discovered during the same screen, is also a mutation in *Smchd1*; however, unpublished data from the Blewitt lab suggests MommeD43 is a gain of function mutant. MommeD43 has gain of function properties in gene expression that result in gain of function phenotypes in *Hox* gene silencing, where *Smchd1* is known to play a role [Jansz et al., 2018a]. This work on MommeD43 is currently in preparation for publication. How the MommeD43 produces a gain of function effect is not clear. This mutant afforded an additional opportunity to understand *Smchd1* function.

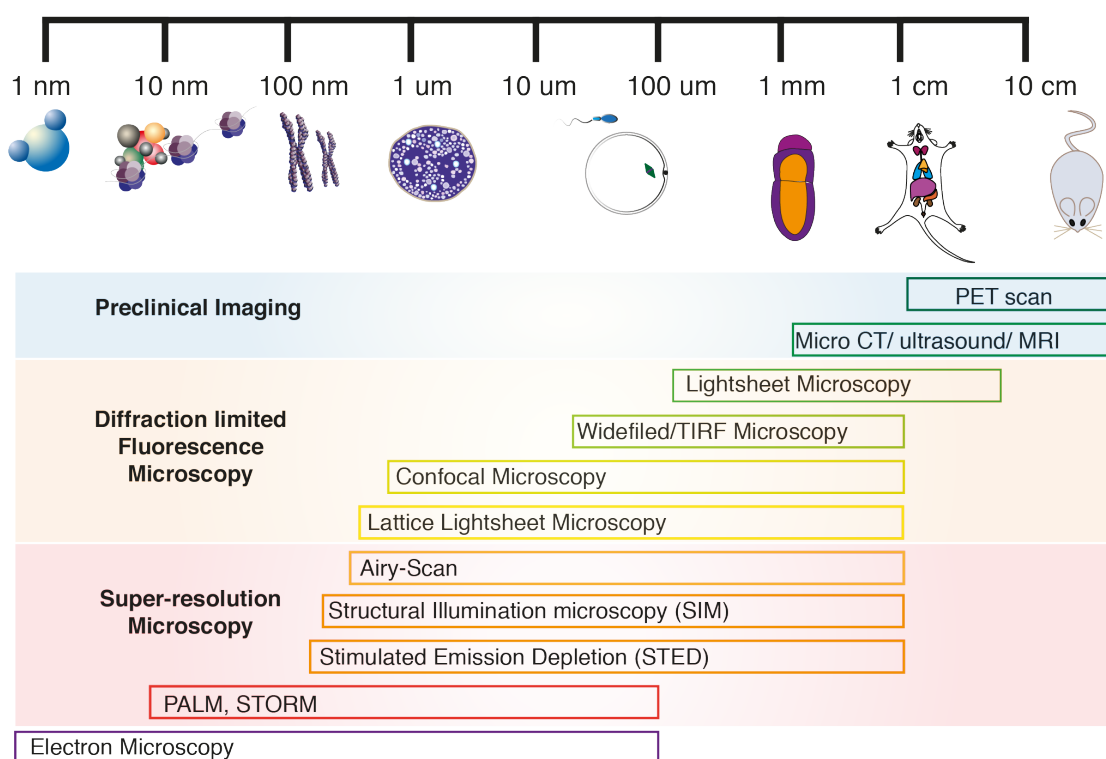


FIGURE 5.1: The observation of biology across a broad range of scales using different microscopy techniques.

Different microscopy techniques depending on size of the object that needs to be differentiated in the experiment. Large organisms (mouse), organs, mouse epiblast stage embryo, germ cells, mammalian cell nucleus, chromosomes, chromatin, and protein molecules displayed right to left in the diagram to correspond to scale.

5.2.2 Microscopy methods

Light microscopy techniques have evolved considerably over the last decade and are no longer just regarded as a tool for observation. Instead they have evolved to enable the sub-cellular measurement of molecules, their structural elements and how they function inside a more complex system, whether that be at the level of single cells or whole organisms. Microscopy as an applied science can broadly address biological questions and is an essential tool in biomedical research, which goes beyond the powerful but simplistic notion that ‘seeing is believing’. There are a number of different microscopy methods that can be used to address questions in biology across a wide range of scales, from the single molecule right up to whole organisms (Figure 5.1)

One of the big developments in the last two decades has been in super-resolution microscopy. In the year 2014 the Nobel prize in chemistry acknowledged the development of these techniques for overcoming the diffraction limit. Indeed, light microscopy is limited by the diffraction limit of light (d), which was calculated by Abbe over 100 years ago. The equation accounts for the numerical aperture (NA) of a microscope’s optical lens and wavelength of light (λ) [Abbe, 1873].

$$d = \frac{\lambda}{2NA}$$

According to Abbe’s diffraction limit a resolution of approximately 250 nm can be achieved by conventional microscopy methods (approximately half the wavelength of light). This is not enough to resolve structures that are smaller than 250 nm, certainly not the structure of a protein or indeed protein dynamics. For over a century Electron Microscopy (EM) methods were the only approaches that could be applied to overcome this diffraction limit (reviewed in Allen [2008]). But EM sample preparation limits usage of fluorophores, and makes it impossible to observe living cells [Allen, 2008]. EM does allow ultrastructural resolution below 1 nm in the best cases; however EM sample preparation is tedious, highly specialized and leads to artefacts. In contrast, living cells can be observed in their natural environment and their features characterized by tagging different proteins or features with

fluorescent proteins. These can then be observed using light microscopy. Emerging techniques based on expansion microscopy promise to overcome Abbe's diffraction limit even more by revealing macromolecular structures with a resolution that is more compatible to EM (reviewed in [Won \[2009\]](#), [Schermelleh et al. \[2010\]](#)).

Abbe's limit can be overcome by two different approaches. Direct approaches, which directly observe the target molecule, involve super resolution microscopy, such as structural illumination microscopy (SIM), Stimulated emission depletion (STED), Photo-activated localization microscopy (PALM), direct stochastic optical reconstruction microscopy (dSTORM) and looking at single molecules [[Schermelleh et al., 2010](#)]. In contrast indirect approaches such as Fluorescence recovery after photobleaching (FRAP), Fluorescence correlation microscopy (FCS), and Fluorescence lifetime imaging (FLIM) or Förster resonance energy transfer (FRET) enable studies on the behaviour and biophysical properties of proteins by looking at the volume of fluorescent molecules and their behaviour (reviewed in [Ishikawa-Ankerhold et al. \[2012\]](#)).

Despite advances in modern optical microscopy techniques there remains a fine balance between spatial resolution (contrast), sample health (photo-toxicity, bleaching), and temporal resolution (speed) among other things [[Schermelleh et al., 2010](#)]. Currently no microscopy technique is able to overcome all these trade-offs. Therefore, it is vital to choose the appropriate techniques to tackle the relevant questions accounting for the limitations in each approach (Figure 5.2). Keeping this in mind I have chosen three advanced microscopy techniques to begin to investigate Smchd1's structure and dynamics in the nucleus.

5.2.2.1 Functional analysis by Fluorescence Recovery after Photobleaching (FRAP)

To observe Smchd1 dynamics in the nucleus I made use of a technique known as Fluorescence Recovery After Photobleaching (FRAP). When carrying out FRAP experiments, fluorescence molecules in a small region of interest (ROI) are irreversibly photobleached using a high intensity laser beam for a short period of time

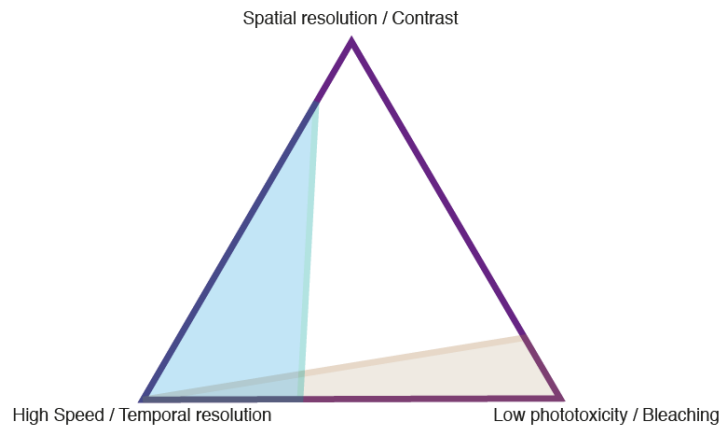


FIGURE 5.2: **Triangle of “Frustration”**

Trade-offs in advanced fluorescence microscopy techniques. In an ideal situation the triangle will be a perfectly symmetrical **purple triangle**, but currently trade-offs are made in order to achieve the desired results. For live cell experiments either the resolution or the speed is sacrificed **brown triangle**, while for the experiments that require higher resolution sample health or the speed is sacrificed **blue triangle** and so on.

and fluorescence recovery within the bleached region is observed immediately after and over time using a low intensity laser [Axelrod et al., 1976] (Figure 5.3a). The laser bleaching inactivates the fluorophore, so the recovery of fluorescence indicates the culmination of the unloading of the bleached protein and reloading of fluorescent protein from elsewhere in the cell. Therefore, FRAP is used to analyse dynamic turnover of the protein, usually during interphase.

In FRAP experiments due to the high intensity of the laser used to bleach the region, and the time duration that the live cells are under duress, it is important to control for photobleaching and other artefacts (Figure 5.4b). FRAP is typically performed using a confocal microscope with a FRAP module. Post data processing enables specific details about protein dynamics to be obtained, for example: whether a particular protein molecule is mobile within a particular cellular volume; whether it is restricted in movement or immobilized potentially due to tethering to another structure or protein; and how long the protein can stay bound to a particular region in living cells (and remain immobile) [van Royen et al., 2009]. Analysis of FRAP also makes use of fitting various mathematical models depending on the desired experimental outcome [Kang et al., 2015] (Figure 5.3c).

Although FRAP was initially developed to look at dynamics of protein mobility through a membrane, it has been proven to be an excellent and very well established methods to get information about chromatin bound proteins. Polycomb group proteins and HP1 [Almeida et al., 2017, Strom et al., 2017] are examples of chromatin proteins studied using FRAP, where they showed fast recovery times of less than 30 seconds. We decided to use this method to observe Smchd1 dynamics in live cells. I prepared the samples and carried out the experiments. Technical expertise was provided by A/Prof. Kelly Rodgers and Dr Niall Geoghegan. Dr Tracy Wilson produced the NLS-acGFP construct. Dr Quentin Gouil and I analysed the data.

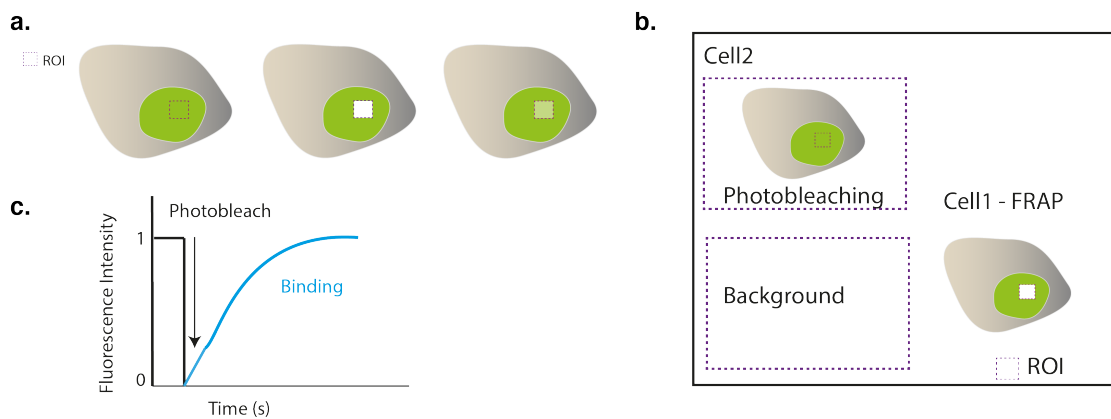


FIGURE 5.3: **Principals of Fluorescence Recovery After Photobleaching (FRAP)**

a. Schematic of a FRAP experimental set up. ROI = region of interest. **b.** Controls for FRAP experiments, by imaging a second nearby cell that is not undergoing bleaching, and observing background fluorescence changes. **c.** A typical FRAP curve.

5.2.2.2 Fast volumetric imaging using lattice light sheet microscopy

A major limitation of conventional light microscopy is phototoxicity, which specifically adversely affects observations of sensitive cellular processes such as cell division. Invention of Lattice LightSheet microscopy overcame this limitation and allowed to look at development of organisms without damaging the live cellular processes [Chen et al., 2014]. Unlike the conventional microscopy methods Light-Sheet microscopy uses a thin sheet of light, which is positioned orthogonally to the detection plane of the camera. As such, only the area captured by the camera

is also excited by the laser. The light sheet provides the means to perform optical sectioning in a 3D volume, which contrasts to confocal where light is captured point by point by way of a pinhole, but still penetrates the sample in areas that are outside of the focal plane. This approach is overall less toxic to live cells and capable of capturing rapid volumes, giving excellent images in 3D with good axial resolution, while keeping phototoxicity to a minimum (Figure 5.4) [Chen et al., 2014]. This also makes it an ideal system to study protein dynamics during mitosis in live cells. Therefore, I chose to study Smchd1's dynamics during cell division using Lattice LightSheet microscopy. Technical expertise was provided by Dr Niall Geoghegan. Dr Niall Geoghegan acquired the data and also reconstructed images. Dr Tracy Willson produced the H2B-mScarlet1 construct. I prepared the samples analysed and interpreted the data.

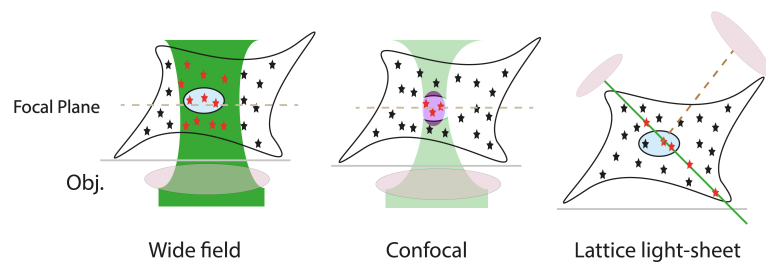


FIGURE 5.4: **Lattice lightsheet microscope set up compared to conventional microscopy**

Comparison of widefield, confocal and lattice lightsheet microscopy methods, Light path in green, illuminated region of the cell in red, objectives in light pink

5.2.2.3 Structural analysis using single molecule imaging methods

To investigate the structure of Smchd1 in the nucleus, going beyond what has already been established by biochemical methods on recombinant protein, I made use of single molecule imaging using a super resolution microscopy technique known as direct stochastic optical reconstruction microscopy (dSTORM). This technique makes use of photoswitchable fluorophores, providing the ability to switch on and off fluorophores at different time points using buffers containing reducing agents, in order to have them blinking (“on-off”) in a stochastic way. Then, localization

microscopy is performed by taking a series of images which can be reconstructed into a final image with a resolution close to 10-20 nm under ideal conditions (Figure 5.5) [Hess et al., 2006, Betzig et al., 2006, Rust et al., 2006, Heilemann et al., 2008]. As Smchd1 is a large protein of about 230 kDa we hypothesized it might be possible to use this approach to obtain some information regarding its macromolecular structure. In this microscopy technique, the buffer conditions and choice of fluorophores are vital to achieve optimal results [Whelan and Bell, 2015, Jimenez et al., 2020]. These experiments were performed in collaboration with Dr Kirstin Elgass (Hudson Institute, Monash University) and Dr Michael Mlodzianoski (Centre for Dynamic Imaging, WEHI). I prepared all the samples and assisted in the optimization of experiments. Dr Michael Mlodzianoski carried out all dSTORM image acquisition and analysis. I contributed to the final interpretation of data.

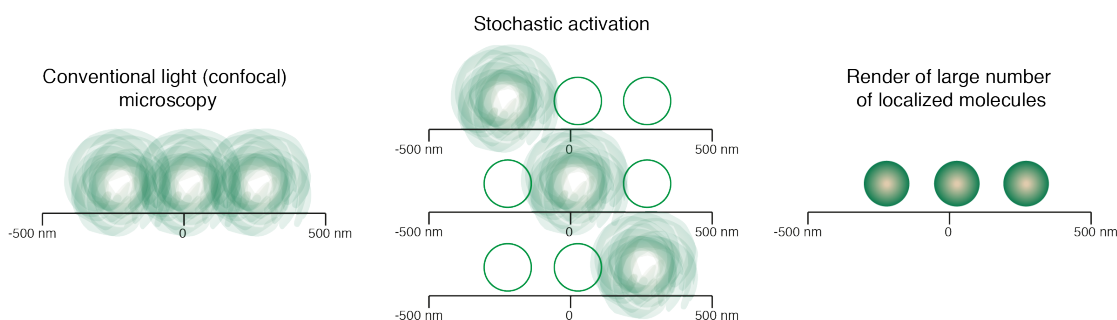


FIGURE 5.5: **Schematic of basic principle of single molecule localisation microscopy**

Individual molecules closer than Abbe's limit cannot be separated by conventional microscopy methods. Single molecule localization microscopy makes use of photoswitchable fluorophore "blinking" to overcome this resolution limitation by taking a series of images overtime and reconstructing to resolve the final molecular structure

5.3 Results

5.3.1 Smchd1 shows slow dynamics during interphase

Smchd1 is a nuclear protein that is dispersed within the nucleus but also forms foci at the Xi and other chromatin bound regions. XCI lasts throughout the lifetime of the female mammal and as Smchd1 is a major player keeping the Xi silent, one could assume that it is rather immobile. However, Smchd1 dynamics weren't previously studied. So I carried out FRAP (Figure 5.4) to understand Smchd1 dynamics during interphase of cell division. We used iMEFs derived from Smchd1-GFP mouse (Figure 5.11b) as they have higher GFP signal which is favourable for live cell imaging. I observed that wild type (wt) Smchd1 turnover was rather slow on the Xi, even after 40 minute of observation Smchd1 was not fully recovered (Figure 5.6)

The slow fluorescence recovery suggested that the fluorescence recovery corresponded to Smchd1-GFP binding to free immobile sites, and that the unbinding rate was much slower than the diffusion rate. We therefore applied a diffusion-uncoupled model of FRAP, fitting a simple exponential recovery, $frap(t) = 1 - Aexp(-K^{off}.t)$, where K^{off} is the unbinding constant (Figure 5.7) [Sprague and McNally, 2005, Kang et al., 2015].

When this model is fitted we observed that the FRAP curve had not yet plateaued by 40 minutes (Figure 5.8a); however, 40 minutes was the longest time period we were able to follow the same nuclei due to phototoxicity, photobleaching and the time constraints of a very long acquisition period for each single nucleus. To test whether the observed recovery after 40 minutes was due to maturation of new Smchd1-GFP, we compared the data to AcGFP with a nuclear localisation signal (NLS), that was not tethered to any other protein. AcGFP is the specific form of GFP that was used to make the Smchd1-GFP mouse model. We found that when the whole nucleus was bleached, AcGFP did not recover at all in 40 minutes (Figure 5.8b). These data suggest that the recovery seen for Smchd1-GFP is due to reloading of Smchd1-GFP rather than maturation of new protein.

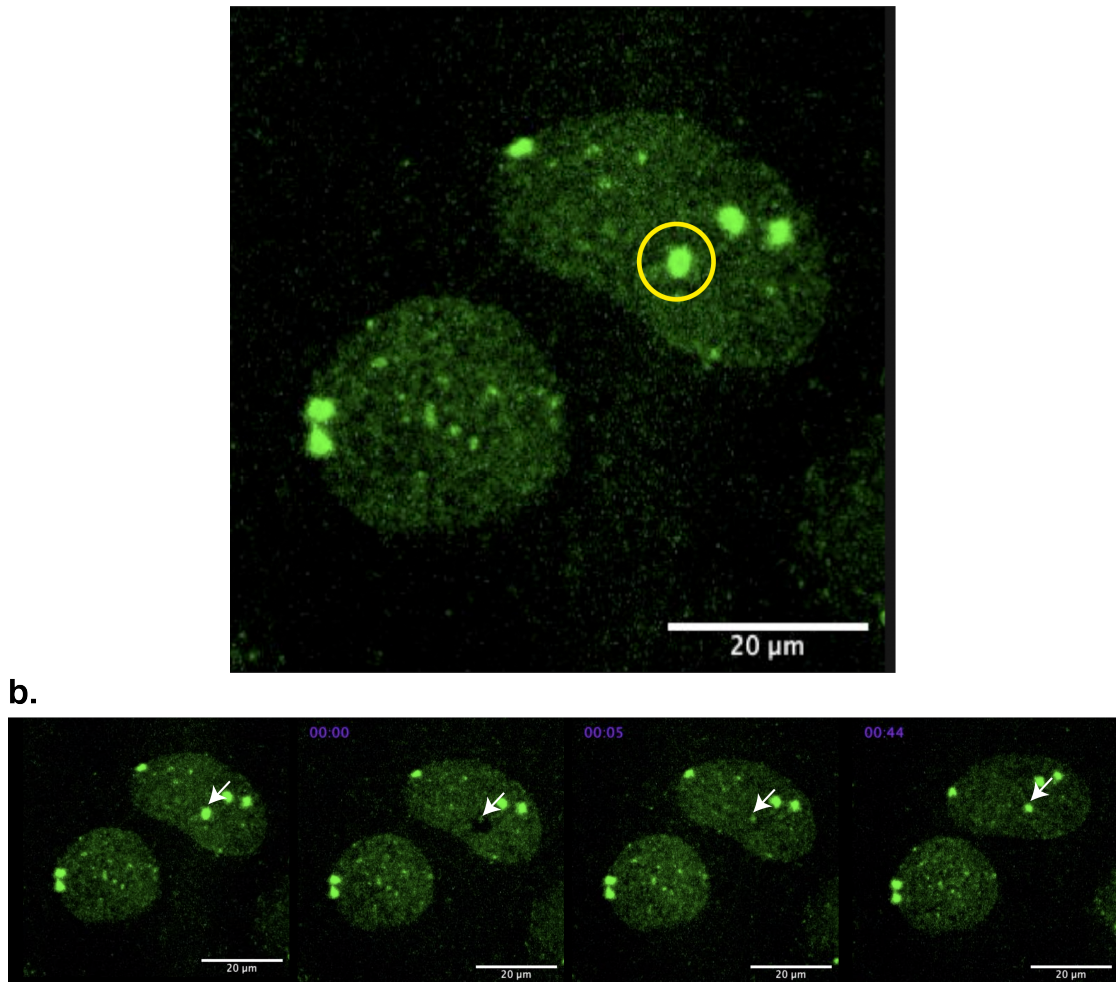


FIGURE 5.6: **FRAP** movie in **Smchd1-GFP** iMEF

a. Example FRAP movie in wt iMEF. Time in minutes. Bright foci are the Xi. Immortalised MEFs are frequently polyploid, so there is more than one Xi. Two cells observed - one where one Xi was bleached and the other that was used for photobleaching normalization. Movie can be viewed using adobe acrobat reader or by following the [link](#) **b.** Stills from the movie. Arrows indicate the Xi that was subjected to bleaching.

Greater than 40 minutes recovery time is quite significant compared to other known heterochromatin marks such as HP1 and polycomb group proteins that recover within 30 sec [Almeida et al., 2017, Strom et al., 2017].

I next compared the wild type data to the gain of function Smchd1 mutant MommeD43. Interestingly, the Smchd1-GFP-MommeD43 mutant showed faster dynamics compared to wild-type Smchd1-GFP (Figure 5.9a). These data suggested that potentially MommeD43 enhances Smchd1 silencing by increasing the rate of turnover of Smchd1 on chromatin.

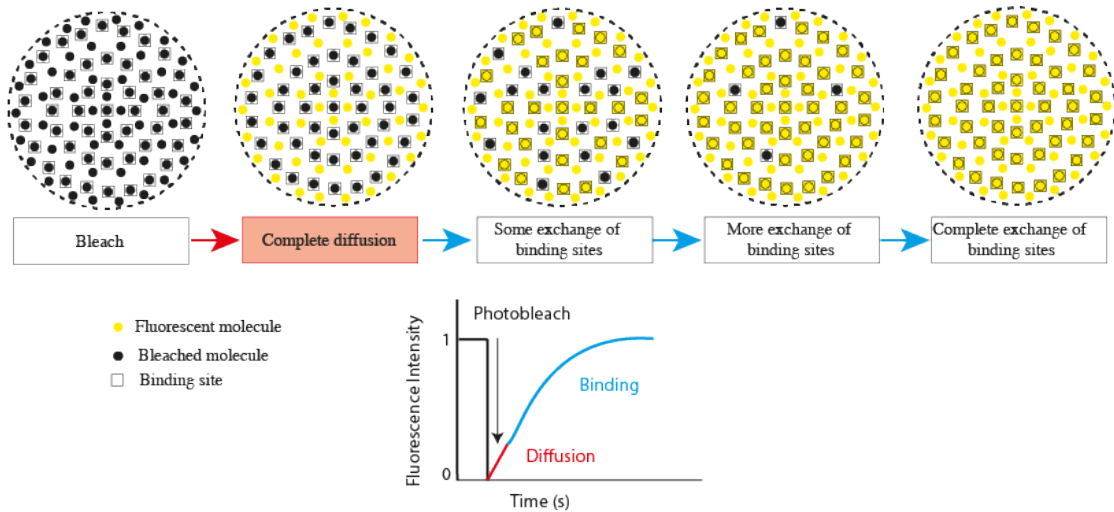


FIGURE 5.7: Diffusion-uncoupled model used to analyse FRAP data

Pictorial representation of diffusion-uncoupled model used to analyse the FRAP data. When the ROI is bleached with the high intensity laser first the nearby molecules diffuse into the ROI. After that binding and unbinding of fluorescent and bleached molecules occurs in a much slower rate.

The data from iMEFs was the first piece of data to suggest active turnover may boost Smchd1 function. I chose to validate these data in primary cell type, NSCs. These cells move rapidly, which meant they frequently moved out of frame when trying to image for 40 minute periods. Therefore, I decided to perform FRAP over a short time interval of 5 minutes. In order to have comparable data in iMEFs and NSCs, I first did the short time FRAP in iMEFs (Figure 5.9b) which again showed that the MommeD43 mutant had a significantly higher unbinding rate compared to wild-type Smchd1. In NSCs the same trend was seen although the values weren't significant (Figure 5.9c). This could be due to the number of cells scored and also due to the difficulty in carrying out this experiment in practice given the rapid movement of NSCs, even over 5 minutes.

Together the FRAP data in NSCs and iMEFs suggests that the MommeD43 mutation increases the turnover rate of Smchd1 on the Xi during interphase. Given that the MommeD43 mutation results in a gain of function effect on Smchd1-mediated silencing, this suggests that turnover is required for silencing.

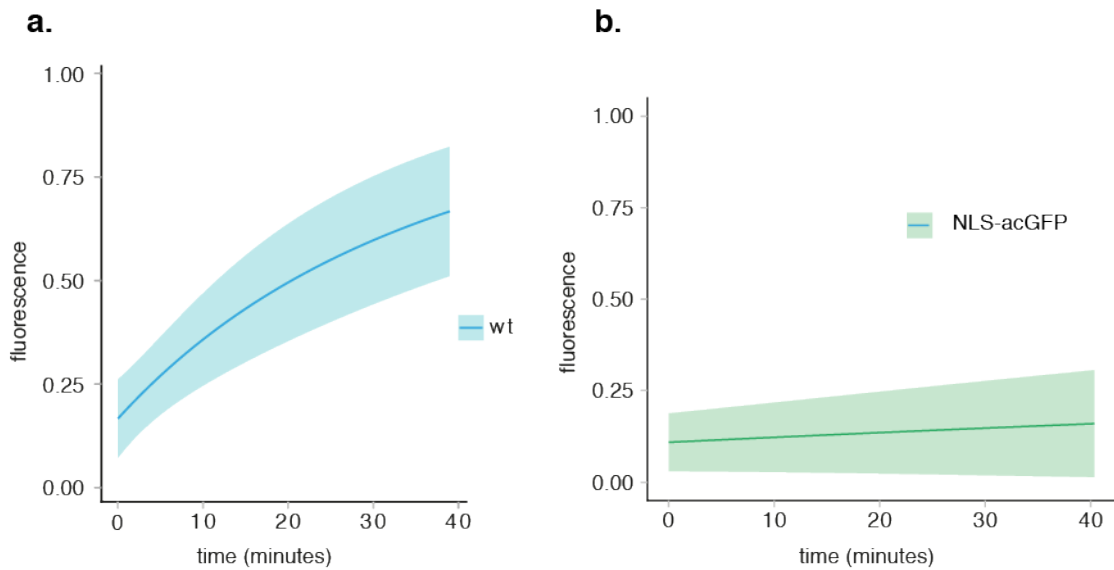


FIGURE 5.8: **Slow dynamics of Smchd1 cannot be attributed to GFP maturation**

a. Recovery curve obtained for Smchd1-GFP wt iMEFs. N=22, two different cell lines scored in more than 3 biological and technical replicates. **b.** Ac-GFP-NLS maturation curve in iMEFs, where the whole nucleus was bleached to analyse protein maturation. N=4. The SEM of each individual recovery curve is shown by the coloured areas over the average line.

5.3.2 Smchd1 does not remain bound to Xi during mitosis

While FRAP data revealed the Smchd1 dynamics during interphase we wanted to understand whether Smchd1 remains bound to the Xi chromosome while the cell is dividing. Other XCI factors such as Long ncRNA *Xist* and PRC2 components Ezh2 were previously known to be bound to the Xi during cell division [Lee et al., 1999, Aoto et al., 2008]. Quite interestingly, when I was imaging Smchd1-GFP cells I observed that in the cells that appeared to be undergoing division there was no Smchd1-GFP fluorescent signal. This could have been due to the mitotic cell being out of the appropriate focal plane. To validate the lack of Smchd1 in mitotic chromatin, I carried out mitotic shake off to enrich for cells that are undergoing cell division and carried out 3D Airy-scan imaging of wild-type and Smchd1-GFP female cells for GFP epifluorescence, and again observed a lack of Smchd1 on mitotic chromosomes (Figure 5.10). However, this kind of imaging only produces a snapshot in time, mostly in 2D. Cell division is quite a dynamic and a sensitive process that needs to be observed in 3D and ideally in living cells.

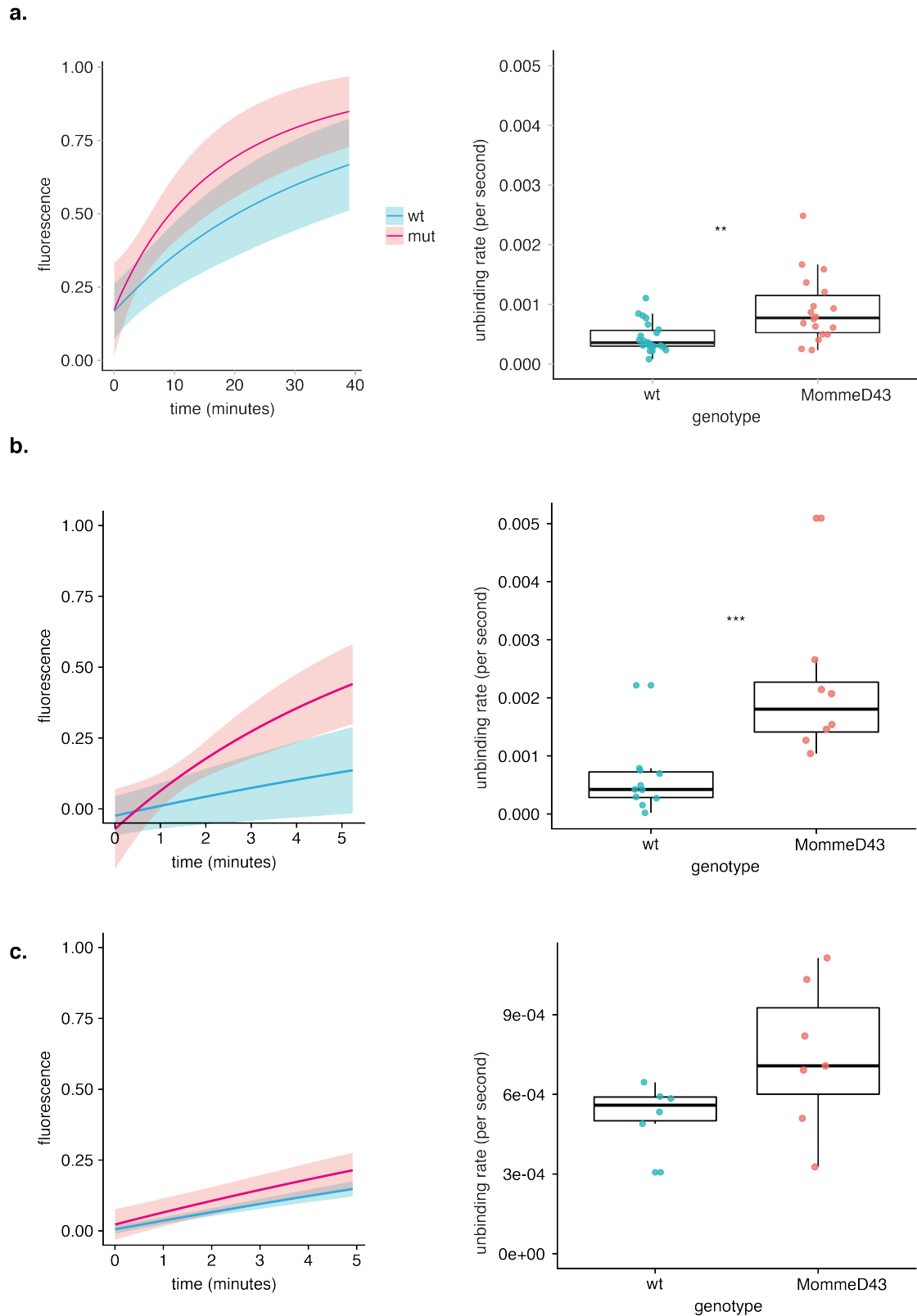


FIGURE 5.9: **Smchd1 – MommeD43 mutant shows faster dynamics**

Left panels: Average recovery curves (line) obtained for Smchd1-GFP mut (MommeD43) and Smchd1-GFP wild-type **a.** in iMEFs long time period (40 minutes) FRAP. wt N = 22, mut N=18. **b.** in iMEFs in short (5 minutes) time period. Wt N = 10, mut N=8. **c.** in NSCs in short (5 minutes) time period wt N = 6, mut N=7. The SEM of each individual recovery curve is shown by the coloured areas over the average line. Right panels: individual calculated unbinding rate for each nucleus, obtained from the individual recovery curves that were used to calculate the average recovery curves shown on the left. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

In order to observe Smchd1-GFP on the chromatin during mitosis using live imaging, we carried out Lattice lightsheet microscopy. Again, we used Smchd1-GFP immortalised MEFs (iMEFs), this time transduced with H2B-mScarlet to mark the chromatin. I observed that Smchd1 is removed when the cell enters into metaphase and it is recovered when the cell moves into anaphase/telophase of the cell cycle (Figure 5.11). For these data, the staging is based on difference observed in chromatin structure. While we didn't have the precise stages, since each image is acquired at 2 minute time intervals we could calculate the time that the cell is without Smchd1. We found that from our 2 datasets (2 cell divisions), Smchd1 is absent for 44 min (average of 42 minute and 46 minute).

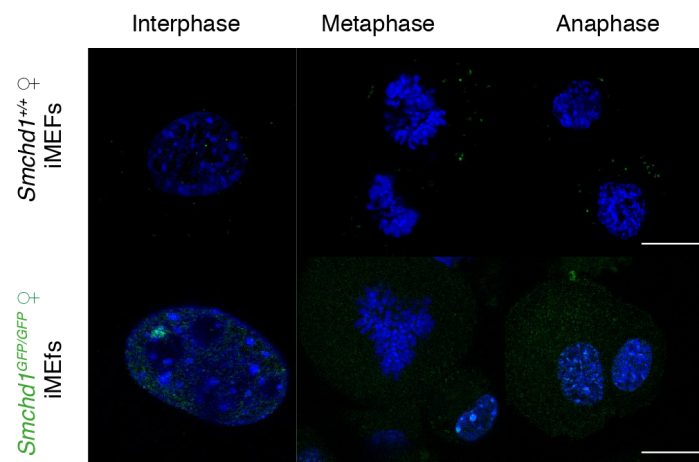


FIGURE 5.10: **Smchd1 doesn't stay bound to Xi during cell division**

Airyscan images obtained for **a.** Smchd1 wt and **b.** Smchd1-GFP wt female iMEFs followed a mitotic shakeoff to enrich for dividing cells. The native acGFP fluorescent signal was observed, with chromatin marked by DAPI. N=5

5.3.3 Macro-molecular structure of full length Smchd1 in the nucleus

Thus far, full length Smchd1 protein has proven to be rather difficult to purify in order to carry out protein structure studies [Gurzau, 2020]. The only available full length information is from negative stain electron microscopy images on recombinant Smchd1 which showed that the Smchd1 dimer is approximately 40 nm

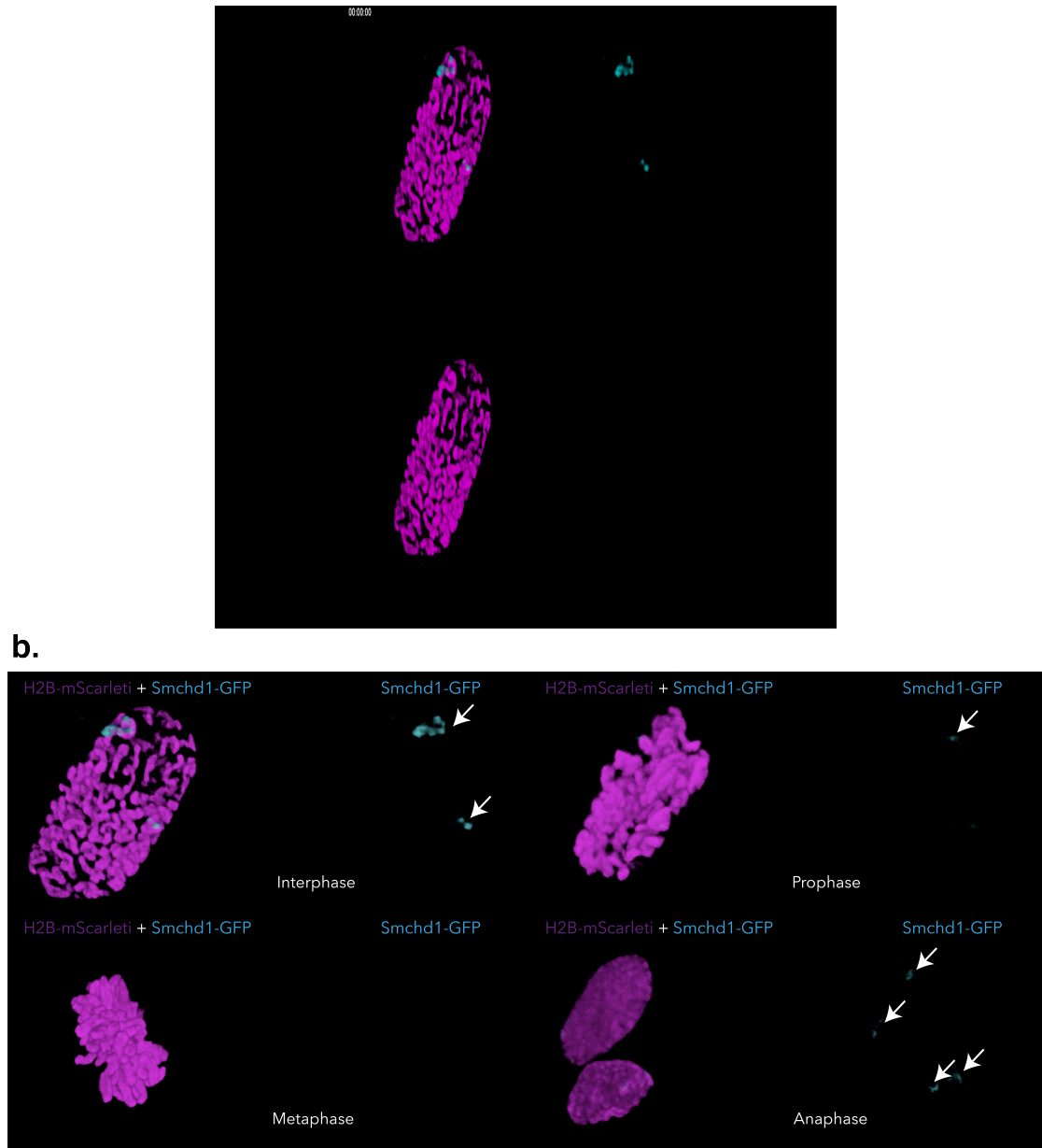


FIGURE 5.11: **Lattice lightsheet movie showing loss of Smchd1 from the Xi chromatin during mitosis**

Example lattice lightsheet movie of (Smchd1) bound to chromatin shown in (H2B-mscarleti) dynamics during cell division in Smchd1-GFP iMEF. Time in minuteutes. Bright foci are the Xi. Since these cells are immortalised there is more than one Xi. Movie can be viewed using adobe acrobat reader or by following [this link](#) **b**. Stills from the movie. Arrows indicate Smchd1 protein localisation

in length [Brideau et al., 2015]. To evaluate the full length of Smchd1 in its native environment, and compare whether it is similar to the recombinant protein, I used the single molecule localization microscopy method, 3D-dSTORM.

For these experiments we used immortalised mouse embryonic fibroblasts (iMEFs) and mouse neural stem cells (NSCs) derived from *Smchd1*-GFP mice. The full length *Smchd1* has a C-terminal acGFP1 tag, after a short two amino acid linker (Jansz et al. NSMB 2018). I decided to use NSCs and iMEFs as they express *Smchd1*-GFP significantly higher than other cell types. In addition, mouse embryonic fibroblasts are large, relatively flat cells that are ideal for STORM experiments (Figure 5.12b) [Jimenez et al., 2020]. I used an anti-GFP antibody to detect the C-terminus of *Smchd1* and an inhouse monoclonal antibody that binds to N-terminal GHKL-ATPase of *Smchd1* to detect *Smchd1*'s N terminus (Figure 5.12a).

One crucial step in STORM experiments is deciding the fluorophore combination used as the end resolution depends on the quality of the fluorophore and its ability to “blink” when stimulated. For this reason, I used Alexa 568 and Alexa 647 secondary antibodies and optimized both the fluorophore concentration and buffer conditions to achieve optimal results.

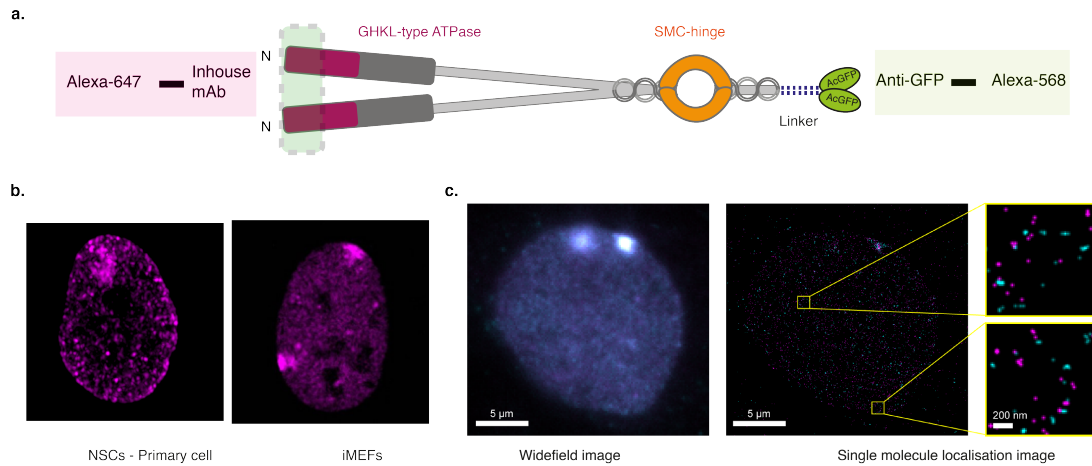


FIGURE 5.12: ***Smchd1* macro molecular structure using 3D-dSTORM microscopy - experimental set up**

a. Schematic of experimental set up. *Smchd1* with the GFP tag at the C-terminal hinge domain. The Anti-GFP primary antibody used in combination with Alexa 568 dye to detect the C-terminus, and in house mAb with Alexa 647 used to detect N-terminus **b.** Confocal images of *Smchd1*-GFP female NSCs and iMEFs. *Smchd1*-GFP is shown in magenta. The single bright focus in the NSC nucleus is the Xi, whereas iMEFs frequently have polyploid so have multiple Xi. **c.** Single molecule localization image. Scale bar corresponds to 5 µm.

5.3.3.1 3D-STORM sample analysis

All image acquisition and analysis was carried out by Dr Michael Mlodzianoski (Centre for Dynamic Imaging, WEHI). I will describe the analysis here with sample data before presenting our final data as these analyses are complex and custom to our experiments.

Due to the large amount of Smchd1 found on the Xi, we chose to analyse the nucleoplasm rather than the Xi, since single molecule imaging requires fluorophore separation meaning the Xi was too densely stained. For each image acquired, the nearest neighbour distance between Alexa 647 and Alexa 568 was calculated (Figure 5.13a, red curve). The curve is composed of a mixture of the Smchd1 probe pairs (both pairs properly identified) and the random distribution of probes without a counterpart at the opposite end of the Smchd1 protein. The lack of a counterpart can be due to several factors: some proteins will not be labelled at both ends, some fluorescent probes will photobleach with too few photons to be detected and/or signal to noise may be too low to reliably localize the molecules. The densities of Alexa 568 and Alexa 647 and the imaged volume were recorded for each image. As a control, random molecule positions were generated based on these values for both fluorophores. The nearest neighbour distributions were calculated identically the same as for the experimental data. The randomized control distributions (Figure 5.13a, blue curve) were compared to the experimental data using the Kolmogorov-Smirnov statistical test to determine if any statistical difference exists between the curves. The random nearest neighbour portion of the distribution from the experimental data can be removed by comparison with the control. A simplex algorithm was then applied to determine the optimal weight of the control to subtract from the experimental distribution. The remaining curve (Figure 5.13a, black curve) represents the nearest neighbour measurements from the fluorescent probe pairs at opposing ends of the SMCHD1 protein. This curve is fit to a Gaussian distribution (Figure 5.13a, green curve) to determine the mean separation between the two fluorophores for each imaged nucleus.

The plot of the mean separation of fluorophores as a function of the Kolmogorov-Smirnov test p-value (Figure 5.13b) reveals a wider range in separation values at larger p-values. The separation lengths calculated in Figure 5.7a only represent the mean distance between the imaged fluorophores. However, the length of the antibody and possible angle between the antibody and the Smchd1 protein should be considered to properly determine the length of Smchd1. More insight into the length of Smchd1 can be gained by performing simulations. (Figure 5.13c) shows a physical representation of the simulations we performed next. A rigid body of length p , representing a protein of fixed length and secondary rigid bodies of length a , representing antibodies, attached at each end of the protein are simulated. The antibodies are free to occupy any angle in 3D and were simulated to a maximum value φ . Molecules at the end of each antibody had localization precisions corresponding to the experimentally measured values. The simulation has the five following variables; p , protein length; a , antibody length; φ , maximum polar angle; lateral localization precision; and axial localization precision. The lateral and axial localization precisions were taken from the measurements of each experiment and had values of 12.5 nm and 39.1 nm, respectively. The antibody length was assumed to be 10 nm. The two remaining variables were tested over a range of conditions. 100,000 iterations were performed for each condition. The nearest neighbour distances were calculated the same as for the experimental data.

5.3.3.2 Smchd1 is approximately 40 nm in length in the nucleus

Each condition tested a different simulated protein length from 30 to 60 nm and a different maximum polar angle from 0 to π . Figure 5.14 shows the mean probe separation for these simulations, using a heat scale to indicate how many of the simulations gave each length for each maximum polar angle. The cyan lines correspond to the experimentally measured mean separation values of 83 nm (Figure 5.14a) and 74.9 nm (Figure 5.14b). The data for NSCs were similar but are not shown. Given the antibody lengths, based on these data the Smchd1 length is calculated to be 41.8 nm for iMEFs and 43.1 nm for NSCs (Table 5.1).

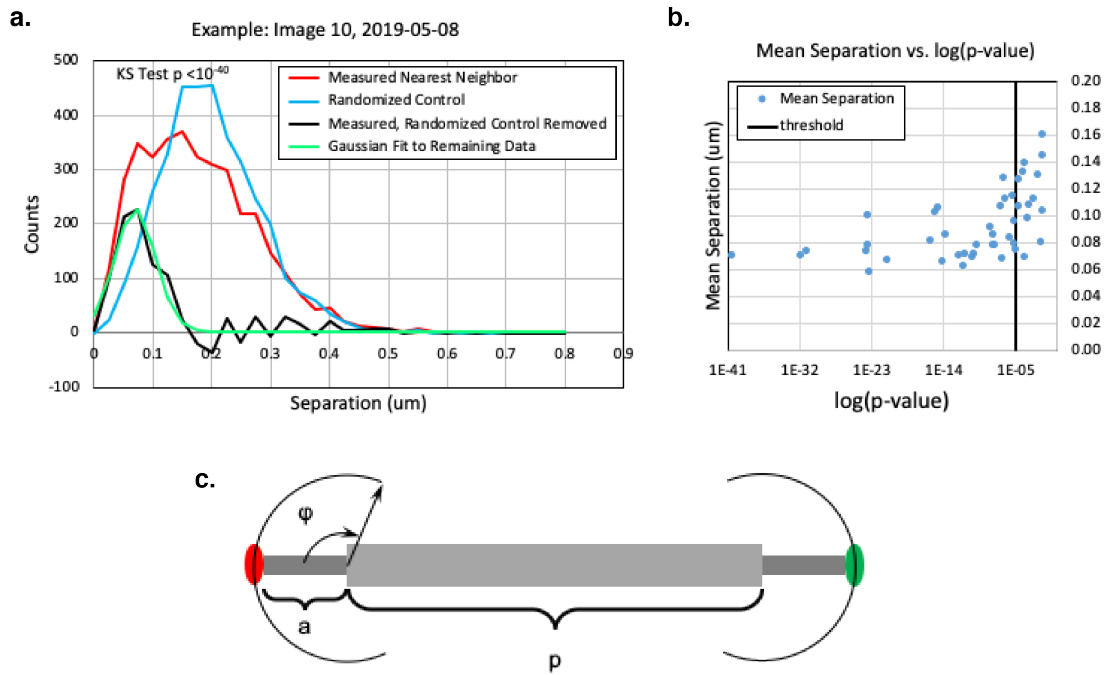


FIGURE 5.13: 3D-dSTORM microscopy - Image analysis

a. Nearest neighbour distributions from a single nucleus. Each imaged Alexa 647 molecule was matched to its nearest Alexa 568 molecule within the nucleus and the separation distance plotted as a histogram (red curve). The same number of molecules for each colour probe were randomly distributed within the same size region as the experimental data. The corresponding separation was calculated (blue curve). The Kolmogorov-Smirnov test was performed on these curves ($p < 10^{-40}$) in the above example. The random distribution was removed from the experimentally measured distribution and was done so using a simplex algorithm to determine the optimal weight of the randomized distribution to remove (black curve). The remaining part of the curve is fit to a Gaussian distribution (green curve). **b.** The mean Smchd1 nearest neighbour separation for each image nucleus is plotted against the logarithm of the p-value from the Kolmogorov-Smirnov test. A threshold of $p < 0.00001$ was chosen. **c.** Physical representation of the structure used for the simulations. Molecule positions (green and red points) were generated using lateral and axial localization precisions corresponding to the experimental measurements. The antibody (short gray bars) length, a , was set to 10 nm and a 3D angle between 0 and ϕ was randomly generated. Nearest neighbour distances were calculated for 100,000 iterations of randomly generated molecules under a range of different conditions. The conditions that were varied for each set of simulations were the protein length (long gray bar), p , and maximum angle range, ϕ . Setting a threshold of $p < 10^{-5}$, resulted in probe separations of 75 ± 18 nm and 83.0 ± 17 nm (mean $\pm$ standard deviation)

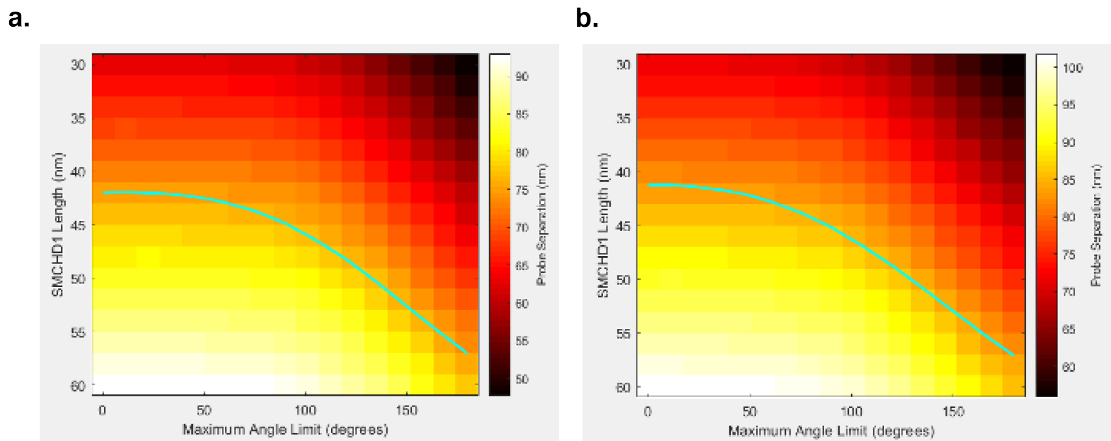


FIGURE 5.14: dSTORM fluorescent Probe separation

a. Results from the simulations performed using mean localization precisions and data from experiments performed on iMEFs. The calculated probe separation as a function of the probe length and maximum angle limit for antibody length of 10 nm. The heat scale (red to yellow) indicates the number of simulations that gave each length for each maximum angle limit. The cyan line represents the experimentally measured 83 nm probe separation. **b.** The same results performed for different replicate taken on a different day. The cyan line corresponds to 75 nm mean probe separation.

TABLE 5.1: Smchd1 is approximately 40 nm in length in both primary and immortalised cells

	iMEFs	NSCs
Total Number of Cells Imaged	30	28
Number of Cells with Smchd1 Separation Significantly Different from a Randomized Distribution using Kolmogorov-Smirnov Test	21	16
Mean Separation (um)	0.088	0.093
Standard Separation (um)	16.000	0.040
Standard Error (um)	0.004	0.010
KS p threshold	0.01	0.01
Median Localization Precision (nm, approximate)	21	24
Estimated SMCHD1 length (from simulations using measured data) (nm)	41.8	43.1

5.4 Discussion

Here I have explored a series of advanced microscopy experiments to look at Smchd1 structure and behaviour and discovered several novel findings. First, I explored Smchd1 dynamics in live cells. Proteins are highly dynamic in their native environment, therefore it is essential to study them in live cells as opposed to fixed cells or using genomics, as those approaches will only give information about snapshots in time rather than the whole picture.

I used a well-established technique known as FRAP to understand Smchd1 dynamics in the Xi region during interphase. Here for the first time I reveal that Smchd1 has a slow turnover on the Xi, where it takes more than 40 minutes to recover. This is slow in comparison to other chromatin bound proteins such as HP1 and some polycomb group proteins which show fast recovery in less than 30 seconds [Almeida et al., 2017, Strom et al., 2017]. However, another SMC family protein, cohesin, takes 16h for recovery, making Smchd1's turnover look comparatively rapid [Rhodes et al., 2017]. Interestingly, our data showed faster dynamics in the gain of function mutant MommeD43 compared to wild-type Smchd1-GFP. It was unexpected that the gain of function mutant MommeD43 would have faster dynamics. Since Smchd1 is a silencing protein, one could expect that the "gain-of-function" could relate to more silencing and therefore slower dynamics. However, it could be that Smchd1 silencing is an active process related to Smchd1's ATPase activity and the MommeD43 mutant could somehow enhance this active process. However, we not yet know whether we are observing an active process, that requires Smchd1's ATPase activity. This should be tested by using a Smchd1 ATPase dead mutant, that has currently been developed in the Blewitt lab. We plan to perform FRAP in such samples as above. Therefore FRAP data give insights into what the role may be for this turnover of Smchd1, suggesting that the turnover may be required for Smchd1-dependent silencing. Given the success with this technique it would be interesting to also use it with other mutant forms of Smchd1 (e.g. the ATPase dead mutant), and where relevant expanded to understand [SMCHD1 related disease mutations](#).

At this stage, we cannot apply the same FRAP technique to other regions in the nucleus due to insufficient signal for Smchd1. Therefore, we cannot use FRAP to assess turnover of Smchd1 at autosomal loci. It will be useful to use much stronger fluorophores such as halo tags such as used by Rhodes et al to carry out the same experiments [Rhodes et al., 2017], so that we may analyse the autosomal foci. This approach could be enhanced by labelling specific autosomal regions using dCas9 tagged with a separate fluorophore and a series of gRNAs to direct dCas9 to a specific region of interest e.g. the *Snrpn* imprinted cluster. In this way it might be possible to analyse Smchd1 turnover specifically at autosomal and disease relevant regions of interest.

Given the my novel findings with FRAP, I am interested to validate and complement these data by employing related techniques such as fluorescence correlation microscopy (FCS). FCS can provide data on dynamic movement of proteins in the nucleus, and will also give information regarding the oligomeric state of Smchd1, which may relate to the following discussion on the higher order structures in the nucleus that make up the Smchd1 nuclear speckles. We have an ongoing collaboration with Dr Elizabeth Hinde (Bio physics, University of Melbourne), an expert in FCS to carry out this technique. In our preliminary studies the MommeD43 mutant form shows faster turn over in the nucleus (regions other than Xi) compared to the wild-type. Data from our lab and others have suggested that Smchd1 is an insulator protein [Chen et al., 2015, Jansz et al., 2018a, Wang et al., 2018, Gdula et al., 2019]. One possibility is that the faster turnover we see by FRAP and potentially by FCS could be due to its insulation effect happening during unloading or reloading of the protein on the chromatin which could be giving the gain of function capability to the MommeD43 mutant.

Next, I have used lattice lightsheet microscopy to understand Smchd1 dynamics during cell division. I used lattice lightsheet microscopy as it is a sensitive method that can be used to achieve high quality volumetric data using less laser power (less photo-toxicity) compared to other microscopy methods [Chen et al., 2014]. I again looked at the Xi where Smchd1 is abundant. As a protein that helps to silence the Xi one might expect it to be remain bound to Xi during cell division

similar to other Xi regulators such as *Xist* and Ezh2 [Lee et al., 1999, Aoto et al., 2008]. Unexpectedly, we observed that Smchd1 does not remain bound to the Xi during cell division. This raises a number of questions, one major question being, if Smchd1 binding and unbinding is proven to be an active process, then why would the cell spend energy in removing Smchd1 during mitosis? It would be favourable for it to stay bound to the Xi similar to other regulators. It also raises the question whether there are active modifiers that help in this process of Smchd1 removal and restoration. In this experiment the stage of cell cycle that the Smchd1 is removed and restored is determined by the changes to the chromatin, however in the future it will be interesting to mark the different stages with cell cycle markers to understand the precise timing of when Smchd1 is lost and reloaded.

While we know the pathway that recruits Smchd1 to the Xi, we have not yet identified the specific factor that recruits Smchd1 directly to the Xi during development [Jansz et al., 2018b,a]. Carrying out a screen using cell cycle factors during G2/M transition or screening for post translational modifications during this stage of cell cycle using mass spectroscopy might provide some clues both for what regulates Smchd1 (un)loading specifically during mitosis, but also loading during development. If we could identify such a factor it will be immensely useful as targeting these factors would provide an alternative mode in which to remove or enhance SMCHD1 function in [SMCHD1 related disease](#). Similarly, we can envisage screens to identify factors that alter Smchd1 turnover during interphase, to go beyond considering Smchd1 mutants.

While protein dynamic data is useful in understanding how the proteins behaviour changes the next step would be to narrow down its structure and find the reasons behind change in protein behaviour by taking a closer look at its structure. In order to do so I used the super resolution microscopy technique dSTORM. Using dSTORM I revealed that the length of Smchd1 is approximately 40 nm. This corresponds to previous finding which used negative stain EM on recombinant Smchd1 protein where they found the full length protein to be 40 nm [Brideau et al., 2015]. What is remarkable, is that our experiment was carried out in the unperturbed Smchd1 protein in the nucleus. The similarity in length not only

validates the length of Smchd1 protein but also shows other avenues to explore protein structure and conformation, which is especially relevant because large proteins such as Smchd1 are hard to produce and purify as a recombinant full length protein, which is required for X-ray crystallography or Cryo-EM.

In our dSTORM analysis we make multiple assumptions, one of them being the antibody distance is 10nm which might not be precise. Therefore, we are currently in the process of developing nano bodies to further optimise this protocol. Nanobodies have proven to produce high quality dSTORM imaging [Virant et al., 2018]. Eliminating primary and secondary antibodies will give more accurate distance measurements.

While it was difficult to optimize the dSTORM protocol, now that it has been optimized it is a reliable method that can be used into wide variety of approaches. We could consider additional simulations as well as imaging of mutant Smchd1. One consideration for simulations would be to consider Smchd1 not as a rigid rod as we have assumed in our analysis, but rather as a potentially flexible ring similar to other SMC family proteins such as Cohesin and Condensin and could potentially have different formations. It would be interesting to further these experiments by examining mutant forms of Smchd1 and compare the structural conformation. This could raise compelling questions such as whether gain of function or loss of function mutants in Smchd1 change these conformations that could be observed? If so, combining FRAP, FCS and also other techniques such as FILM-FRET [Schmid and Hugel, 2020]], which allows us to understand Smchd1's chromatin binding and occupancy in relation to chromatin compaction, we could explore the role for these potential different conformations.

It would also be interesting to use dSTORM look at different Smchd1 enriched regions in the nucleus to understand whether they form structural compartments. Similar experiments have been carried out on long ncRNA NEAT which forms paraspeckles in the nucleus using another super resolution microscopy technique 3D-SIM [West et al., 2016]. Prior work using novel genomic approaches such as Hi-C shows Smchd1 forms long range chromatin interactions at autosomal clustered

regions [Chen et al., 2015, Jansz et al., 2018a, Wang et al., 2018, Gdula et al., 2019]. But how exactly Smchd1 forms these long range interactions is yet to be discovered. Therefore super resolution microscopy could provide an answer to whether these regions are brought together by a specific higher order arrangement of Smchd1 in clusters. Knowing the architecture of Smchd1 in the nucleus will help us understand its function and how it is altered in Smchd1 mutants.

For all the imaging studies we utilised Smchd1-GFP mouse that contains an endogenous GFP tag at the N terminus of Smchd1 protein which functionally identical to the wild type Smchd1 protein as observed by genomic assays (data not shown). However, it'll be interesting to estimate the effect on protein turn over due to the GFP-tag at the N-terminal end. All studies here have been performed in mouse cells; however to consider whether what is found is truly relevant to therapy, it is first necessary to ask whether human SMCHD1 behaves the same as mouse Smchd1. To start with we could use 293T cells (human cell line) as they can be easily manipulated to carry out above mentioned microscopy approaches. Next, we could proceed to using patient derived cells such as human iPSCs (induced pluripotent stem cells) to differentiate into relevant cell types for imaging and screening.

Taken together these advance microscopy methods provides multiple new avenues to understand the structure and function of Smchd1 protein. Using these techniques we could potentially narrow down the exact mechanism by which Smchd1 works, and and provide novel approaches for targeted therapy.

Chapter 6

General Discussion

Smchd1 was identified based on its role in transgene silencing and characterised as an epigenetic modifier important in XCI in 2008 [Blewitt et al., 2005, 2008]. In a short span of time more discoveries were made on SMCHD1 and its relevance in two distinct human diseases [Lemmers et al., 2012, Gordon et al., 2017]. Therefore, there has been growing interest in understanding Smchd1's role in different regions in the genome and its mechanism of function. Much of this work is focused on the zygotic form of Smchd1 and particularly its role in XCI [Jansz et al., 2018b,a, Wang et al., 2018, 2019, Gdula et al., 2019]. However, to date there has not been much information on maternal Smchd1 or its role in imprinted regions. In addition, there is a gap in the understanding of Smchd1's molecular mechanism at a cellular level. During my PhD I shed some light into these aspects of Smchd1: first exploring the requirement for maternal Smchd1 in imprinted gene regulation; and second by exploring novel imaging approaches to understand how Smchd1 carries out its function in the nucleus.

6.0.1 Aim 1: To evaluate whether Smchd1 is a maternal effect protein that has a role in genomic imprinting

To date any and all information on Smchd1 was of its zygotic form, and the role of the maternal form of Smchd1 was only hinted by Lathman Lab [Midic

et al., 2018]. But these previous studies on the maternal form of Smchd1 does not involve comprehensive genomics and mouse genetic approaches to look at its role as a maternal effect protein. Maternal proteins play crucial roles during early embryonic development. Absence of certain maternal proteins such as *Dnmt10*, *Trim28*, *Stella* and others can lead to embryonic lethality. But perhaps what is more crucial is maternal effect proteins that play roles by hindering development causing variable effects later in the development of the embryo. As stated in my introduction, only a limited set of maternal effect genes have been tested and play roles in imprinted gene regulation. In chapter three I explored Smchd1's function as a maternal effect gene, using mouse models of maternal deletion and allele-specific genomics. First, I found that maternal Smchd1 is present until E2.75 stage of embryo development (Figure 6.1a). Next, I discovered that maternal Smchd1 does indeed regulate a set of autosomal imprinted genes.

Until this discovery our knowledge on Smchd1's function at imprinted genes was limited to zygotic Smchd1 and its role in *Snrpn* cluster (*Ndn*, *Peg12*, *Magel2*, and *Mkrn3* genes) and *Igf2r-Airn* cluster genes (*Slc22a3*) [Mould et al., 2013, Gendrel et al., 2013, Chen et al., 2015]. I was not only able to expand this list of imprinted genes regulated by zygotic Smchd1 but importantly I also found imprinted genes that are affected by maternal deletion of Smchd1 until at least E14.5. I specifically found that maternal deletion of Smchd1 not only influenced autosomal imprinted genes but also iXCI in E14.5 placentae, more than 11 days after zygotic Smchd1 has been switched on. These data suggest that maternal Smchd1 sets up an epigenetic memory, and in its absence there are long-lived effects on gene expression. In future, analysing genomic data from reciprocal cross where Smchd1 is paternally deleted, will further validate that effects we see in these genes are solely due to maternal effect.

Apart from the effects of loss of Smchd1 at mid-gestation, I was also able to analyse the effect of maternal Smchd1 during the preimplantation period, when it is having its direct effects. The most striking change I found in E2.75 embryos was relaxation of imprinting at the *Xist* long ncRNA expression, in both male and female embryos. In addition, we found relaxation of imprinting at a few additional

autosomal imprinted genes, however at least one *Smchd1*-sensitive imprinted gene did not yet show relaxation of imprinting at this time. Since I found that zygotic *Smchd1* also plays a role at many of these genes, it seems that maternal *Smchd1* may establish a chromatin state that is subsequently maintained by zygotic *Smchd1*.

I discovered that *Smchd1* regulates imprinted genes that belong to both classical canonical imprinted clusters in addition to rather novel non-canonical imprinted clusters. Interestingly, unlike many known maternal effect genes that regulate imprinted gene expression, maternal deletion of *Smchd1* does not affect primary DNA methylation imprints, either preimplantation or later in gestation. Since

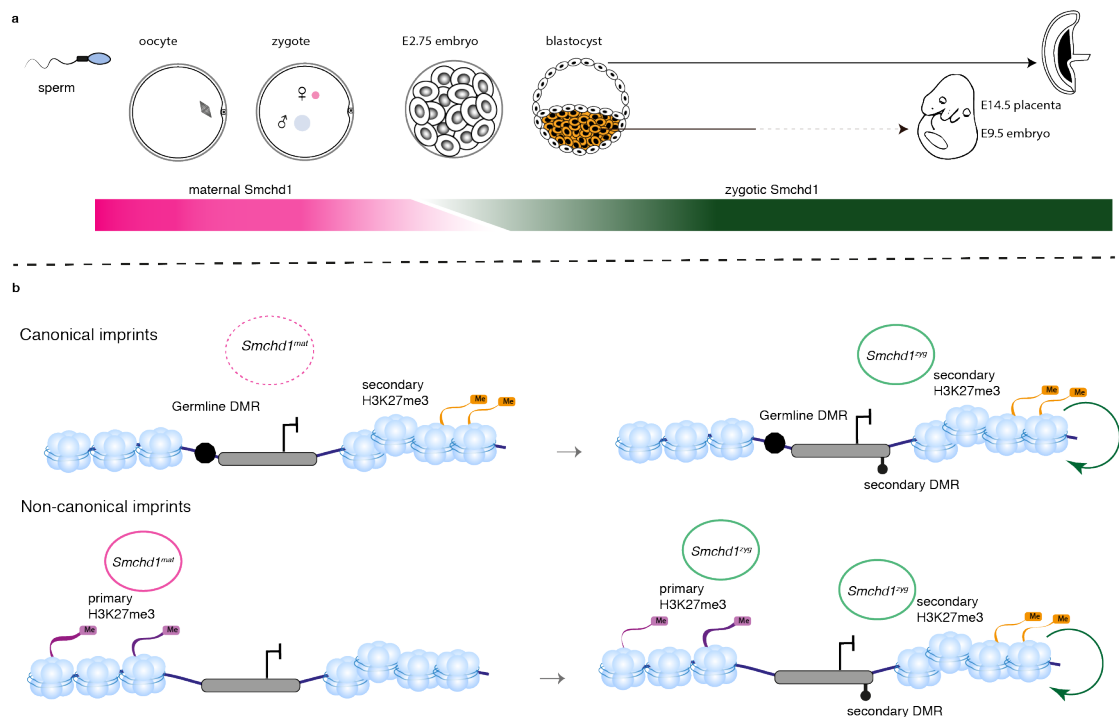


FIGURE 6.1: Maternal *Smchd1* translates the imprints to establish a heritable chromatin state required for imprinted expression.

a. Developmental windows of activity of **maternal** and **zygotic** *Smchd1*. **b.** Proposed model illustrating the regulation of imprinted genes by *Smchd1*. Both **maternal** and **zygotic** *Smchd1* contribute to an epigenetic memory downstream of polycomb repressive histone marks.

the majority of imprinted genes affected by maternal deletion of *Smchd1* are non-canonical imprinted genes (including *Xist*) it is interesting to consider where *Smchd1* could integrate into the known mechanisms that control non-canonical imprinted gene expression. Thus far only the *Eed* and the relevant demethylase *Kdm6b* are known to regulate non-canonical imprinted gene expression regulation [Inoue et al., 2017b,a, Harris et al., 2019]. *Eed* was also discovered to regulate *Xist* expression. Non-canonical imprinted genes including *Xist* are regulated via primary H3K27me3 imprints, laid down by PRC2. With this previous work in mind when we assess our current findings, it is possible that *Smchd1* acts subsequent to primary H3K27me3 imprints (Figure 6.1 b). This hypothesis is supported by several key pieces of data. Firstly, we observe the most striking relaxation of imprinting at both non-canonical and canonical imprinted genes is in *Smchd1* zygotic-deleted (*zygΔ*) samples. Because these samples do not affect *Smchd1* levels in preimplantation development (prior to E2.75) there should be no effect on primary H3K27me3 imprints. Previous work shows that zygotic null samples did not affect H3K27me3 imprints at canonical imprinted genes that are sensitive to *Smchd1*, that contains a secondary H3K27me3 imprints [Chen et al., 2015, Jansz et al., 2018a]. Previous data also shows that *Smchd1* is recruited to the Xi via PRC mediated pathway [Jansz et al., 2018b, Wang et al., 2019]. Therefore, it is possible that a similar pathway could also exist at imprinted genes to recruit *Smchd1*. In future, this hypothesis can be validated by analysing enrichment of H3K27me3 using ChIP-seq or CUT&RUN in relevant tissues in the absence of maternal *Smchd1*.

Our lab and others have previously demonstrated that *Smchd1* is an insulator protein that helps to form long range chromatin interactions at its targeted regions including imprinted genes [Jansz et al., 2018a, Gdula et al., 2019, Wang et al., 2018, 2019]. Absence of *Smchd1* causes increase in Ctf binding, and increase in transcriptionally active histone marks H3K4me2 or H3K4me3 [Chen et al., 2015, Mould et al., 2013, Gdula et al., 2019, Wang et al., 2018], and indeed even an increase in H3K27me3 on the Xi [Jansz et al., 2018a, Wang et al., 2018]. This led me to hypothesise that the presence of maternal *Smchd1* could potentially

insulate against other epigenetic factors. Therefore, it is possible that maternal Smchd1 reads select imprinted genes based in part on polycomb marks, and transforms their chromatin state into an epigenetic memory required for their imprinted expression. This chromatin state is then maintained by the presence of zygotic Smchd1 (Figure 6.1 b).

6.0.2 Aim 2: To unravel the Smchd1 macro molecular structure and dynamics in the nucleus

While the hypothesis I arrived in aim one is interesting and is supported by a number of novel genomic approaches, it still remains unanswered how exactly the Smchd1 protein mediates its function. Genomic approaches are a wonderful way to distinguish the mechanistic potential of a protein; however the protocols tend to treat cells with harsh chemicals during preparation and tend to disrupt the natural environment of the cell. Therefore, in my second aim I set out to explore how the Smchd1 protein behaves in as close as possible to an unperturbed state.

I used live cell microscopy techniques to understand Smchd1 dynamics, particularly the timing of Smchd1 unloading and loading during mitosis, and binding and unbinding during interphase, both on the Smchd1-rich Xi. What is quite important is that I was able to compare the wild-type Smchd1 dynamics during interphase to a gain of function mutant form MommeD43. By including the mutant I was able to obtain data suggesting that binding and unbinding likely influence function; highlighting the importance of using these novel approaches to tackle how different mutant forms alter function. There was not sufficient scope within this study to look at all other disease related mutants but we now have a foundation to look at the impact and the consequences it may have.

Interestingly, there are quite remarkable studies that uses Lattice LightSheet microscopy to visualise mouse embryo development without disrupting the embryo health [Chen et al., 2014, McDole et al., 2018]. Therefore this approach could potentially be used to look into how maternal Smchd1 dynamics change during

development. By partnering with different mutants in Smchd1 or other epigenetic factors, we could test how Smchd1 dynamics contribute to its function as a maternal effect gene, and also how it is recruited to imprinted sites.

The next step was to understand the structure of the Smchd1 protein in the nucleus. To do this I used super resolution microscopy. During my PhD I was able to successfully implement this technique, and ask a basic structural question. But this method can be expanded further to understand the secondary structure that the unperturbed protein forms within the nucleus and how it may or may not change in mutant forms. It will also be interesting to track the Smchd1 foci changes through the development, by carrying out dSTORM in oocyte, and developing preimplantation embryo. By taking this approach we may be able to obtain insights into how maternal Smchd1 modulates the chromatin in the oocyte or preimplantation. Perhaps Smchd1 forms foci based on a higher order chromatin structure that brings together many maternal Smchd1-regulated genes; something that we could observe using dSTORM.

In conclusion, the work I have presented in my thesis has not only contributed to our current understanding on Smchd1 function but also opened up new avenues and techniques to facilitate future discoveries on this fascinating epigenetic modifier. This information could be crucial in tackling SMCHD1 related diseases mutations and potentially lead to development of therapeutic solutions.

Appendix A

Differential imprinted expression in *Smchd1* maternal null vs wildtype male E2.75 embryos

Differential imprinted methylation in E2.75 male embryos, *Smchd1* mat Δ vs wt. A methylated allele is called differentially methylated if the multiple-testing corrected p-value is below 0.05.

Differential imprinted expression in male E2.75 B6Cast embryos, 5mchd1 mat α vs wt.											
ensembl_gene_id	gene	chromosome	location	cluster	expressed_allele	expressed	imprinted	sufficient_counts	padj		diff_imprinted
ENSMUSG00000030199	Etv6	6	6 G1	dist 6	P	TRUE	TRUE	TRUE	4.63966E-05		TRUE
ENSMUSG00000025764	Jade1	3	3 B	cent 3	P	TRUE	TRUE	TRUE	7.71948E-05		TRUE
ENSMUSG00000100826	Shhg14	7	7 C-B5	cent 7	P	TRUE	TRUE	TRUE	0.03569431		TRUE
ENSMUSG00000019188	H13	2	2 H1	dist 2	isoform-specific	TRUE	TRUE	TRUE	0.042176602		TRUE
ENSMUSG00000040013	Fkbp6	5	5 G2	misc 5	P	TRUE	TRUE	TRUE	0.000291215		FALSE
ENSMUSG00000024140	Epas1	17	17 E4	cent 17	P	TRUE	TRUE	TRUE	0.021662138		FALSE
ENSMUSG00000068129	Cst7	2	2 NA	dist 2	M	TRUE	TRUE	TRUE	0.068580741		FALSE
ENSMUSG00000000948	Gm38393	7	7 C-B5	cent 7	P	TRUE	TRUE	TRUE	0.208532987		FALSE
ENSMUSG00000037280	Galnt6	15	15 F1	misc 15	P	TRUE	TRUE	TRUE	0.260034168		FALSE
ENSMUSG00000021848	Otv2	14	14 C1	prox 14	P	TRUE	TRUE	TRUE	0.693991137		FALSE
ENSMUSG00000061186	Sfmbt2	2	2 A1	prox 2	P	TRUE	TRUE	TRUE	0.77861518		FALSE
ENSMUSG00000082536	Gm13456	2	2 NA	prox 2	M	TRUE	TRUE	TRUE	0.77861518		FALSE
ENSMUSG00000042814	Mcts2	2	2 H1	dist 2	P	TRUE	TRUE	TRUE	0.77861518		FALSE
ENSMUSG00000032359	Ctsh	9	9 E3.1	dist 9	P	TRUE	TRUE	TRUE	0.77861518		FALSE
ENSMUSG00000020634	Ubxn2a	12	12 NA	prox 12	P	TRUE	TRUE	TRUE	0.77861518		FALSE
ENSMUSG00000015143	Actn1	12	12 C3	misc 12	P	TRUE	TRUE	TRUE	0.77861518		FALSE
ENSMUSG00000022641	Bbx	16	16 B5	cent 16	P	TRUE	TRUE	TRUE	0.77861518		FALSE
ENSMUSG00000023861	Mpc1	17	17 A1	prox 17	M	TRUE	TRUE	TRUE	0.77861518		FALSE
ENSMUSG00000062078	Ok	17	17 A1	prox 17	M	TRUE	TRUE	TRUE	0.77861518		FALSE
ENSMUSG00000024423	Impact	18	18 A2-B2	cent 18	P	TRUE	TRUE	TRUE	0.77861518		FALSE
ENSMUSG00000017843	Ppp2r5c	12	12 F1	dist 12	P	TRUE	TRUE	TRUE	0.833044374		FALSE
ENSMUSG00000028859	Csf3r	4	4 NA	misc 4	M	TRUE	TRUE	TRUE	0.856989057		FALSE
ENSMUSG00000031665	Sall1	8	8 NA	misc 8	P	TRUE	TRUE	TRUE	0.895466485		FALSE
ENSMUSG00000037221	Mospd3	5	5 G2	misc 5	P	TRUE	TRUE	TRUE		1	FALSE
ENSMUSG00000002265	Peg3	7	7 A2-B1	prox 7	P	TRUE	TRUE	TRUE		1	FALSE
ENSMUSG00000102252	Snrpn	7	7 C-B5	cent 7	P	TRUE	TRUE	TRUE		1	FALSE
ENSMUSG00000101609	Kcnq1ot1	7	7 F5	dist 7	P	TRUE	TRUE	TRUE		1	FALSE
ENSMUSG00000031714	Gab1	8	8 C2	misc 8	P	TRUE	TRUE	TRUE		1	FALSE
ENSMUSG00000018707	Dync1h1	12	12 F1	dist 12	P	TRUE	TRUE	TRUE		1	FALSE
ENSMUSG00000022464	Slc38a4	15	15 F1	misc 15	P	TRUE	TRUE	TRUE		1	FALSE
ENSMUSG00000023830	Igf2r	17	17 A1	prox 17	M	TRUE	TRUE	TRUE		1	FALSE
ENSMUSG00000025967	Eef1b2	1	1 C2	cent 1	P	TRUE	FALSE	FALSE	NA		FALSE
ENSMUSG00000027520	Zdbf2	1	1 C2	cent 1	P	TRUE	TRUE	FALSE	NA		FALSE
ENSMUSG00000009876	Cox4i2	2	2 H1	dist 2	P	TRUE	TRUE	FALSE	NA		FALSE
ENSMUSG00000027523	Gnas	2	2 H4	dist 2	M	TRUE	TRUE	FALSE	NA		FALSE
ENSMUSG00000051855	Mest	6	6 A3	prox 6	isoform-specific	TRUE	TRUE	FALSE	NA		FALSE
ENSMUSG00000025607	Copg2	6	6 A3	prox 6	M	TRUE	FALSE	FALSE	NA		FALSE
ENSMUSG00000045752	Tssc4	7	7 F5	dist 7	M	TRUE	FALSE	FALSE	NA		FALSE
ENSMUSG00000010760	Phlda2	7	7 F5	dist 7	M	TRUE	FALSE	FALSE	NA		FALSE
ENSMUSG00000040111	Gramd1b	9	9 A5.1	misc 9	P	TRUE	TRUE	FALSE	NA		FALSE
ENSMUSG00000019795	Pcm1	10	10 A1	misc 10	P	TRUE	FALSE	FALSE	NA		FALSE
ENSMUSG00000019814	Ltv1	10	10 NA	cent 10	M	TRUE	FALSE	FALSE	NA		FALSE
ENSMUSG00000025348	Itga7	10	10 NA	dist 10	M	TRUE	TRUE	FALSE	NA		FALSE
ENSMUSG00000061469	Gm12569	11	11 B1.3	misc 11	P	TRUE	TRUE	FALSE	NA		FALSE
ENSMUSG00000021266	Wars	12	12 F1	dist 12	P	TRUE	FALSE	FALSE	NA		FALSE
ENSMUSG00000061517	Sox21	14	14 E4	misc 14	P	TRUE	TRUE	FALSE	NA		FALSE
ENSMUSG00000022139	Mbnl2	14	14 E4	misc 14	P	TRUE	TRUE	FALSE	NA		FALSE
ENSMUSG00000068391	Chrac1	15	15 D3	cent 15	M	TRUE	FALSE	FALSE	NA		FALSE
ENSMUSG00000036698	Ago2	15	15 D3	cent 15	M	TRUE	FALSE	FALSE	NA		FALSE
ENSMUSG00000023169	Slc38a1	15	15 F1	misc 15	P	TRUE	TRUE	FALSE	NA		FALSE
ENSMUSG00000022462	Slc38a2	15	15 F1	misc 15	P	TRUE	TRUE	FALSE	NA		FALSE
ENSMUSG00000022952	Runx1	16	16 C4	misc 16	P	TRUE	TRUE	FALSE	NA		FALSE
ENSMUSG00000073460	Pnlcd1	17	17 NA	prox 17	M	TRUE	FALSE	FALSE	NA		FALSE
ENSMUSG00000067049	Unc93a	17	17 NA	prox 17	M	TRUE	TRUE	FALSE	NA		FALSE
ENSMUSG00000046856	Gpr1	1	1 C2	cent 1	P	FALSE	NA	FALSE	NA		NA
ENSMUSG00000025964	Adam23	1	1 C2	cent 1	P	FALSE	NA	FALSE	NA		NA
ENSMUSG00000026395	Ptprc	1	1 NA	misc 1	M	FALSE	NA	FALSE	NA		NA
ENSMUSG00000026358	Rgs1	1	1 NA	misc 1	M	FALSE	NA	FALSE	NA		NA
ENSMUSG00000026471	Mr1	1	1 G3	misc 1	M	FALSE	NA	FALSE	NA		NA
ENSMUSG00000026470	Stx6	1	1 G3	misc 1	M	TRUE	TRUE	FALSE	NA		NA
ENSMUSG00000033722	BCD34090	1	1 G3	misc 1	M	FALSE	NA	FALSE	NA		NA
ENSMUSG00000026581	Sell	1	1 NA	misc 1	M	FALSE	NA	FALSE	NA		NA
ENSMUSG00000073494	Sh2d1b2	1	1 NA	misc 1	M	FALSE	NA	FALSE	NA		NA
ENSMUSG00000038421	Fcrla	1	1 NA	misc 1	M	FALSE	NA	FALSE	NA		NA
NA	Zdbf2linc	1	1 C2	cent 1	P	FALSE	NA	FALSE	NA		NA
NA	XLOC_04784	2	2	prox 2	P	FALSE	NA	FALSE	NA		NA
ENSMUSG00000086748	Gm13261	2	2 NA	prox 2	P	FALSE	NA	FALSE	NA		NA
ENSMUSG00000026770	Il2ra	2	2 NA	prox 2	M	FALSE	NA	FALSE	NA		NA
ENSMUSG00000087701	Gm13493	2	2 NA	prox 2	M	TRUE	NA	FALSE	NA		NA
ENSMUSG00000026970	Rbms1	2	2 C1.2	prox 2	P	TRUE	TRUE	FALSE	NA		NA
ENSMUSG00000081833	Gm13669	2	2 NA	prox 2	M	TRUE	TRUE	FALSE	NA		NA
ENSMUSG00000016458	Wt1	2	2 E2-E3	prox 2	M	FALSE	NA	FALSE	NA		NA
ENSMUSG00000040093	Bmf	2	2 NA	prox 2	M	TRUE	FALSE	FALSE	NA		NA
ENSMUSG00000042745	Id1	2	2 H1	dist 2	P	TRUE	TRUE	FALSE	NA		NA
ENSMUSG00000007659	Bcl2l1	2	2 H1	dist 2	P	TRUE	FALSE	FALSE	NA		NA
ENSMUSG00000027469	Tpx2	2	2 H1	dist 2	P	TRUE	FALSE	FALSE	NA		NA
ENSMUSG00000067787	Bicap	2	2 H1	dist 2	isoform-specific	TRUE	TRUE	FALSE	NA		NA
ENSMUSG00000067786	Nnat	2	2 H1	dist 2	P	TRUE	NA	FALSE	NA		NA
ENSMUSG00000074622	Matfb	2	2 H2	dist 2	M	FALSE	NA	FALSE	NA		NA
ENSMUSG00000027551	Zfp64	2	2 H3	dist 2	P	TRUE	TRUE	FALSE	NA		NA
ENSMUSG0000008999	Bmp7	2	2 H3	dist 2	P	FALSE	NA	FALSE	NA		NA
ENSMUSG00000027522	Stx16	2	2 H4	dist 2	M	TRUE	TRUE	FALSE	NA		NA
ENSMUSG00000039263	Npepl1	2	2 H4	dist 2	M	TRUE	TRUE	FALSE	NA		NA
ENSMUSG00000065521	Mir296	2	2 H4	dist 2	P	FALSE	NA	FALSE	NA		NA
ENSMUSG00000065410	Mir298	2	2 H4	dist 2	P	FALSE	NA	FALSE	NA		NA
ENSMUSG00000086537	Nespas	2	2 H4	dist 2	P	TRUE	NA	FALSE	NA		NA
ENSMUSG00000027524	Edn3	2	2 H4	dist 2	P	FALSE	NA	FALSE	NA		NA

NA	NESP55	2	2 H4	dist 2	M	FALSE	NA	FALSE	NA	NA
NA	GNASXL	2	2 H4	dist 2	P	FALSE	NA	FALSE	NA	NA
NA	Exon-1A	2	2 H4	dist 2	P	FALSE	NA	FALSE	NA	NA
NA	GS-alpha	2	2 H4	dist 2	M	FALSE	NA	FALSE	NA	NA
NA	F7	2	2 H4	dist 2	M	FALSE	NA	FALSE	NA	NA
ENSMUSG00000001865	Cpa3	3	NA	misc 3	M	FALSE	NA	FALSE	NA	NA
ENSMUSG00000105260	Gm40040	3	NA	cent 3	P	TRUE	TRUE	FALSE	NA	NA
ENSMUSG000000097639	Platr4	3	3 B	cent 3	P	TRUE	TRUE	FALSE	NA	NA
ENSMUSG00000106223	2400006e01	3	3 B	cent 3	P	TRUE	TRUE	FALSE	NA	NA
ENSMUSG000000027955	Gask1b	3	3 E3	misc 3	P	TRUE	TRUE	FALSE	NA	NA
ENSMUSG000000028078	Dclk2	3	NA	misc 3	M	TRUE	TRUE	FALSE	NA	NA
ENSMUSG000000056071	S100a9	3	NA	misc 3	M	FALSE	NA	FALSE	NA	NA
ENSMUSG000000087119	Atg4a-ps	3	NA	misc 3	M	FALSE	NA	FALSE	NA	NA
ENSMUSG000000028339	Col15a1	4	NA	misc 4	M	FALSE	NA	FALSE	NA	NA
ENSMUSG000000083219	Gm11410	4	NA	misc 4	M	TRUE	TRUE	FALSE	NA	NA
ENSMUSG000000028581	Laptm5	4	NA	misc 4	M	FALSE	NA	FALSE	NA	NA
ENSMUSG000000029096	Htra3	5	5 B3	prox 5	M	FALSE	NA	FALSE	NA	NA
ENSMUSG000000029304	Spp1	5	NA	misc 5	M	FALSE	NA	FALSE	NA	NA
ENSMUSG000000097287	D130017N08	5	5 G2	misc 5	P	FALSE	NA	FALSE	NA	NA
ENSMUSG000000042817	Flt3	5	NA	misc 5	P	FALSE	NA	FALSE	NA	NA
ENSMUSG000000023964	Calcr	6	6 A1	prox 6	M	FALSE	NA	FALSE	NA	NA
ENSMUSG000000029664	Tfp12	6	6 A1	prox 6	M	FALSE	NA	FALSE	NA	NA
ENSMUSG000000015189	Casd1	6	6 A1	prox 6	M	TRUE	TRUE	FALSE	NA	NA
ENSMUSG000000004631	Sgce	6	6 A1	prox 6	P	TRUE	TRUE	FALSE	NA	NA
ENSMUSG000000092035	Peg10	6	6 A1	prox 6	P	FALSE	NA	FALSE	NA	NA
ENSMUSG000000032827	Ppp1r9a	6	6 A1	prox 6	M	TRUE	TRUE	FALSE	NA	NA
ENSMUSG000000042607	Asb4	6	6 A1	prox 6	M	FALSE	NA	FALSE	NA	NA
ENSMUSG000000029775	Klhdcl10	6	6 A3	prox 6	M	TRUE	TRUE	FALSE	NA	NA
ENSMUSG000000065501	Mir335	6	6 A3	prox 6	P	FALSE	NA	FALSE	NA	NA
ENSMUSG000000039032	Tsga13	6	6 A3.3	prox 6	M	FALSE	NA	FALSE	NA	NA
ENSMUSG000000073209	Klf14	6	6 A3	prox 6	M	FALSE	NA	FALSE	NA	NA
ENSMUSG000000076490	Trbc1	6	NA	prox 6	M	FALSE	NA	FALSE	NA	NA
ENSMUSG000000029804	Herc3	6	6 B3	prox 6	isoform-specific	TRUE	TRUE	FALSE	NA	NA
ENSMUSG000000055430	Nap1l5	6	6 C1	prox 6	P	FALSE	NA	FALSE	NA	NA
ENSMUSG000000025889	Snca	6	NA	prox 6	M	FALSE	NA	FALSE	NA	NA
ENSMUSG000000030114	Klrg1	6	NA	dist 6	M	FALSE	NA	FALSE	NA	NA
ENSMUSG000000108090	Gm6637	6	NA	dist 6	M	FALSE	NA	FALSE	NA	NA
NA	Cogp2os1	6	6 A3	prox 6	P	FALSE	NA	FALSE	NA	NA
ENSMUSG00000108113	Zim2	7	7 A2-B1	prox 7	M	FALSE	NA	FALSE	NA	NA
ENSMUSG000000002266	Zim1	7	7 A2-B1	prox 7	M	FALSE	NA	FALSE	NA	NA
ENSMUSG0000000051527	Usp29	7	7 A2-B1	prox 7	P	TRUE	FALSE	FALSE	NA	NA
ENSMUSG00000108043	Zim3	7	7 A2-B1	prox 7	M	FALSE	NA	FALSE	NA	NA
ENSMUSG0000000109176	Zfp264	7	7 A2-B1	prox 7	P	TRUE	TRUE	FALSE	NA	NA
ENSMUSG000000002602	Axl	7	7 A3-B1	misc 7	M	FALSE	NA	FALSE	NA	NA
ENSMUSG000000030579	Tyrobp	7	NA	misc 7	M	FALSE	NA	FALSE	NA	NA
ENSMUSG000000025326	Ube3a	7	7 C-B5	cent 7	M	TRUE	FALSE	FALSE	NA	NA
ENSMUSG000000077191	Snord64	7	7C-B5	cent 7	P	FALSE	NA	FALSE	NA	NA
ENSMUSG000000089375	Snord107	7	7C-B5	cent 7	P	FALSE	NA	FALSE	NA	NA
ENSMUSG00000102627	Snurf	7	7 C-B5	cent 7	P	FALSE	NA	FALSE	NA	NA
NA	XLOC_07595	7	7	cent 7	P	FALSE	NA	FALSE	NA	NA
ENSMUSG00000109006	B230209E15	7	7 C-B5	cent 7	P	FALSE	NA	FALSE	NA	NA
ENSMUSG00000109394	A230057D08	7	7 C-B5	cent 7	P	FALSE	NA	FALSE	NA	NA
ENSMUSG00000109321	A330076H08	7	7 C-B5	cent 7	P	FALSE	NA	FALSE	NA	NA
ENSMUSG000000045813	Gm9801	7	7 C-B5	cent 7	P	FALSE	NA	FALSE	NA	NA
ENSMUSG000000033585	Ndn	7	7 C-B5	cent 7	P	TRUE	TRUE	FALSE	NA	NA
ENSMUSG000000056972	Magel2	7	7 C-B5	cent 7	P	FALSE	NA	FALSE	NA	NA
ENSMUSG000000070527	Mktn3	7	7 C-B5	cent 7	P	TRUE	TRUE	FALSE	NA	NA
ENSMUSG000000070526	Peg12	7	7 C-B5	cent 7	P	FALSE	NA	FALSE	NA	NA
ENSMUSG000000030638	Sh3gl3	7	7 D3	misc 7	P	FALSE	NA	FALSE	NA	NA
ENSMUSG000000030724	Cd19	7	NA	distcent 7	M	FALSE	NA	FALSE	NA	NA
ENSMUSG000000030782	Tgfb11	7	7 F3	distcent 7	M	FALSE	NA	FALSE	NA	NA
ENSMUSG000000030847	Bag3	7	7 F3	distcent 7	M	TRUE	FALSE	FALSE	NA	NA
ENSMUSG000000042105	Inpp5f	7	7 F3	distcent 7	P	TRUE	TRUE	FALSE	NA	NA
ENSMUSG000000045777	Ifitm10	7	7 F5	dist 7	M	FALSE	NA	FALSE	NA	NA
ENSMUSG000000018819	Lsp1	7	7 F5	dist 7	M	FALSE	NA	FALSE	NA	NA
ENSMUSG000000087090	Nctc1	7	7 F5	dist 7	P	FALSE	NA	FALSE	NA	NA
ENSMUSG000000000031	H19	7	7 F5	dist 7	M	FALSE	NA	FALSE	NA	NA
ENSMUSG000000048583	Igf2	7	7 F5	dist 7	P	FALSE	NA	FALSE	NA	NA
ENSMUSG000000086266	Igf2os	7	7 F5	dist 7	P	FALSE	NA	FALSE	NA	NA
ENSMUSG000000000215	Ins2	7	7 F5	dist 7	P	FALSE	NA	FALSE	NA	NA
ENSMUSG000000000214	Th	7	7 F5	dist 7	M	FALSE	NA	FALSE	NA	NA
ENSMUSG000000009248	Ascl2	7	7 F5	dist 7	M	FALSE	NA	FALSE	NA	NA
ENSMUSG0000000059277	R74862	7	7 F5	dist 7	M	FALSE	NA	FALSE	NA	NA
ENSMUSG0000000037706	Cd81	7	7 F5	dist 7	M	FALSE	NA	FALSE	NA	NA
ENSMUSG000000009545	Kcnq1	7	7 F5	dist 7	M	FALSE	NA	FALSE	NA	NA
ENSMUSG000000037664	Cdkn1c	7	7 F5	dist 7	M	TRUE	TRUE	FALSE	NA	NA
NA	AF313042	7	7 F5	dist 7	M	FALSE	NA	FALSE	NA	NA
ENSMUSG000000000154	Slc22a18	7	7 F5	dist 7	M	FALSE	NA	FALSE	NA	NA
ENSMUSG000000059119	Nap1l4	7	7 F5	dist 7	M	TRUE	FALSE	FALSE	NA	NA
ENSMUSG0000000045362	Tnfrsf26	7	7 F5	dist 7	M	FALSE	NA	FALSE	NA	NA
ENSMUSG00000109251	E230032D23	7	7	dist 7	M	FALSE	NA	FALSE	NA	NA
ENSMUSG000000010751	Tnfrsf22	7	7 F5	dist 7	M	TRUE	TRUE	FALSE	NA	NA
ENSMUSG000000037613	Tnfrsf23	7	7 F5	dist 7	M	TRUE	TRUE	FALSE	NA	NA
ENSMUSG000000037606	Osbpl5	7	7 F5	dist 7	M	TRUE	TRUE	FALSE	NA	NA
ENSMUSG000000031075	Ano1	7	7 F5	dist 7	M	FALSE	NA	FALSE	NA	NA
NA	Peg3os	7	7 A1	prox 7	P	FALSE	NA	FALSE	NA	NA
NA	Mktn3os	7	7C-B5	cent 7	P	FALSE	NA	FALSE	NA	NA
NA	BM117114	7	7C-B5	cent 7	P	FALSE	NA	FALSE	NA	NA
NA	Pec2	7	7C-B5	cent 7	P	FALSE	NA	FALSE	NA	NA

NA	BB077283	7	7C-B5	cent 7	P	FALSE	NA	FALSE	NA	NA
NA	Pec3	7	7C-B5	cent 7	P	FALSE	NA	FALSE	NA	NA
NA	Snord116_c	7	7C-B5	cent 7	P	FALSE	NA	FALSE	NA	NA
NA	Snord115_c	7	7C-B5	cent 7	P	FALSE	NA	FALSE	NA	NA
NA	U08093	7	7C-B5	cent 7	P	FALSE	NA	FALSE	NA	NA
NA	PIHit	7	7 F5	dist 7	P	FALSE	NA	FALSE	NA	NA
NA	Tsc1	8	8	misc 8	P	FALSE	NA	FALSE	NA	NA
ENSMUSG00000031962	Cdh15	8	8 E1	misc 8	isoform-specific	FALSE	NA	FALSE	NA	NA
ENSMUSG00000032094	Cd3d	9	NA	misc 9	M	FALSE	NA	FALSE	NA	NA
ENSMUSG00000032299	Comm4d	9	NA	misc 9	P	TRUE	FALSE	FALSE	NA	NA
ENSMUSG00000032280	Tle3	9	9 B	misc 9	P	TRUE	TRUE	FALSE	NA	NA
ENSMUSG00000040204	Pclaf	9	NA	misc 9	P	FALSE	NA	FALSE	NA	NA
ENSMUSG00000032422	Snx14	9	9 E3.1	dist 9	P	TRUE	TRUE	FALSE	NA	NA
ENSMUSG00000065596	Mir184	9	9 E3.1	dist 9	P	FALSE	NA	FALSE	NA	NA
ENSMUSG00000098627	4930524008	9	9 E3.1	dist 9	P	FALSE	NA	FALSE	NA	NA
ENSMUSG00000032356	Rasgrf1	9	9 E3.1	dist 9	P	TRUE	TRUE	FALSE	NA	NA
ENSMUSG00000032454	Rbp2	9	9 E3.3	dist 9	P	FALSE	NA	FALSE	NA	NA
NA	AS4	9	9 E3.1	dist 9	P	FALSE	NA	FALSE	NA	NA
ENSMUSG00000019817	Plagl1	10	10 A2	cent 10	P	FALSE	NA	FALSE	NA	NA
ENSMUSG00000062866	Phactr2	10	10 A1	cent 10	M	FALSE	NA	FALSE	NA	NA
ENSMUSG00000037202	Prf1	10	NA	mid 10	M	FALSE	NA	FALSE	NA	NA
ENSMUSG00000000290	Itgb2	10	NA	mid 10	M	FALSE	NA	FALSE	NA	NA
ENSMUSG00000025383	Il23a	10	NA	dist 10	M	TRUE	TRUE	FALSE	NA	NA
ENSMUSG00000020182	Odc	11	11 A1	prox 11	P	TRUE	FALSE	FALSE	NA	NA
ENSMUSG00000020176	Grb10	11	11 A1	prox 11	isoform-specific	TRUE	TRUE	FALSE	NA	NA
NA	XLOC_01103	11	11	misc 11	M	FALSE	NA	FALSE	NA	NA
NA	Tsc2	11	11	misc 11	M	FALSE	NA	FALSE	NA	NA
ENSMUSG00000020173	Cobl	11	11 A1	prox 11	M	FALSE	NA	FALSE	NA	NA
ENSMUSG00000051650	B3gnt2	11	11 A3.2	prox 11	M	FALSE	NA	FALSE	NA	NA
ENSMUSG00000051355	Comm1	11	11 A3	prox 11	isoform-specific	TRUE	FALSE	FALSE	NA	NA
ENSMUSG00000044068	Zrsr1	11	11 A3	prox 11	P	TRUE	TRUE	FALSE	NA	NA
ENSMUSG00000020399	Havcr2	11	NA	misc 11	M	FALSE	NA	FALSE	NA	NA
ENSMUSG00000036545	Adamts2	11	11 B1.3	misc 11	P	TRUE	TRUE	FALSE	NA	NA
ENSMUSG00000086646	Platr20	11	11 B1.3	misc 11	P	FALSE	NA	FALSE	NA	NA
ENSMUSG00000045942	BC049762	11	11 B1.3	misc 11	P	FALSE	NA	FALSE	NA	NA
ENSMUSG00000085860	2410003L11	11	11 D	misc 11	P	TRUE	TRUE	FALSE	NA	NA
ENSMUSG00000043439	Epop	11	NA	misc 11	P	TRUE	TRUE	FALSE	NA	NA
ENSMUSG00000035112	Wnk4	11	11 D	misc 11	M	FALSE	NA	FALSE	NA	NA
ENSMUSG00000039963	Ccdc40	11	11 E2	misc 11	P	TRUE	FALSE	FALSE	NA	NA
ENSMUSG00000066632	Pgk1-rs7	12	NA	misc 12	M	FALSE	NA	FALSE	NA	NA
ENSMUSG00000112546	Gm6992	12	NA	misc 12	P	FALSE	NA	FALSE	NA	NA
NA	Tsc4	12	12	misc 12	P	FALSE	NA	FALSE	NA	NA
ENSMUSG00000021136	Smoc1	12	12 D1	misc 12	P	TRUE	TRUE	FALSE	NA	NA
ENSMUSG00000047415	Gpr68	12	12 F1	dist 12	M	FALSE	NA	FALSE	NA	NA
ENSMUSG00000040867	Begain	12	12 F1	dist 12	P	FALSE	NA	FALSE	NA	NA
ENSMUSG00000040856	DLK1	12	12 F1	dist 12	P	FALSE	NA	FALSE	NA	NA
ENSMUSG00000112977	Gm34081	12	12	dist 12	M	FALSE	NA	FALSE	NA	NA
ENSMUSG00000021268	Meg3	12	12 F1	dist 12	M	FALSE	NA	FALSE	NA	NA
ENSMUSG00000065526	Mir337	12	12 F1	dist 12	M	FALSE	NA	FALSE	NA	NA
ENSMUSG00000085925	Rtl1	12	12 F1	dist 12	P	FALSE	NA	FALSE	NA	NA
ENSMUSG00000070080	Mir431	12	12 F1	dist 12	M	FALSE	NA	FALSE	NA	NA
ENSMUSG00000070072	Mir433	12	12 F1	dist 12	M	FALSE	NA	FALSE	NA	NA
ENSMUSG00000070076	Mir127	12	12 F1	dist 12	M	FALSE	NA	FALSE	NA	NA
ENSMUSG00000070133	Mir434	12	12 F1	dist 12	M	FALSE	NA	FALSE	NA	NA
ENSMUSG00000087855	Mir432	12	12 F1	dist 12	M	FALSE	NA	FALSE	NA	NA
ENSMUSG00000104858	Mir136	12	12 F1	dist 12	M	FALSE	NA	FALSE	NA	NA
ENSMUSG00000097451	Rian	12	12 F1	dist 12	M	FALSE	NA	FALSE	NA	NA
ENSMUSG00000065433	Mir370	12	12 F1	dist 12	M	FALSE	NA	FALSE	NA	NA
ENSMUSG00000088294	AF357425	12	12 F1	dist 12	M	FALSE	NA	FALSE	NA	NA
NA	AF357426	12	12 F1	dist 12	M	FALSE	NA	FALSE	NA	NA
NA	AF357359	12	12 F1	dist 12	M	FALSE	NA	FALSE	NA	NA
ENSMUSG00000064621	Gm26922	12	12 F1	dist 12	M	FALSE	NA	FALSE	NA	NA
ENSMUSG00000065749	AF357428	12	12 F1	dist 12	M	FALSE	NA	FALSE	NA	NA
ENSMUSG00000065022	AF357341	12	12 F1	dist 12	M	FALSE	NA	FALSE	NA	NA
ENSMUSG00000065477	Mir411	12	12 F1	dist 12	M	FALSE	NA	FALSE	NA	NA
ENSMUSG00000104582	Mir380	12	12 F1	dist 12	M	FALSE	NA	FALSE	NA	NA
ENSMUSG00000105853	Mir376b	12	12 F1	dist 12	M	FALSE	NA	FALSE	NA	NA
ENSMUSG00000097391	Mirg	12	12 F1	dist 12	M	FALSE	NA	FALSE	NA	NA
ENSMUSG00000065426	Mir134	12	12 F1	dist 12	M	FALSE	NA	FALSE	NA	NA
ENSMUSG00000065448	Mir154	12	12 F1	dist 12	M	FALSE	NA	FALSE	NA	NA
ENSMUSG00000065497	Mir410	12	12 F1	dist 12	M	FALSE	NA	FALSE	NA	NA
ENSMUSG00000075707	Dio3	12	12 F1	dist 12	P	FALSE	NA	FALSE	NA	NA
ENSMUSG00000104213	Ighd	12	NA	dist 12	M	FALSE	NA	FALSE	NA	NA
NA	DLK1 down	12	12 F1	NA	P	FALSE	NA	FALSE	NA	NA
NA	Mico1	12	12 F1	NA	M	FALSE	NA	FALSE	NA	NA
NA	Mico1os	12	12 F1	NA	M	FALSE	NA	FALSE	NA	NA
NA	anti-Rtl1	12	12 F1	dist 12	M	FALSE	NA	FALSE	NA	NA
ENSMUSG00000053101	Gpr141	13	NA	misc 13	M	FALSE	NA	FALSE	NA	NA
ENSMUSG00000018259	Prl8a2	13	NA	cent 13	M	FALSE	NA	FALSE	NA	NA
ENSMUSG00000016477	E2f3	13	13 A3.2	cent 13	P	TRUE	TRUE	FALSE	NA	NA
ENSMUSG00000044231	Nhlrc1	13	13 A5	cent 13	P	FALSE	NA	FALSE	NA	NA
ENSMUSG00000047842	Dras2	13	NA	cent 13	M	FALSE	NA	FALSE	NA	NA
ENSMUSG00000041773	Enc1	13	13 D1	dist 13	P	TRUE	TRUE	FALSE	NA	NA
ENSMUSG00000046082	Tmem174	13	13 D1	dist 13	P	FALSE	NA	FALSE	NA	NA
ENSMUSG00000021699	Pde4d	13	13 D2	dist 13	P	FALSE	NA	FALSE	NA	NA
ENSMUSG00000098682	Otx2os1	14	14 C1	prox 14	P	FALSE	NA	FALSE	NA	NA
ENSMUSG00000040314	Ctsg	14	NA	prox 14	M	FALSE	NA	FALSE	NA	NA
ENSMUSG00000022156	Gzme	14	NA	prox 14	M	FALSE	NA	FALSE	NA	NA
ENSMUSG00000059256	Gzmd	14	NA	prox 14	M	FALSE	NA	FALSE	NA	NA

ENSMUSG00000040284	Gzmg		14	NA	prox_14	M		FALSE	NA	FALSE	NA	NA
ENSMUSG00000015441	Gzmf		14	NA	prox_14	M		FALSE	NA	FALSE	NA	NA
ENSMUSG00000079186	Gzmc		14	NA	prox_14	M		FALSE	NA	FALSE	NA	NA
ENSMUSG00000033470	Cysltr2		14	NA	cent_14	M		FALSE	NA	FALSE	NA	NA
ENSMUSG00000034997	Htr2a		14	14 D2	cent_14	M		FALSE	NA	FALSE	NA	NA
ENSMUSG00000036760	Kcnk9		15	15 D3	cent_15	M		FALSE	NA	FALSE	NA	NA
ENSMUSG00000047921	Traopp9		15	15 D3	cent_15	isoform-specific		TRUE	TRUE	FALSE	NA	NA
ENSMUSG000000106847	Peg13		15	15 D3	cent_15	P		FALSE	NA	FALSE	NA	NA
ENSMUSG00000000489	Pdgfb		15	15 E1	cent_15	P		FALSE	NA	FALSE	NA	NA
ENSMUSG000000116295	Gm32885		15		15 misc_15	P		FALSE	NA	FALSE	NA	NA
ENSMUSG00000022657	Cd96		16	NA	cent_16	M		FALSE	NA	FALSE	NA	NA
ENSMUSG00000022651	Retnlg		16	NA	cent_16	M		FALSE	NA	FALSE	NA	NA
ENSMUSG000000097694	G73001380		16	16 B5	cent_16	P		FALSE	NA	FALSE	NA	NA
ENSMUSG00000022639	Dubr		16	NA	cent_16	P		TRUE	TRUE	FALSE	NA	NA
NA	XLOC_03743		17		17 prox_17	M		FALSE	NA	FALSE	NA	NA
ENSMUSG00000069729	Arid1b		17	17 A1	prox_17	M		TRUE	FALSE	FALSE	NA	NA
ENSMUSG00000034265	Zdhhc14		17	17A1	prox_17	M		FALSE	NA	FALSE	NA	NA
ENSMUSG00000073471	Rsph3a		17	17A1	prox_17	M		FALSE	NA	FALSE	NA	NA
ENSMUSG00000023868	Pde10a		17	17A1	prox_17	M		FALSE	NA	FALSE	NA	NA
ENSMUSG00000023873	1700010114		17	17 A1	prox_17	M		FALSE	NA	FALSE	NA	NA
ENSMUSG00000037196	Pacrg		17	17 A1	prox_17	M		FALSE	NA	FALSE	NA	NA
ENSMUSG00000023826	Prkn		17	17 A1	prox_17	M		TRUE	TRUE	FALSE	NA	NA
NA	XLOC_03763		17		17 prox_17	M		FALSE	NA	FALSE	NA	NA
ENSMUSG00000023828	Sic22a3		17	17 A1	prox_17	M		FALSE	NA	FALSE	NA	NA
ENSMUSG00000040966	Sic22a2		17	17 A1	prox_17	M		FALSE	NA	FALSE	NA	NA
ENSMUSG00000078247	Airn		17	17 A1	prox_17	P		FALSE	NA	FALSE	NA	NA
ENSMUSG000000089769	Gm16574		17	NA	prox_17	P		FALSE	NA	FALSE	NA	NA
ENSMUSG000000089817	Gm7162		17	NA	prox_17	P		FALSE	NA	FALSE	NA	NA
ENSMUSG00000048826	Dact2		17	17 A2	prox_17	M		FALSE	NA	FALSE	NA	NA
ENSMUSG00000023886	Smoc2		17	17 A2	prox_17	M		FALSE	NA	FALSE	NA	NA
ENSMUSG00000023885	Thbs2		17	17 A2	prox_17	M		FALSE	NA	FALSE	NA	NA
ENSMUSG00000036594	H2-Aa		17	NA	misc_17	M		TRUE	TRUE	FALSE	NA	NA
NA	CrossFerre		17	X A5	NA	M		FALSE	NA	FALSE	NA	NA
ENSMUSG00000043458	Pcdhb12		18	18 B3	misc_18	M		FALSE	NA	FALSE	NA	NA
ENSMUSG00000024610	Cd74		18	NA	misc_18	M		TRUE	TRUE	FALSE	NA	NA
ENSMUSG00000044906	4930503L19		18	NA	misc_18	P		TRUE	TRUE	FALSE	NA	NA
ENSMUSG00000075010	AW112010		19	NA	misc_19	M		FALSE	NA	FALSE	NA	NA
ENSMUSG00000024673	Ms4a1		19	NA	misc_19	M		FALSE	NA	FALSE	NA	NA
ENSMUSG00000052387	Trpm3		19	NA	misc_19	M		FALSE	NA	FALSE	NA	NA
ENSMUSG00000048720	Tbc1d12		19	19 C3	cent_19	P		FALSE	NA	FALSE	NA	NA
ENSMUSG000000089679	Gm16299		19	19 D2	cent_19	P		FALSE	NA	FALSE	NA	NA
ENSMUSG00000095180	Rhox5	X		X A3.3	cent_X	M		FALSE	NA	FALSE	NA	NA
ENSMUSG00000073125	Xlr3b	X		X A7	dist_X	M		FALSE	NA	FALSE	NA	NA
ENSMUSG00000067768	Xlr4b	X		X A7	dist_X	M		FALSE	NA	FALSE	NA	NA
ENSMUSG00000031362	Xlr4c	X		X A7	dist_X	M		FALSE	NA	FALSE	NA	NA
ENSMUSG00000085715	Tsix	X		X D	dist_X	P		FALSE	NA	FALSE	NA	NA
ENSMUSG00000086503	Xist	X		X D	dist_X	P		FALSE	NA	FALSE	NA	NA
ENSMUSG00000097571	Jpx	X		X D	dist_X	P		FALSE	NA	FALSE	NA	NA
ENSMUSG00000086370	Ftx	X		X D	dist_X	M		FALSE	NA	FALSE	NA	NA
ENSMUSG00000031330	Zcchc13	X		X D	dist_X	M		FALSE	NA	FALSE	NA	NA

Bibliography

- M. E. Blewitt, A. V. Gendrel, Z. Y. Pang, D. B. Sparrow, N. Whitelaw, J. M. Craig, A. Apedaile, D. J. Hilton, S. L. Dunwoodie, N. Brockdorff, G. F. Kay, and E. Whitelaw. Smchd1, containing a structural-maintenance-of-chromosomes hinge domain, has a critical role in x inactivation. *Nature Genetics*, 40(5):663–669, 2008. ISSN 1061-4036. doi: 10.1038/ng.142. URL [<GotoISI>://WOS:000255366700037](#).
- A. W. Mould, Z. Y. Pang, M. Pakusch, I. D. Tonks, M. Stark, D. Carrie, P. Mukhopadhyay, A. Seidel, J. J. Ellis, J. Deakin, M. J. Wakefield, L. Krause, M. E. Blewitt, and G. F. Kay. Smchd1 regulates a subset of autosomal genes subject to monoallelic expression in addition to being critical for x inactivation. *Epigenetics & Chromatin*, 6, 2013. ISSN 1756-8935. doi: Artn1910.1186/1756-8935-6-19. URL [<GotoISI>://WOS:000321520900001](#).
- A. V. Gendrel, Y. A. Tang, M. Suzuki, J. Godwin, T. B. Nesterova, J. M. Greally, E. Heard, and N. Brockdorff. Epigenetic functions of smchd1 repress gene clusters on the inactive x chromosome and on autosomes. *Mol Cell Biol*, 33(16): 3150–65, 2013. ISSN 1098-5549 (Electronic) 0270-7306 (Linking). doi: 10.1128/MCB.00145-13. URL <https://www.ncbi.nlm.nih.gov/pubmed/23754746>.
- R. J. Lemmers, R. Tawil, L. M. Petek, J. Balog, G. J. Block, G. W. Santen, A. M. Amell, P. J. van der Vliet, R. Almomani, K. R. Straasheijm, Y. D. Krom, R. Klooster, Y. Sun, J. T. den Dunnen, Q. Helmer, C. M. Donlin-Smith, G. W. Padberg, B. G. van Engelen, J. C. de Greef, A. M. Aartsma-Rus, R. R. Frants, M. de Visser, C. Desnuelle, S. Sacconi, G. N. Filippova, B. Bakker, M. J. Bamshad, S. J. Tapscott, D. G. Miller, and S. M. van der Maarel. Digenic

- inheritance of an *smchd1* mutation and an *fshd*-permissive *d4z4* allele causes facioscapulohumeral muscular dystrophy type 2. *Nat Genet*, 44(12):1370–4, 2012. ISSN 1546-1718 (Electronic) 1061-4036 (Linking). doi: 10.1038/ng.2454. URL <https://www.ncbi.nlm.nih.gov/pubmed/23143600>.
- C. T. Gordon, S. Xue, G. Yigit, H. Filali, K. Chen, N. Rosin, K. I. Yoshiura, M. Oufadem, T. J. Beck, R. McGowan, A. C. Magee, J. Altmuller, C. Dion, H. Thiele, A. D. Gurzau, P. Nurnberg, D. Meschede, W. Muhlbauer, N. Okamoto, V. Varghese, R. Irving, S. Sigaudy, D. Williams, S. F. Ahmed, C. Bonnard, M. K. Kong, I. Ratbi, N. Fejjal, M. Fikri, S. C. Elalaoui, H. Reigstad, C. Bole-Feysot, P. Nitschke, N. Ragge, N. Levy, G. Tuncbilek, A. S. Teo, M. L. Cunningham, A. Sefiani, H. Kayserili, J. M. Murphy, C. Chatdokmaiprai, A. M. Hillmer, D. Wattanasirichaigoon, S. Lyonnet, F. Magdinier, A. Javed, M. E. Blewitt, J. Amiel, B. Wollnik, and B. Reversade. De novo mutations in *smchd1* cause bosma arhinia microphthalmia syndrome and abrogate nasal development. *Nat Genet*, 49(2):249–255, 2017. ISSN 1546-1718 (Electronic) 1061-4036 (Linking). doi: 10.1038/ng.3765. URL <https://www.ncbi.nlm.nih.gov/pubmed/28067911>.
- N. D. Shaw, H. Brand, Z. A. Kupchinsky, H. Bengani, L. Plummer, T. I. Jones, S. Erdin, K. A. Williamson, J. Rainger, A. Stortchevoi, K. Samocha, B. B. Curren, D. S. Dunican, R. L. Collins, J. R. Willer, A. Lek, M. Lek, M. Nassan, S. Pereira, T. Kammin, D. Lucente, A. Silva, C. M. Seabra, C. Chiang, Y. An, M. Ansari, J. K. Rainger, S. Joss, J. C. Smith, M. F. Lippincott, S. S. Singh, N. Patel, J. W. Jing, J. R. Law, N. Ferraro, A. Verloes, A. Rauch, K. Steindl, M. Zweier, I. Scheer, D. Sato, N. Okamoto, C. Jacobsen, J. Tryggestad, S. Chernausk, L. A. Schimmenti, B. Brasseur, C. Cesaretti, J. E. Garcia-Ortiz, T. P. Buitrago, O. P. Silva, J. D. Hoffman, W. Muhlbauer, K. W. Ruprecht, B. L. Loeys, M. Shino, A. M. Kaandl, C. H. Cho, C. C. Morton, R. R. Meehan, V. van Heyningen, E. C. Liao, R. Balasubramanian, J. E. Hall, S. B. Seminara, D. Macarthur, S. A. Moore, K. I. Yoshiura, J. F. Gusella, J. A. Marsh, Jr. Graham, J. M., A. E. Lin, N. Katsanis, P. L. Jones, Jr. Crowley, W. F.,

- E. E. Davis, D. R. FitzPatrick, and M. E. Talkowski. Smchd1 mutations associated with a rare muscular dystrophy can also cause isolated arhinia and bosma arhinia microphthalmia syndrome. *Nat Genet*, 49(2):238–248, 2017. ISSN 1546-1718 (Electronic) 1061-4036 (Linking). doi: 10.1038/ng.3743. URL <https://www.ncbi.nlm.nih.gov/pubmed/28067909>.
- Natasha Jansz, Andrew Keniry, Marie Trussart, Heidi Bildsoe, Tamara Beck, Ian D. Tonks, Arne W. Mould, Peter Hickey, Kelsey Breslin, Megan Iminoff, Matthew E. Ritchie, Edwina McGlinn, Graham F. Kay, James M. Murphy, and Marnie E. Blewitt. Smchd1 regulates long-range chromatin interactions on the inactive x chromosome and at hox clusters. *Nature Structural & Molecular Biology*, 2018a. ISSN 1545-9993 1545-9985. doi: 10.1038/s41594-018-0111-z.
- K. Chen, J. Hu, D. L. Moore, R. Liu, S. A. Kessans, K. Breslin, I. S. Lucet, A. Keniry, H. S. Leong, C. L. Parish, D. J. Hilton, R. J. Lemmers, S. M. van der Maarel, P. E. Czabotar, R. C. Dobson, M. E. Ritchie, G. F. Kay, J. M. Murphy, and M. E. Blewitt. Genome-wide binding and mechanistic analyses of smchd1-mediated epigenetic regulation. *Proc Natl Acad Sci U S A*, 112(27):E3535–44, 2015. ISSN 1091-6490 (Electronic) 0027-8424 (Linking). doi: 10.1073/pnas.1504232112. URL <https://www.ncbi.nlm.nih.gov/pubmed/26091879>.
- C. H. Waddington. The epigenotype. 1942. *Int J Epidemiol*, 41(1):10–3, 1942. ISSN 1464-3685 (Electronic) 0300-5771 (Linking). doi: 10.1093/ije/dyr184. URL <https://www.ncbi.nlm.nih.gov/pubmed/22186258>.
- U. Deichmann. Epigenetics: The origins and evolution of a fashionable topic. *Dev Biol*, 416(1):249–254, 2016. ISSN 1095-564X (Electronic) 0012-1606 (Linking). doi: 10.1016/j.ydbio.2016.06.005. URL <https://www.ncbi.nlm.nih.gov/pubmed/27291929>.
- J. M. Greally. A user’s guide to the ambiguous word ‘epigenetics’. *Nat Rev Mol Cell Biol*, 19(4):207–208, 2018. ISSN 1471-0080 (Electronic) 1471-0072 (Linking). doi: 10.1038/nrm.2017.135. URL <https://www.ncbi.nlm.nih.gov/pubmed/29339796>.

- V. E. A. (Vincenzo E. A.) Russo, Arthur D Riggs, and Robert A Martienssen. *Epigenetic mechanisms of gene regulation*. Cold Spring Harbour Laboratory Press, Cold Spring Harbour, NY, 1996. ISBN 0879694904.
- A. D. Riggs. X chromosome inactivation, differentiation, and DNA methylation revisited, with a tribute to susumu ohno. *Cytogenet Genome Res*, 99(1-4):17–24, 2002. ISSN 1424-859X (Electronic) 1424-8581 (Linking). doi: 10.1159/000071569. URL <https://www.ncbi.nlm.nih.gov/pubmed/12900540>.
- A. V. Probst, E. Dunleavy, and G. Almouzni. Epigenetic inheritance during the cell cycle. *Nat Rev Mol Cell Biol*, 10(3):192–206, 2009. ISSN 1471-0080 (Electronic) 1471-0072 (Linking). doi: 10.1038/nrm2640. URL <https://www.ncbi.nlm.nih.gov/pubmed/19234478>.
- M. E. Oomen and J. Dekker. Epigenetic characteristics of the mitotic chromosome in 1d and 3d. *Crit Rev Biochem Mol Biol*, 52(2):185–204, 2017. ISSN 1549-7798 (Electronic) 1040-9238 (Linking). doi: 10.1080/10409238.2017.1287160. URL <https://www.ncbi.nlm.nih.gov/pubmed/28228067>.
- Bruce Alberts, Alexander Johnson, Julian Lewis, Martin Raff, Keith Roberts, and Peter Walter. *Chromosomal DNA and Its Packaging in the Chromatin Fiber*, book section 4, pages 173–276. Garland Science, New York, 2002.
- G. Felsenfeld and M. Groudine. Controlling the double helix. *Nature*, 421(6921):448–53, 2003. ISSN 0028-0836 (Print) 0028-0836 (Linking). doi: 10.1038/nature01411. URL <https://www.ncbi.nlm.nih.gov/pubmed/12540921>.
- A. L. Olins and D. E. Olins. Spheroid chromatin units (v bodies). *Science*, 183(4122):330–2, 1974. ISSN 0036-8075 (Print) 0036-8075 (Linking). doi: 10.1126/science.183.4122.330. URL <https://www.ncbi.nlm.nih.gov/pubmed/4128918>.
- Sepideh Khorasanizadeh. The nucleosome. *Cell*, 116(2):259–272, 2004. ISSN 00928674. doi: 10.1016/s0092-8674(04)00044-3.

- K. Luger, A. W. Mader, R. K. Richmond, D. F. Sargent, and T. J. Richmond. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature*, 389(6648):251–60, 1997. ISSN 0028-0836 (Print) 0028-0836 (Linking). doi: 10.1038/38444. URL <https://www.ncbi.nlm.nih.gov/pubmed/9305837>.
- A. J. Bannister and T. Kouzarides. Regulation of chromatin by histone modifications. *Cell Res*, 21(3):381–95, 2011. ISSN 1748-7838 (Electronic) 1001-0602 (Linking). doi: 10.1038/cr.2011.22. URL <https://www.ncbi.nlm.nih.gov/pubmed/21321607>.
- S. A. Jacobs and S. Khorasanizadeh. Structure of hp1 chromodomain bound to a lysine 9-methylated histone h3 tail. *Science*, 295(5562):2080–3, 2002. ISSN 0036-8075. doi: 10.1126/science.1069473.
- J. Nuebler, G. Fudenberg, M. Imakaev, N. Abdennur, and L. A. Mirny. Chromatin organization by an interplay of loop extrusion and compartmental segregation. *Proc Natl Acad Sci U S A*, 115(29):E6697–E6706, 2018. ISSN 1091-6490 (Electronic) 0027-8424 (Linking). doi: 10.1073/pnas.1717730115. URL <https://www.ncbi.nlm.nih.gov/pubmed/29967174>.
- S. L. Klemm, Z. Shipony, and W. J. Greenleaf. Chromatin accessibility and the regulatory epigenome. *Nat Rev Genet*, 20(4):207–220, 2019. ISSN 1471-0064 (Electronic) 1471-0056 (Linking). doi: 10.1038/s41576-018-0089-8. URL <https://www.ncbi.nlm.nih.gov/pubmed/30675018>.
- R. C. Allshire and H. D. Madhani. Ten principles of heterochromatin formation and function. *Nat Rev Mol Cell Biol*, 19(4):229–244, 2018. ISSN 1471-0080 (Electronic) 1471-0072 (Linking). doi: 10.1038/nrm.2017.119. URL <https://www.ncbi.nlm.nih.gov/pubmed/29235574>.
- C. D. Allis and T. Jenuwein. The molecular hallmarks of epigenetic control. *Nat Rev Genet*, 17(8):487–500, 2016. ISSN 1471-0064 (Electronic) 1471-0056 (Linking). doi: 10.1038/nrg.2016.59. URL <https://www.ncbi.nlm.nih.gov/pubmed/27346641>.

- R. Holliday and J. E. Pugh. DNA modification mechanisms and gene activity during development. *Science*, 187(4173):226–32, 1975. ISSN 0036-8075 (Print) 0036-8075 (Linking). doi: 10.1126/science.187.4173.226. URL <https://www.ncbi.nlm.nih.gov/pubmed/1111098>.
- A. D. Riggs. X inactivation, differentiation, and DNA methylation. *Cytogenetics and cell genetics*, 14(1):9–25, 1975. ISSN 0301-0171. doi: 10.1159/000130315. URL <http://europepmc.org/abstract/MED/1093816https://doi.org/10.1159/000130315>.
- R. D. Hotchkiss. The quantitative separation of purines, pyrimidines, and nucleosides by paper chromatography. *J Biol Chem*, 175(1):315–32, 1948. ISSN 0021-9258 (Print) 0021-9258.
- E. Li and Y. Zhang. DNA methylation in mammals. *Cold Spring Harb Perspect Biol*, 6(5):a019133, 2014. ISSN 1943-0264 (Electronic) 1943-0264 (Linking). doi: 10.1101/cshperspect.a019133. URL <https://www.ncbi.nlm.nih.gov/pubmed/24789823>.
- F. Lyko. The DNA methyltransferase family: a versatile toolkit for epigenetic regulation. *Nat Rev Genet*, 19(2):81–92, 2018. ISSN 1471-0064 (Electronic) 1471-0056 (Linking). doi: 10.1038/nrg.2017.80. URL <https://www.ncbi.nlm.nih.gov/pubmed/29033456>.
- En Li, Timothy H. Bestor, and Rudolf Jaenisch. Targeted mutation of the dna methyltransferase gene results in embryonic lethality. *Cell*, 69(6):915–926, 1992. ISSN 00928674. doi: 10.1016/0092-8674(92)90611-f.
- Masaki Okano, Daphne W. Bell, Daniel A. Haber, and En Li. Dna methyltransferases dnmt3a and dnmt3b are essential for de novo methylation and mammalian development. *Cell*, 99(3):247–257, 1999. ISSN 00928674. doi: 10.1016/s0092-8674(00)81656-6.
- R. S. Hansen, C. Wijmenga, P. Luo, A. M. Stanek, T. K. Canfield, C. M. Weemaes, and S. M. Gartler. The dnmt3b dna methyltransferase gene is mutated in the icf

- immunodeficiency syndrome. *Proc Natl Acad Sci U S A*, 96(25):14412–7, 1999. ISSN 0027-8424 (Print) 0027-8424 (Linking). doi: 10.1073/pnas.96.25.14412. URL <https://www.ncbi.nlm.nih.gov/pubmed/10588719>.
- M. Y. Shah and J. D. Licht. Dnmt3a mutations in acute myeloid leukemia. *Nat Genet*, 43(4):289–90, 2011. ISSN 1546-1718 (Electronic) 1061-4036 (Linking). doi: 10.1038/ng0411-289. URL <https://www.ncbi.nlm.nih.gov/pubmed/21445072>.
- J. Winkelmann, L. Lin, B. Schormair, B. R. Kornum, J. Faraco, G. Plazzi, A. Melberg, F. Cornelio, A. E. Urban, F. Pizza, F. Poli, F. Grubert, T. Wieland, E. Graf, J. Hallmayer, T. M. Strom, and E. Mignot. Mutations in dnmt1 cause autosomal dominant cerebellar ataxia, deafness and narcolepsy. *Hum Mol Genet*, 21(10):2205–10, 2012. ISSN 1460-2083 (Electronic) 0964-6906 (Linking). doi: 10.1093/hmg/dds035. URL <https://www.ncbi.nlm.nih.gov/pubmed/22328086>.
- C. J. Klein, M. V. Botuyan, Y. Wu, C. J. Ward, G. A. Nicholson, S. Hammans, K. Hojo, H. Yamanishi, A. R. Karpf, D. C. Wallace, M. Simon, C. Lander, L. A. Boardman, J. M. Cunningham, G. E. Smith, W. J. Litchy, B. Boes, E. J. Atkinson, S. Middha, B. Dyck PJ, J. E. Parisi, G. Mer, D. I. Smith, and P. J. Dyck. Mutations in dnmt1 cause hereditary sensory neuropathy with dementia and hearing loss. *Nat Genet*, 43(6):595–600, 2011. ISSN 1546-1718 (Electronic) 1061-4036 (Linking). doi: 10.1038/ng.830. URL <https://www.ncbi.nlm.nih.gov/pubmed/21532572>.
- Catherine McKeon, Hiroaki Ohkubo, Ira Pastan, and Benoit de Crombrughe. Unusual methylation pattern of the $\alpha 2(i)$ collagen gene. *Cell*, 29(1):203–210, 1982. ISSN 00928674. doi: 10.1016/0092-8674(82)90104-0.
- Adrian Bird, Mary Taggart, Marianne Frommer, Orlando J. Miller, and Donald Macleod. A fraction of the mouse genome that is derived from islands of non-methylated, CpG-rich DNA. *Cell*, 40(1):91–99, 1985. ISSN 00928674. doi: 10.1016/0092-8674(85)90312-5.

- J. M. Craig and W. A. Bickmore. The distribution of CpG islands in mammalian chromosomes. *Nat Genet*, 7(3):376–82, 1994. ISSN 1061-4036 (Print) 1061-4036 (Linking). doi: 10.1038/ng0794-376. URL <https://www.ncbi.nlm.nih.gov/pubmed/7920655>.
- S. F. Wolf, D. J. Jolly, K. D. Lunnen, T. Friedmann, and B. R. Migeon. Methylation of the hypoxanthine phosphoribosyltransferase locus on the human x chromosome: implications for x-chromosome inactivation. *Proc Natl Acad Sci U S A*, 81(9):2806–10, 1984. ISSN 0027-8424 (Print) 0027-8424 (Linking). doi: 10.1073/pnas.81.9.2806. URL <https://www.ncbi.nlm.nih.gov/pubmed/6585829>.
- A. C. Ferguson-Smith, H. Sasaki, B. M. Cattanch, and M. A. Surani. Parental-origin-specific epigenetic modification of the mouse h19 gene. *Nature*, 362(6422):751–5, 1993. ISSN 0028-0836 (Print) 0028-0836 (Linking). doi: 10.1038/362751a0. URL <https://www.ncbi.nlm.nih.gov/pubmed/8469285>.
- D. Macleod, J. Charlton, J. Mullins, and A. P. Bird. Sp1 sites in the mouse aprt gene promoter are required to prevent methylation of the cpg island. *Genes Dev*, 8(19):2282–92, 1994. ISSN 0890-9369 (Print) 0890-9369 (Linking). doi: 10.1101/gad.8.19.2282. URL <https://www.ncbi.nlm.nih.gov/pubmed/7958895>.
- Y. Yin, E. Morgunova, A. Jolma, E. Kaasinen, B. Sahu, S. Khund-Sayeed, P. K. Das, T. Kivioja, K. Dave, F. Zhong, K. R. Nitta, M. Taipale, A. Popov, P. A. Ginno, S. Domcke, J. Yan, D. Schubeler, C. Vinson, and J. Taipale. Impact of cytosine methylation on dna binding specificities of human transcription factors. *Science*, 356(6337), 2017. ISSN 1095-9203 (Electronic) 0036-8075 (Linking). doi: 10.1126/science.aaj2239. URL <https://www.ncbi.nlm.nih.gov/pubmed/28473536>.
- A. M. Lindroth, X. Cao, J. P. Jackson, D. Zilberman, C. M. McCallum, S. Henikoff, and S. E. Jacobsen. Requirement of CHROMOMETHYLASE3 for maintenance of CpXpG methylation. *Science*, 292(5524):2077–80, 2001. ISSN 0036-8075 (Print) 0036-8075 (Linking). doi: 10.1126/science.1059745. URL <https://www.ncbi.nlm.nih.gov/pubmed/11349138>.

- R. Lister, M. Pelizzola, R. H. Downen, R. D. Hawkins, G. Hon, J. Tonti-Filippini, J. R. Nery, L. Lee, Z. Ye, Q. M. Ngo, L. Edsall, J. Antosiewicz-Bourget, R. Stewart, V. Ruotti, A. H. Millar, J. A. Thomson, B. Ren, and J. R. Ecker. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature*, 462(7271):315–22, 2009. ISSN 1476-4687 (Electronic) 0028-0836 (Linking). doi: 10.1038/nature08514. URL <https://www.ncbi.nlm.nih.gov/pubmed/19829295>.
- H. S. Jang, W. J. Shin, J. E. Lee, and J. T. Do. CpG and non-cpG methylation in epigenetic gene regulation and brain function. *Genes (Basel)*, 8(6), 2017. ISSN 2073-4425 (Print) 2073-4425 (Linking). doi: 10.3390/genes8060148. URL <https://www.ncbi.nlm.nih.gov/pubmed/28545252>.
- M. Tahiliani, K. P. Koh, Y. Shen, W. A. Pastor, H. Bandukwala, Y. Brudno, S. Agarwal, L. M. Iyer, D. R. Liu, L. Aravind, and A. Rao. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian dna by mll partner tet1. *Science*, 324(5929):930–5, 2009. ISSN 1095-9203 (Electronic) 0036-8075 (Linking). doi: 10.1126/science.1170116. URL <https://www.ncbi.nlm.nih.gov/pubmed/19372391>.
- S. Ito, A. C. D’Alessio, O. V. Taranova, K. Hong, L. C. Sowers, and Y. Zhang. Role of tet proteins in 5mc to 5hmc conversion, es-cell self-renewal and inner cell mass specification. *Nature*, 466(7310):1129–33, 2010. ISSN 1476-4687 (Electronic) 0028-0836 (Linking). doi: 10.1038/nature09303. URL <https://www.ncbi.nlm.nih.gov/pubmed/20639862>.
- S. C. Wu and Y. Zhang. Active dna demethylation: many roads lead to rome. *Nat Rev Mol Cell Biol*, 11(9):607–20, 2010. ISSN 1471-0080 (Electronic) 1471-0072 (Linking). doi: 10.1038/nrm2950. URL <https://www.ncbi.nlm.nih.gov/pubmed/20683471>.
- Carina Y. Howell, Timothy H. Bestor, Feng Ding, Keith E. Latham, Carmen Mertineit, Jacquetta M. Trasler, and J. Richard Chaillet. Genomic imprinting disrupted by a maternal effect mutation in the dnmt1 gene. *Cell*, 104(6):829–838, 2001. ISSN 00928674. doi: 10.1016/s0092-8674(01)00280-x.

- P. Tessarz and T. Kouzarides. Histone core modifications regulating nucleosome structure and dynamics. *Nat Rev Mol Cell Biol*, 15(11):703–8, 2014. ISSN 1471-0080 (Electronic) 1471-0072 (Linking). doi: 10.1038/nrm3890. URL <https://www.ncbi.nlm.nih.gov/pubmed/25315270>.
- P. Tropberger and R. Schneider. Scratching the (lateral) surface of chromatin regulation by histone modifications. *Nat Struct Mol Biol*, 20(6):657–61, 2013. ISSN 1545-9985 (Electronic) 1545-9985 (Linking). doi: 10.1038/nsmb.2581. URL <https://www.ncbi.nlm.nih.gov/pubmed/23739170>.
- T. Kouzarides. Chromatin modifications and their function. *Cell*, 128(4):693–705, 2007. ISSN 0092-8674 (Print) 0092-8674 (Linking). doi: 10.1016/j.cell.2007.02.005. URL <https://www.ncbi.nlm.nih.gov/pubmed/17320507>.
- A. F. Kebede, R. Schneider, and S. Daujat. Novel types and sites of histone modifications emerge as players in the transcriptional regulation contest. *FEBS J*, 282(9):1658–74, 2015. ISSN 1742-4658 (Electronic) 1742-464X (Linking). doi: 10.1111/febs.13047. URL <https://www.ncbi.nlm.nih.gov/pubmed/25220185>.
- M. Lawrence, S. Daujat, and R. Schneider. Lateral thinking: How histone modifications regulate gene expression. *Trends Genet*, 32(1):42–56, 2016. ISSN 0168-9525 (Print) 0168-9525 (Linking). doi: 10.1016/j.tig.2015.10.007. URL <https://www.ncbi.nlm.nih.gov/pubmed/26704082>.
- V. G. Allfrey, R. Faulkner, and A. E. Mirsky. Acetylation and methylation of histones and their possible role in the regulation of rna synthesis. *Proc Natl Acad Sci U S A*, 51:786–94, 1964. ISSN 0027-8424 (Print) 0027-8424 (Linking). doi: 10.1073/pnas.51.5.786. URL <https://www.ncbi.nlm.nih.gov/pubmed/14172992>.
- M. Grunstein. Histone acetylation in chromatin structure and transcription. *Nature*, 389(6649):349–52, 1997. ISSN 0028-0836 (Print) 0028-0836 (Linking). doi: 10.1038/38664. URL <https://www.ncbi.nlm.nih.gov/pubmed/9311776>.

- E. L. Greer and Y. Shi. Histone methylation: a dynamic mark in health, disease and inheritance. *Nat Rev Genet*, 13(5):343–57, 2012. ISSN 1471-0064 (Electronic) 1471-0056 (Linking). doi: 10.1038/nrg3173. URL <https://www.ncbi.nlm.nih.gov/pubmed/22473383>.
- Vincenzo Pirrotta. Polycomb the genome: Pcg, trxg, and chromatin silencing. *Cell*, 93(3):333–336, 1998. ISSN 00928674. doi: 10.1016/s0092-8674(00)81162-9.
- Andrés Gaytán de Ayala Alonso, Luis Gutiérrez, Cornelia Fritsch, Bernadett Papp, Dirk Beuchle, and Jürg Müller. A genetic screen identifies novel polycomb group genes in drosophila. *Genetics*, 176(4):2099–2108, 2007. ISSN 0016-6731 1943-2631. doi: 10.1534/genetics.107.075739.
- H. Wang, L. Wang, H. Erdjument-Bromage, M. Vidal, P. Tempst, R. S. Jones, and Y. Zhang. Role of histone h2a ubiquitination in polycomb silencing. *Nature*, 431(7010):873–8, 2004. ISSN 1476-4687 (Electronic) 0028-0836 (Linking). doi: 10.1038/nature02985. URL <https://www.ncbi.nlm.nih.gov/pubmed/15386022>.
- Sergi Aranda, Gloria Mas, and Luciano Di Croce. Regulation of gene transcription by polycomb proteins. *Science Advances*, 1(11), 2015. ISSN 2375-2548. doi: 10.1126/sciadv.1500737.
- R. Cao, L. Wang, H. Wang, L. Xia, H. Erdjument-Bromage, P. Tempst, R. S. Jones, and Y. Zhang. Role of histone h3 lysine 27 methylation in polycomb-group silencing. *Science*, 298(5595):1039–43, 2002. ISSN 1095-9203 (Electronic) 0036-8075 (Linking). doi: 10.1126/science.1076997. URL <https://www.ncbi.nlm.nih.gov/pubmed/12351676>.
- A. Kuzmichev, K. Nishioka, H. Erdjument-Bromage, P. Tempst, and D. Reinberg. Histone methyltransferase activity associated with a human multiprotein complex containing the enhancer of zeste protein. *Genes Dev*, 16(22):2893–905, 2002. ISSN 0890-9369 (Print) 0890-9369 (Linking). doi: 10.1101/gad.1035902. URL <https://www.ncbi.nlm.nih.gov/pubmed/12435631>.
- Y. B. Schwartz and V. Pirrotta. A new world of polycombs: unexpected partnerships and emerging functions. *Nat Rev Genet*, 14(12):853–64, 2013. ISSN

- 1471-0064 (Electronic) 1471-0056 (Linking). doi: 10.1038/nrg3603. URL <https://www.ncbi.nlm.nih.gov/pubmed/24217316>.
- W. Zhou, P. Zhu, J. Wang, G. Pascual, K. A. Ohgi, J. Lozach, C. K. Glass, and M. G. Rosenfeld. Histone h2a monoubiquitination represses transcription by inhibiting rna polymerase ii transcriptional elongation. *Mol Cell*, 29(1):69–80, 2008. ISSN 1097-2765 (Print) 1097-2765 (Linking). doi: 10.1016/j.molcel.2007.11.002. URL <https://www.ncbi.nlm.nih.gov/pubmed/18206970>.
- D. J. Grau, B. A. Chapman, J. D. Garlick, M. Borowsky, N. J. Francis, and R. E. Kingston. Compaction of chromatin by diverse polycomb group proteins requires localized regions of high charge. *Genes Dev*, 25(20):2210–21, 2011. ISSN 1549-5477 (Electronic) 0890-9369 (Linking). doi: 10.1101/gad.172882.11. URL <https://www.ncbi.nlm.nih.gov/pubmed/22012622>.
- R. Eskeland, M. Leeb, G. R. Grimes, C. Kress, S. Boyle, D. Sproul, N. Gilbert, Y. Fan, A. I. Skoultschi, A. Wutz, and W. A. Bickmore. Ring1b compacts chromatin structure and represses gene expression independent of histone ubiquitination. *Mol Cell*, 38(3):452–64, 2010. ISSN 1097-4164 (Electronic) 1097-2765 (Linking). doi: 10.1016/j.molcel.2010.02.032. URL <https://www.ncbi.nlm.nih.gov/pubmed/20471950>.
- S. Kundu, F. Ji, H. Sunwoo, G. Jain, J. T. Lee, R. I. Sadreyev, J. Dekker, and R. E. Kingston. Polycomb repressive complex 1 generates discrete compacted domains that change during differentiation. *Mol Cell*, 65(3):432–446 e5, 2017. ISSN 1097-4164 (Electronic) 1097-2765 (Linking). doi: 10.1016/j.molcel.2017.01.009. URL <https://www.ncbi.nlm.nih.gov/pubmed/28157505>.
- S. Boyle, I. M. Flyamer, I. Williamson, D. Sengupta, W. A. Bickmore, and R. S. Illingworth. A central role for canonical prc1 in shaping the 3d nuclear landscape. *Genes Dev*, 2020. ISSN 1549-5477 (Electronic) 0890-9369 (Linking). doi: 10.1101/gad.336487.120. URL <https://www.ncbi.nlm.nih.gov/pubmed/32439634>.

- N. P. Blackledge, A. M. Farcas, T. Kondo, H. W. King, J. F. McGouran, L. L. Hanssen, S. Ito, S. Cooper, K. Kondo, Y. Koseki, T. Ishikura, H. K. Long, T. W. Sheahan, N. Brockdorff, B. M. Kessler, H. Koseki, and R. J. Klose. Variant prc1 complex-dependent h2a ubiquitylation drives prc2 recruitment and polycomb domain formation. *Cell*, 157(6):1445–59, 2014. ISSN 1097-4172 (Electronic) 0092-8674 (Linking). doi: 10.1016/j.cell.2014.05.004. URL <https://www.ncbi.nlm.nih.gov/pubmed/24856970>.
- N. Jansz, T. Nesterova, A. Keniry, M. Iminoff, P. F. Hickey, G. Pintacuda, O. Masui, S. Kobelke, N. Geoghegan, K. A. Breslin, T. A. Willson, K. Rogers, G. F. Kay, A. H. Fox, H. Koseki, N. Brockdorff, J. M. Murphy, and M. E. Blewitt. Smchd1 targeting to the inactive x is dependent on the xist-hnrnpk-prc1 pathway. *Cell Rep*, 25(7):1912–1923 e9, 2018b. ISSN 2211-1247 (Electronic). doi: 10.1016/j.celrep.2018.10.044. URL <https://www.ncbi.nlm.nih.gov/pubmed/30428357>.
- R. R. Pandey, T. Mondal, F. Mohammad, S. Enroth, L. Redrup, J. Komorowski, T. Nagano, D. Mancini-Dinardo, and C. Kanduri. Kcnq1ot1 antisense noncoding rna mediates lineage-specific transcriptional silencing through chromatin-level regulation. *Mol Cell*, 32(2):232–46, 2008. ISSN 1097-4164 (Electronic) 1097-2765 (Linking). doi: 10.1016/j.molcel.2008.08.022. URL <https://www.ncbi.nlm.nih.gov/pubmed/18951091>.
- P. B. Becker and J. L. Workman. Nucleosome remodeling and epigenetics. *Cold Spring Harb Perspect Biol*, 5(9), 2013. ISSN 1943-0264 (Electronic) 1943-0264 (Linking). doi: 10.1101/cshperspect.a017905. URL <https://www.ncbi.nlm.nih.gov/pubmed/24003213>.
- I. Barozzi, M. Simonatto, S. Bonifacio, L. Yang, R. Rohs, S. Ghisletti, and G. Natoli. Coregulation of transcription factor binding and nucleosome occupancy through dna features of mammalian enhancers. *Mol Cell*, 54(5):844–857, 2014. ISSN 1097-4164 (Electronic) 1097-2765 (Linking). doi: 10.1016/j.molcel.2014.04.006. URL <https://www.ncbi.nlm.nih.gov/pubmed/24813947>.

- Ahmed H. Hassan, Philippe Prochasson, Kristen E. Neely, Scott C. Galasinski, Mark Chandy, Michael J. Carrozza, and Jerry L. Workman. Function and selectivity of bromodomains in anchoring chromatin-modifying complexes to promoter nucleosomes. *Cell*, 111(3):369–379, 2002. ISSN 00928674. doi: 10.1016/s0092-8674(02)01005-x.
- S. A. Denslow and P. A. Wade. The human mi-2/nurd complex and gene regulation. *Oncogene*, 26(37):5433–8, 2007. ISSN 0950-9232 (Print) 0950-9232 (Linking). doi: 10.1038/sj.onc.1210611. URL <https://www.ncbi.nlm.nih.gov/pubmed/17694084>.
- C. P. Ponting, P. L. Oliver, and W. Reik. Evolution and functions of long non-coding rnas. *Cell*, 136(4):629–41, 2009. ISSN 1097-4172 (Electronic) 0092-8674 (Linking). doi: 10.1016/j.cell.2009.02.006. URL <https://www.ncbi.nlm.nih.gov/pubmed/19239885>.
- R. W. Carthew and E. J. Sontheimer. Origins and mechanisms of mirnas and sirnas. *Cell*, 136(4):642–55, 2009. ISSN 1097-4172 (Electronic) 0092-8674 (Linking). doi: 10.1016/j.cell.2009.01.035. URL <https://www.ncbi.nlm.nih.gov/pubmed/19239886>.
- C. D. Malone and G. J. Hannon. Small rnas as guardians of the genome. *Cell*, 136(4):656–68, 2009. ISSN 1097-4172 (Electronic) 0092-8674 (Linking). doi: 10.1016/j.cell.2009.01.045. URL <https://www.ncbi.nlm.nih.gov/pubmed/19239887>.
- Y. Okazaki, M. Furuno, T. Kasukawa, J. Adachi, H. Bono, S. Kondo, I. Nikaido, N. Osato, R. Saito, H. Suzuki, I. Yamanaka, H. Kiyosawa, K. Yagi, Y. Tomaru, Y. Hasegawa, A. Nogami, C. Schonbach, T. Gojobori, R. Baldarelli, D. P. Hill, C. Bult, D. A. Hume, J. Quackenbush, L. M. Schriml, A. Kanapin, H. Matsuda, S. Batalov, K. W. Beisel, J. A. Blake, D. Bradt, V. Brusic, C. Chothia, L. E. Corbani, S. Cousins, E. Dalla, T. A. Dragani, C. F. Fletcher, A. Forrest, K. S. Frazer, T. Gaasterland, M. Gariboldi, C. Gissi, A. Godzik, J. Gough, S. Grimmond, S. Gustincich, N. Hirokawa, I. J. Jackson, E. D. Jarvis, A. Kanai,

- H. Kawaji, Y. Kawasaki, R. M. Kedzierski, B. L. King, A. Konagaya, I. V. Kurochkin, Y. Lee, B. Lenhard, P. A. Lyons, D. R. Maglott, L. Maltais, L. Marchionni, L. McKenzie, H. Miki, T. Nagashima, K. Numata, T. Okido, W. J. Pavan, G. Pertea, G. Pesole, N. Petrovsky, R. Pillai, J. U. Pontius, D. Qi, S. Ramachandran, T. Ravasi, J. C. Reed, D. J. Reed, J. Reid, B. Z. Ring, M. Ringwald, A. Sandelin, C. Schneider, C. A. Semple, M. Setou, K. Shimada, R. Sultana, Y. Takenaka, M. S. Taylor, R. D. Teasdale, M. Tomita, R. Verardo, L. Wagner, C. Wahlestedt, Y. Wang, Y. Watanabe, C. Wells, L. G. Wilming, A. Wynshaw-Boris, M. Yanagisawa, et al. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cdnas. *Nature*, 420 (6915):563–73, 2002. ISSN 0028-0836 (Print) 0028-0836 (Linking). doi: 10.1038/nature01266. URL <https://www.ncbi.nlm.nih.gov/pubmed/12466851>.
- B. Uszczynska-Ratajczak, J. Lagarde, A. Frankish, R. Guigo, and R. Johnson. Towards a complete map of the human long non-coding rna transcriptome. *Nat Rev Genet*, 19(9):535–548, 2018. ISSN 1471-0064 (Electronic) 1471-0056 (Linking). doi: 10.1038/s41576-018-0017-y. URL <https://www.ncbi.nlm.nih.gov/pubmed/29795125>.
- F. P. Marchese, I. Raimondi, and M. Huarte. The multidimensional mechanisms of long noncoding rna function. *Genome Biol*, 18(1):206, 2017. ISSN 1474-760X (Electronic) 1474-7596 (Linking). doi: 10.1186/s13059-017-1348-2. URL <https://www.ncbi.nlm.nih.gov/pubmed/29084573>.
- M. J. Delas and G. J. Hannon. Incrnas in development and disease: from functions to mechanisms. *Open Biol*, 7(7), 2017. ISSN 2046-2441 (Electronic) 2046-2441 (Linking). doi: 10.1098/rsob.170121. URL <https://www.ncbi.nlm.nih.gov/pubmed/28747406>.
- G. Pintacuda, G. Wei, C. Roustan, B. A. Kirmizitas, N. Solcan, A. Cerase, A. Castello, S. Mohammed, B. Moindrot, T. B. Nesterova, and N. Brockdorff. hnrnpk recruits pcgf3/5-prc1 to the xist rna b-repeat to establish polycomb-mediated chromosomal silencing. *Mol Cell*, 68(5):955–969 e10, 2017. ISSN

- 1097-4164 (Electronic) 1097-2765 (Linking). doi: 10.1016/j.molcel.2017.11.013. URL <https://www.ncbi.nlm.nih.gov/pubmed/29220657>.
- A. Bousard, A. C. Raposo, J. J. Zylicz, C. Picard, V. B. Pires, Y. Qi, C. Gil, L. Syx, H. Y. Chang, E. Heard, and S. T. da Rocha. The role of xist-mediated polycomb recruitment in the initiation of x-chromosome inactivation. *EMBO Rep*, page e48019, 2019. ISSN 1469-3178 (Electronic) 1469-221X (Linking). doi: 10.15252/embr.201948019. URL <https://www.ncbi.nlm.nih.gov/pubmed/31456285>.
- C. W. Hanna, R. Perez-Palacios, L. Gahurova, M. Schubert, F. Krueger, L. Biggins, S. Andrews, M. Colome-Tatche, D. Bourc’his, W. Dean, and G. Kelsey. Endogenous retroviral insertions drive non-canonical imprinting in extra-embryonic tissues. *Genome Biol*, 20(1):225, 2019. ISSN 1474-760X (Electronic) 1474-7596 (Linking). doi: 10.1186/s13059-019-1833-x. URL <https://www.ncbi.nlm.nih.gov/pubmed/31665063>.
- T. R. Mercer, M. E. Dinger, and J. S. Mattick. Long non-coding rnas: insights into functions. *Nat Rev Genet*, 10(3):155–9, 2009. ISSN 1471-0064 (Electronic) 1471-0056 (Linking). doi: 10.1038/nrg2521. URL <https://www.ncbi.nlm.nih.gov/pubmed/19188922>.
- R. W. Yao, Y. Wang, and L. L. Chen. Cellular functions of long noncoding rnas. *Nat Cell Biol*, 21(5):542–551, 2019. ISSN 1476-4679 (Electronic) 1465-7392 (Linking). doi: 10.1038/s41556-019-0311-8. URL <https://www.ncbi.nlm.nih.gov/pubmed/31048766>.
- E. Lieberman-Aiden, N. L. van Berkum, L. Williams, M. Imakaev, T. Ragoczy, A. Telling, I. Amit, B. R. Lajoie, P. J. Sabo, M. O. Dorschner, R. Sandstrom, B. Bernstein, M. A. Bender, M. Groudine, A. Gnirke, J. Stamatoyannopoulos, L. A. Mirny, E. S. Lander, and J. Dekker. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*, 326(5950):289–93, 2009. ISSN 1095-9203 (Electronic) 0036-8075 (Linking). doi: 10.1126/science.1181369. URL <https://www.ncbi.nlm.nih.gov/pubmed/19815776>.

- J. R. Hughes, N. Roberts, S. McGowan, D. Hay, E. Giannoulatou, M. Lynch, M. De Gobbi, S. Taylor, R. Gibbons, and D. R. Higgs. Analysis of hundreds of cis-regulatory landscapes at high resolution in a single, high-throughput experiment. *Nat Genet*, 46(2):205–12, 2014. ISSN 1546-1718 (Electronic) 1061-4036 (Linking). doi: 10.1038/ng.2871. URL <https://www.ncbi.nlm.nih.gov/pubmed/24413732>.
- P. J. Skene and S. Henikoff. An efficient targeted nuclease strategy for high-resolution mapping of dna binding sites. *Elife*, 6, 2017. ISSN 2050-084X (Electronic) 2050-084X (Linking). doi: 10.7554/eLife.21856. URL <https://www.ncbi.nlm.nih.gov/pubmed/28079019>.
- R. Kempfer and A. Pombo. Methods for mapping 3d chromosome architecture. *Nat Rev Genet*, 21(4):207–226, 2020. ISSN 1471-0064 (Electronic) 1471-0056 (Linking). doi: 10.1038/s41576-019-0195-2. URL <https://www.ncbi.nlm.nih.gov/pubmed/31848476>.
- M. J. Rust, M. Bates, and X. Zhuang. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (storm). *Nat Methods*, 3(10):793–5, 2006. ISSN 1548-7091 (Print) 1548-7091 (Linking). doi: 10.1038/nmeth929. URL <https://www.ncbi.nlm.nih.gov/pubmed/16896339>.
- B. Bonev and G. Cavalli. Organization and function of the 3d genome. *Nat Rev Genet*, 17(11):661–678, 2016. ISSN 1471-0064 (Electronic) 1471-0056 (Linking). doi: 10.1038/nrg.2016.112. URL <https://www.ncbi.nlm.nih.gov/pubmed/27739532>.
- Stefan Schoenfelder and Peter Fraser. Long-range enhancer–promoter contacts in gene expression control. *Nature Reviews Genetics*, 20(8):437–455, 2019. ISSN 1471-0056 1471-0064. doi: 10.1038/s41576-019-0128-0.
- A. S. Weintraub, C. H. Li, A. V. Zamudio, A. A. Sigova, N. M. Hannett, D. S. Day, B. J. Abraham, M. A. Cohen, B. Nabet, D. L. Buckley, Y. E. Guo, D. Hnisz,

- R. Jaenisch, J. E. Bradner, N. S. Gray, and R. A. Young. Yy1 is a structural regulator of enhancer-promoter loops. *Cell*, 171(7):1573–1588 e28, 2017. ISSN 1097-4172 (Electronic) 0092-8674 (Linking). doi: 10.1016/j.cell.2017.11.008. URL <https://www.ncbi.nlm.nih.gov/pubmed/29224777>.
- S. S. P. Rao, S. C. Huang, B. Glenn St Hilaire, J. M. Engreitz, E. M. Perez, K. R. Kieffer-Kwon, A. L. Sanborn, S. E. Johnstone, G. D. Bascom, I. D. Bochkov, X. Huang, M. S. Shamim, J. Shin, D. Turner, Z. Ye, A. D. Omer, J. T. Robinson, T. Schlick, B. E. Bernstein, R. Casellas, E. S. Lander, and E. L. Aiden. Cohesin loss eliminates all loop domains. *Cell*, 171(2):305–320 e24, 2017. ISSN 1097-4172 (Electronic) 0092-8674 (Linking). doi: 10.1016/j.cell.2017.09.026. URL <https://www.ncbi.nlm.nih.gov/pubmed/28985562>.
- I. F. Davidson, B. Bauer, D. Goetz, W. Tang, G. Wutz, and J. M. Peters. Dna loop extrusion by human cohesin. *Science*, 366(6471):1338–1345, 2019. ISSN 1095-9203 (Electronic) 0036-8075 (Linking). doi: 10.1126/science.aaz3418. URL <https://www.ncbi.nlm.nih.gov/pubmed/31753851>.
- A. Murrell, S. Heeson, and W. Reik. Interaction between differentially methylated regions partitions the imprinted genes *igf2* and *h19* into parent-specific chromatin loops. *Nat Genet*, 36(8):889–93, 2004. ISSN 1061-4036 (Print) 1061-4036 (Linking). doi: 10.1038/ng1402. URL <https://www.ncbi.nlm.nih.gov/pubmed/15273689>.
- J. R. Dixon, S. Selvaraj, F. Yue, A. Kim, Y. Li, Y. Shen, M. Hu, J. S. Liu, and B. Ren. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*, 485(7398):376–80, 2012. ISSN 1476-4687 (Electronic) 0028-0836 (Linking). doi: 10.1038/nature11082. URL <https://www.ncbi.nlm.nih.gov/pubmed/22495300>.
- M. Jordan Rowley and Victor G. Corces. Organizational principles of 3d genome architecture. *Nature Reviews Genetics*, 19(12):789–800, 2018. ISSN 1471-0056 1471-0064. doi: 10.1038/s41576-018-0060-8.

- Q. Szabo, F. Bantignies, and G. Cavalli. Principles of genome folding into topologically associating domains. *Sci Adv*, 5(4):eaaw1668, 2019. ISSN 2375-2548 (Electronic) 2375-2548 (Linking). doi: 10.1126/sciadv.aaw1668. URL <https://www.ncbi.nlm.nih.gov/pubmed/30989119>.
- D. Noordermeer, M. Leleu, P. Schorderet, E. Joye, F. Chabaud, and D. Duboule. Temporal dynamics and developmental memory of 3d chromatin architecture at hox gene loci. *Elife*, 3:e02557, 2014. ISSN 2050-084X (Electronic) 2050-084X (Linking). doi: 10.7554/eLife.02557. URL <https://www.ncbi.nlm.nih.gov/pubmed/24843030>.
- E. Rodriguez-Carballo, L. Lopez-Delisle, N. Yakushiji-Kaminatsui, A. Ullate-Agote, and D. Duboule. Impact of genome architecture on the functional activation and repression of hox regulatory landscapes. *BMC Biol*, 17(1):55, 2019. ISSN 1741-7007 (Electronic) 1741-7007 (Linking). doi: 10.1186/s12915-019-0677-x. URL <https://www.ncbi.nlm.nih.gov/pubmed/31299961>.
- M. Franke, D. M. Ibrahim, G. Andrey, W. Schwarzer, V. Heinrich, R. Schopflin, K. Kraft, R. Kempfer, I. Jerkovic, W. L. Chan, M. Spielmann, B. Timmermann, L. Wittler, I. Kurth, P. Cambiaso, O. Zuffardi, G. Houge, L. Lambie, F. Brancati, A. Pombo, M. Vingron, F. Spitz, and S. Mundlos. Formation of new chromatin domains determines pathogenicity of genomic duplications. *Nature*, 538(7624):265–269, 2016. ISSN 1476-4687 (Electronic) 0028-0836 (Linking). doi: 10.1038/nature19800. URL <https://www.ncbi.nlm.nih.gov/pubmed/27706140>.
- D. Hnisz, A. S. Weintraub, D. S. Day, A. L. Valton, R. O. Bak, C. H. Li, J. Goldmann, B. R. Lajoie, Z. P. Fan, A. A. Sigova, J. Reddy, D. Borges-Rivera, T. I. Lee, R. Jaenisch, M. H. Porteus, J. Dekker, and R. A. Young. Activation of proto-oncogenes by disruption of chromosome neighborhoods. *Science*, 351(6280):1454–1458, 2016. ISSN 1095-9203 (Electronic) 0036-8075 (Linking). doi: 10.1126/science.aad9024. URL <https://www.ncbi.nlm.nih.gov/pubmed/26940867>.

- B. van Steensel and A. S. Belmont. Lamina-associated domains: Links with chromosome architecture, heterochromatin, and gene repression. *Cell*, 169(5):780–791, 2017. ISSN 1097-4172 (Electronic) 0092-8674 (Linking). doi: 10.1016/j.cell.2017.04.022. URL <https://www.ncbi.nlm.nih.gov/pubmed/28525751>.
- D. Peric-Hupkes, W. Meuleman, L. Pagie, S. W. Bruggeman, I. Solovei, W. Bruggeman, S. Graf, P. Flicek, R. M. Kerkhoven, M. van Lohuizen, M. Reinders, L. Wessels, and B. van Steensel. Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation. *Mol Cell*, 38(4):603–13, 2010. ISSN 1097-4164 (Electronic) 1097-2765 (Linking). doi: 10.1016/j.molcel.2010.03.016. URL <https://www.ncbi.nlm.nih.gov/pubmed/20513434>.
- I. Solovei, A. S. Wang, K. Thanisch, C. S. Schmidt, S. Krebs, M. Zwerger, T. V. Cohen, D. Devys, R. Foisner, L. Peichl, H. Herrmann, H. Blum, D. Engelkamp, C. L. Stewart, H. Leonhardt, and B. Joffe. Lbr and lamin a/c sequentially tether peripheral heterochromatin and inversely regulate differentiation. *Cell*, 152(3):584–98, 2013. ISSN 1097-4172 (Electronic) 0092-8674 (Linking). doi: 10.1016/j.cell.2013.01.009. URL <https://www.ncbi.nlm.nih.gov/pubmed/23374351>.
- S. Palikyras and A. Papantonis. Modes of phase separation affecting chromatin regulation. *Open Biol*, 9(10):190167, 2019. ISSN 2046-2441 (Electronic) 2046-2441 (Linking). doi: 10.1098/rsob.190167. URL <https://www.ncbi.nlm.nih.gov/pubmed/31615334>.
- A. Zirkel, M. Nikolic, K. Sofiadis, J. P. Mallm, C. A. Brackley, H. Gothe, O. Drechsel, C. Becker, J. Altmüller, N. Josipovic, T. Georgomanolis, L. Brant, J. Franzen, M. Koker, E. G. Gusmao, I. G. Costa, R. T. Ullrich, W. Wagner, V. Roukos, P. Nurnberg, D. Marenduzzo, K. Rippe, and A. Papantonis. Hmgb2 loss upon senescence entry disrupts genomic organization and induces ctcf clustering across cell types. *Mol Cell*, 70(4):730–744 e6, 2018. ISSN 1097-4164 (Electronic) 1097-2765 (Linking). doi: 10.1016/j.molcel.2018.03.030. URL <https://www.ncbi.nlm.nih.gov/pubmed/29706538>.
- A. G. Larson, D. Elnatan, M. M. Keenen, M. J. Trnka, J. B. Johnston, A. L. Burlingame, D. A. Agard, S. Redding, and G. J. Narlikar. Liquid droplet

- formation by hp1alpha suggests a role for phase separation in heterochromatin. *Nature*, 547(7662):236–240, 2017. ISSN 1476-4687 (Electronic) 0028-0836 (Linking). doi: 10.1038/nature22822. URL <https://www.ncbi.nlm.nih.gov/pubmed/28636604>.
- A. R. Strom, A. V. Emelyanov, M. Mir, D. V. Fyodorov, X. Darzacq, and G. H. Karpen. Phase separation drives heterochromatin domain formation. *Nature*, 547(7662):241–245, 2017. ISSN 1476-4687 (Electronic) 0028-0836 (Linking). doi: 10.1038/nature22989. URL <https://www.ncbi.nlm.nih.gov/pubmed/28636597>.
- C. P. Brangwynne, T. J. Mitchison, and A. A. Hyman. Active liquid-like behavior of nucleoli determines their size and shape in xenopus laevis oocytes. *Proc Natl Acad Sci U S A*, 108(11):4334–9, 2011. ISSN 1091-6490 (Electronic) 0027-8424 (Linking). doi: 10.1073/pnas.1017150108. URL <https://www.ncbi.nlm.nih.gov/pubmed/21368180>.
- F. Frottin, F. Schueder, S. Tiwary, R. Gupta, R. Korner, T. Schlichthaerle, J. Cox, R. Jungmann, F. U. Hartl, and M. S. Hipp. The nucleolus functions as a phase-separated protein quality control compartment. *Science*, 365(6451):342–347, 2019. ISSN 1095-9203 (Electronic) 0036-8075 (Linking). doi: 10.1126/science.aaw9157. URL <https://www.ncbi.nlm.nih.gov/pubmed/31296649>.
- T. Cremer and M. Cremer. Chromosome territories. *Cold Spring Harb Perspect Biol*, 2(3):a003889, 2010. ISSN 1943-0264 (Electronic) 1943-0264 (Linking). doi: 10.1101/cshperspect.a003889. URL <https://www.ncbi.nlm.nih.gov/pubmed/20300217>.
- M. R. Matarazzo, S. Boyle, M. D’Esposito, and W. A. Bickmore. Chromosome territory reorganization in a human disease with altered dna methylation. *Proc Natl Acad Sci U S A*, 104(42):16546–51, 2007. ISSN 0027-8424 (Print) 0027-8424 (Linking). doi: 10.1073/pnas.0702924104. URL <https://www.ncbi.nlm.nih.gov/pubmed/17923676>.

- W. Reik, W. Dean, and J. Walter. Epigenetic reprogramming in mammalian development. *Science*, 293(5532):1089–93, 2001. ISSN 0036-8075 (Print) 0036-8075 (Linking). doi: 10.1126/science.1063443. URL <https://www.ncbi.nlm.nih.gov/pubmed/11498579>.
- Wendy Dean, Fátima Santos, and Wolf Reik. Epigenetic reprogramming in early mammalian development and following somatic nuclear transfer. *Seminars in Cell & Developmental Biology*, 14(1):93–100, 2003. ISSN 10849521. doi: 10.1016/s1084-9521(02)00141-6.
- M. Hemberger, W. Dean, and W. Reik. Epigenetic dynamics of stem cells and cell lineage commitment: digging waddington’s canal. *Nat Rev Mol Cell Biol*, 10(8):526–37, 2009. ISSN 1471-0080 (Electronic) 1471-0072 (Linking). doi: 10.1038/nrm2727. URL <https://www.ncbi.nlm.nih.gov/pubmed/19603040>.
- K. Kim, A. Doi, B. Wen, K. Ng, R. Zhao, P. Cahan, J. Kim, M. J. Aryee, H. Ji, L. I. Ehrlich, A. Yabuuchi, A. Takeuchi, K. C. Cunniff, H. Hongguang, S. McKinney-Freeman, O. Naveiras, T. J. Yoon, R. A. Irizarry, N. Jung, J. Seita, J. Hanna, P. Murakami, R. Jaenisch, R. Weissleder, S. H. Orkin, I. L. Weissman, A. P. Feinberg, and G. Q. Daley. Epigenetic memory in induced pluripotent stem cells. *Nature*, 467(7313):285–90, 2010. ISSN 1476-4687 (Electronic) 0028-0836 (Linking). doi: 10.1038/nature09342. URL <https://www.ncbi.nlm.nih.gov/pubmed/20644535>.
- J. Oswald, S. Engemann, N. Lane, W. Mayer, A. Olek, R. Fundele, W. Dean, W. Reik, and J. Walter. Active demethylation of the paternal genome in the mouse zygote. *Current Biology*, 10(8):475–478, 2000. ISSN 09609822. doi: 10.1016/s0960-9822(00)00448-6.
- W. Mayer, A. Niveleau, J. Walter, R. Fundele, and T. Haaf. Demethylation of the zygotic paternal genome. *Nature*, 403(6769):501–2, 2000. ISSN 0028-0836 (Print) 0028-0836 (Linking). doi: 10.1038/35000656. URL <https://www.ncbi.nlm.nih.gov/pubmed/10676950>.

- P. Sassone-Corsi. Unique chromatin remodeling and transcriptional regulation in spermatogenesis. *Science*, 296(5576):2176–8, 2002. ISSN 1095-9203 (Electronic) 0036-8075 (Linking). doi: 10.1126/science.1070963. URL <https://www.ncbi.nlm.nih.gov/pubmed/12077401>.
- D. Miller, M. Brinkworth, and D. Iles. Paternal dna packaging in spermatozoa: more than the sum of its parts? dna, histones, protamines and epigenetics. *Reproduction*, 139(2):287–301, 2010. ISSN 1741-7899 (Electronic) 1470-1626 (Linking). doi: 10.1530/REP-09-0281. URL <https://www.ncbi.nlm.nih.gov/pubmed/19759174>.
- Y. H. Jung, M. E. G. Sauria, X. Lyu, M. S. Cheema, J. Ausio, J. Taylor, and V. G. Corces. Chromatin states in mouse sperm correlate with embryonic and adult regulatory landscapes. *Cell Rep*, 18(6):1366–1382, 2017. ISSN 2211-1247 (Electronic). doi: 10.1016/j.celrep.2017.01.034. URL <https://www.ncbi.nlm.nih.gov/pubmed/28178516>.
- Y. Seki, K. Hayashi, K. Itoh, M. Mizugaki, M. Saitou, and Y. Matsui. Extensive and orderly reprogramming of genome-wide chromatin modifications associated with specification and early development of germ cells in mice. *Dev Biol*, 278(2):440–58, 2005. ISSN 0012-1606 (Print) 0012-1606 (Linking). doi: 10.1016/j.ydbio.2004.11.025. URL <https://www.ncbi.nlm.nih.gov/pubmed/15680362>.
- P. Hajkova, K. Ancelin, T. Waldmann, N. Lacoste, U. C. Lange, F. Cesari, C. Lee, G. Almouzni, R. Schneider, and M. A. Surani. Chromatin dynamics during epigenetic reprogramming in the mouse germ line. *Nature*, 452(7189):877–81, 2008. ISSN 1476-4687 (Electronic) 0028-0836 (Linking). doi: 10.1038/nature06714. URL <https://www.ncbi.nlm.nih.gov/pubmed/18354397>.
- X. Liu, C. Wang, W. Liu, J. Li, C. Li, X. Kou, J. Chen, Y. Zhao, H. Gao, H. Wang, Y. Zhang, Y. Gao, and S. Gao. Distinct features of h3k4me3 and h3k27me3 chromatin domains in pre-implantation embryos. *Nature*, 537(7621):558–562, 2016. ISSN 1476-4687 (Electronic) 0028-0836 (Linking). doi: 10.1038/nature19362. URL <https://www.ncbi.nlm.nih.gov/pubmed/27626379>.

- J. A. Dahl, I. Jung, H. Aanes, G. D. Greggains, A. Manaf, M. Lerdrup, G. Li, S. Kuan, B. Li, A. Y. Lee, S. Preissl, I. Jermstad, M. H. Haugen, R. Suganthan, M. Bjoras, K. Hansen, K. T. Dalen, P. Fedorcsak, B. Ren, and A. Klungland. Broad histone h3k4me3 domains in mouse oocytes modulate maternal-to-zygotic transition. *Nature*, 537(7621):548–552, 2016. ISSN 1476-4687 (Electronic) 0028-0836 (Linking). doi: 10.1038/nature19360. URL <https://www.ncbi.nlm.nih.gov/pubmed/27626377>.
- A. Burton and M. E. Torres-Padilla. Epigenetic reprogramming and development: a unique heterochromatin organization in the preimplantation mouse embryo. *Brief Funct Genomics*, 9(5-6):444–54, 2010. ISSN 2041-2657 (Electronic) 2041-2649 (Linking). doi: 10.1093/bfgp/elq027. URL <https://www.ncbi.nlm.nih.gov/pubmed/21186177>.
- J. Gassler, H. B. Brandao, M. Imakaev, I. M. Flyamer, S. Ladstatter, W. A. Bickmore, J. M. Peters, L. A. Mirny, and K. Tachibana. A mechanism of cohesin-dependent loop extrusion organizes zygotic genome architecture. *EMBO J*, 36(24):3600–3618, 2017. ISSN 1460-2075 (Electronic) 0261-4189 (Linking). doi: 10.15252/emboj.201798083. URL <https://www.ncbi.nlm.nih.gov/pubmed/29217590>.
- C. B. Hug and J. M. Vaquerizas. The birth of the 3d genome during early embryonic development. *Trends Genet*, 34(12):903–914, 2018. ISSN 0168-9525 (Print) 0168-9525 (Linking). doi: 10.1016/j.tig.2018.09.002. URL <https://www.ncbi.nlm.nih.gov/pubmed/30292539>.
- F. Santos, B. Hendrich, W. Reik, and W. Dean. Dynamic reprogramming of dna methylation in the early mouse embryo. *Dev Biol*, 241(1):172–82, 2002. ISSN 0012-1606 (Print) 0012-1606 (Linking). doi: 10.1006/dbio.2001.0501. URL <https://www.ncbi.nlm.nih.gov/pubmed/11784103>.
- S. Feng, S. E. Jacobsen, and W. Reik. Epigenetic reprogramming in plant and animal development. *Science*, 330(6004):622–7, 2010. ISSN 1095-9203 (Electronic) 0036-8075 (Linking). doi: 10.1126/science.1190614. URL <https://www.ncbi.nlm.nih.gov/pubmed/21030646>.

- J. R. Edwards, O. Yarychkivska, M. Boulard, and T. H. Bestor. Dna methylation and dna methyltransferases. *Epigenetics Chromatin*, 10:23, 2017. ISSN 1756-8935 (Electronic) 1756-8935 (Linking). doi: 10.1186/s13072-017-0130-8. URL <https://www.ncbi.nlm.nih.gov/pubmed/28503201>.
- S. Guibert, T. Forne, and M. Weber. Global profiling of dna methylation erasure in mouse primordial germ cells. *Genome Res*, 22(4):633–41, 2012. ISSN 1549-5469 (Electronic) 1088-9051 (Linking). doi: 10.1101/gr.130997.111. URL <https://www.ncbi.nlm.nih.gov/pubmed/22357612>.
- C. Popp, W. Dean, S. Feng, S. J. Cokus, S. Andrews, M. Pellegrini, S. E. Jacobsen, and W. Reik. Genome-wide erasure of dna methylation in mouse primordial germ cells is affected by aid deficiency. *Nature*, 463(7284):1101–5, 2010. ISSN 1476-4687 (Electronic) 0028-0836 (Linking). doi: 10.1038/nature08829. URL <https://www.ncbi.nlm.nih.gov/pubmed/20098412>.
- P. Hajkova, S. J. Jeffries, C. Lee, N. Miller, S. P. Jackson, and M. A. Surani. Genome-wide reprogramming in the mouse germ line entails the base excision repair pathway. *Science*, 329(5987):78–82, 2010. ISSN 1095-9203 (Electronic) 0036-8075 (Linking). doi: 10.1126/science.1187945. URL <https://www.ncbi.nlm.nih.gov/pubmed/20595612>.
- N. Lane, W. Dean, S. Erhardt, P. Hajkova, A. Surani, J. Walter, and W. Reik. Resistance of iaps to methylation reprogramming may provide a mechanism for epigenetic inheritance in the mouse. *Genesis*, 35(2):88–93, 2003. ISSN 1526-954X (Print) 1526-954X (Linking). doi: 10.1002/gene.10168. URL <https://www.ncbi.nlm.nih.gov/pubmed/12533790>.
- M. A. Eckersley-Maslin, C. Alda-Catalinas, and W. Reik. Dynamics of the epigenetic landscape during the maternal-to-zygotic transition. *Nat Rev Mol Cell Biol*, 2018. ISSN 1471-0080 (Electronic) 1471-0072 (Linking). doi: 10.1038/s41580-018-0008-z. URL <https://www.ncbi.nlm.nih.gov/pubmed/29686419>.
- E. Heard, J. Chaumeil, O. Masui, and I. Okamoto. Mammalian x-chromosome inactivation: an epigenetics paradigm. *Cold Spring Harb Symp Quant Biol*, 69:

- 89–102, 2004. ISSN 0091-7451 (Print) 0091-7451 (Linking). doi: 10.1101/sqb.2004.69.89. URL <https://www.ncbi.nlm.nih.gov/pubmed/16117637>.
- M. L. Barr and E. G. Bertram. A morphological distinction between neurones of the male and female, and the behaviour of the nucleolar satellite during accelerated nucleoprotein synthesis. *Nature*, 163(4148):676, 1949. ISSN 0028-0836 (Print) 0028-0836 (Linking). doi: 10.1038/163676a0. URL <https://www.ncbi.nlm.nih.gov/pubmed/18120749>.
- S. Ohno, W. D. Kaplan, and R. Kinosita. Formation of the sex chromatin by a single x-chromosome in liver cells of rattus norvegicus. *Experimental Cell Research*, 18(2):415–418, 1959. ISSN 00144827. doi: 10.1016/0014-4827(59)90031-x.
- M. F. Lyon. Gene action in the x-chromosome of the mouse (*mus musculus* l.). *Nature*, 190:372–3, 1961. ISSN 0028-0836 (Print) 0028-0836 (Linking). doi: 10.1038/190372a0. URL <https://www.ncbi.nlm.nih.gov/pubmed/13764598>.
- M. F. Lyon. Sex chromatin and gene action in the mammalian x-chromosome. *American journal of human genetics*, 14(2):135–148, 1962. ISSN 0002-9297 1537-6605. URL <https://pubmed.ncbi.nlm.nih.gov/14467629https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1932279/>.
- M. F. Lyon. Attempts to test the inactive-x theory of dosage compensation in mammals. *Genetical Research*, 4(01), 1963. ISSN 0016-6723 1469-5073. doi: 10.1017/s0016672300003451.
- D. W. Cooper, J. L. VandeBerg, G. B. Sharman, and W. E. Poole. Phosphoglycerate kinase polymorphism in kangaroos provides further evidence for paternal x inactivation. *Nat New Biol*, 230(13):155–7, 1971. ISSN 0090-0028 (Print) 0090-0028 (Linking). doi: 10.1038/newbio230155a0. URL <https://www.ncbi.nlm.nih.gov/pubmed/5279474>.
- N. Takagi and M. Sasaki. Preferential inactivation of the paternally derived x chromosome in the extraembryonic membranes of the mouse. *Nature*, 256(5519):640–2, 1975. ISSN 0028-0836 (Print) 0028-0836. doi: 10.1038/256640a0.

- F. Xue, X. C. Tian, F. Du, C. Kubota, M. Taneja, A. Dinnyes, Y. Dai, H. Levine, L. V. Pereira, and X. Yang. Aberrant patterns of x chromosome inactivation in bovine clones. *Nat Genet*, 31(2):216–20, 2002. ISSN 1061-4036 (Print) 1061-4036 (Linking). doi: 10.1038/ng900. URL <https://www.ncbi.nlm.nih.gov/pubmed/12032569>.
- N. Brockdorff, A. Ashworth, G. F. Kay, P. Cooper, S. Smith, V. M. McCabe, D. P. Norris, G. D. Penny, D. Patel, and S. Rastan. Conservation of position and exclusive expression of mouse xist from the inactive x chromosome. *Nature*, 351(6324):329–31, 1991. ISSN 0028-0836 (Print) 0028-0836 (Linking). doi: 10.1038/351329a0. URL <https://www.ncbi.nlm.nih.gov/pubmed/2034279>.
- N. Brockdorff. Localized accumulation of xist rna in x chromosome inactivation. *Open Biol*, 9(12):190213, 2019. ISSN 2046-2441 (Electronic) 2046-2441 (Linking). doi: 10.1098/rsob.190213. URL <https://www.ncbi.nlm.nih.gov/pubmed/31795917>.
- G. D. Penny, G. F. Kay, S. A. Sheardown, S. Rastan, and N. Brockdorff. Requirement for xist in x chromosome inactivation. *Nature*, 379(6561):131–7, 1996. ISSN 0028-0836 (Print) 0028-0836 (Linking). doi: 10.1038/379131a0. URL <https://www.ncbi.nlm.nih.gov/pubmed/8538762>.
- J. T. Lee and R. Jaenisch. Long-range cis effects of ectopic x-inactivation centres on a mouse autosome. *Nature*, 386(6622):275–9, 1997. ISSN 0028-0836 (Print) 0028-0836 (Linking). doi: 10.1038/386275a0. URL <https://www.ncbi.nlm.nih.gov/pubmed/9069285>.
- Y. Hasegawa, N. Brockdorff, S. Kawano, K. Tsutui, K. Tsutui, and S. Nakagawa. The matrix protein hnnp u is required for chromosomal localization of xist rna. *Dev Cell*, 19(3):469–76, 2010. ISSN 1878-1551 (Electronic) 1534-5807 (Linking). doi: 10.1016/j.devcel.2010.08.006. URL <https://www.ncbi.nlm.nih.gov/pubmed/20833368>.
- J. T. Lee, L. S. Davidow, and D. Warshawsky. Tsix, a gene antisense to xist at the x-inactivation centre. *Nat Genet*, 21(4):400–4, 1999. ISSN 1061-4036 (Print)

- 1061-4036 (Linking). doi: 10.1038/7734. URL <https://www.ncbi.nlm.nih.gov/pubmed/10192391>.
- P. Navarro, I. Chambers, V. Karwacki-Neisius, C. Chureau, C. Morey, C. Rougeulle, and P. Avner. Molecular coupling of xist regulation and pluripotency. *Science*, 321(5896):1693–5, 2008. ISSN 1095-9203 (Electronic) 0036-8075 (Linking). doi: 10.1126/science.1160952. URL <https://www.ncbi.nlm.nih.gov/pubmed/18802003>.
- S. Augui, E. P. Nora, and E. Heard. Regulation of x-chromosome inactivation by the x-inactivation centre. *Nat Rev Genet*, 12(6):429–42, 2011. ISSN 1471-0064 (Electronic) 1471-0056 (Linking). doi: 10.1038/nrg2987. URL <https://www.ncbi.nlm.nih.gov/pubmed/21587299>.
- E. Maclary, M. Hinten, C. Harris, and S. Kalantry. Long noncoding rnas in the x-inactivation center. *Chromosome Res*, 21(6-7):601–614, 2013. ISSN 1573-6849 (Electronic) 0967-3849 (Linking). doi: 10.1007/s10577-013-9396-2. URL <https://www.ncbi.nlm.nih.gov/pubmed/24297756>.
- Jeannie T. Lee, William M. Strauss, Jessica A. Dausman, and Rudolf Jaenisch. A 450 kb transgene displays properties of the mammalian x-inactivation center. *Cell*, 86(1):83–94, 1996. ISSN 00928674. doi: 10.1016/s0092-8674(00)80079-3.
- A. Wutz, T. P. Rasmussen, and R. Jaenisch. Chromosomal silencing and localization are mediated by different domains of xist rna. *Nat Genet*, 30(2):167–74, 2002. ISSN 1061-4036 (Print) 1061-4036 (Linking). doi: 10.1038/ng820. URL <https://www.ncbi.nlm.nih.gov/pubmed/11780141>.
- Susumu Ohno. *Sex chromosomes and sex linked genes*, volume 1 of *Monographs on Endocrinology*. Springer Berlin Heidelberg, 1967. ISBN 978-3-642-88178-7. doi: 10.1007/978-3-642-88178-7.
- I. Jonkers, T. S. Barakat, E. M. Achame, K. Monkhorst, A. Kenter, E. Rentmeester, F. Grosveld, J. A. Grootegoed, and J. Gribnau. Rnf12 is an x-encoded dose-dependent activator of x chromosome inactivation. *Cell*, 139(5):999–1011,

2009. ISSN 1097-4172 (Electronic) 0092-8674 (Linking). doi: 10.1016/j.cell.2009.10.034. URL <https://www.ncbi.nlm.nih.gov/pubmed/19945382>.
- T. S. Barakat, N. Gunhanlar, C. G. Pardo, E. M. Achame, M. Ghazvini, R. Boers, A. Kenter, E. Rentmeester, J. A. Grootegoed, and J. Gribnau. Rnf12 activates xist and is essential for x chromosome inactivation. *PLoS Genet*, 7(1):e1002001, 2011. ISSN 1553-7404 (Electronic) 1553-7390 (Linking). doi: 10.1371/journal.pgen.1002001. URL <https://www.ncbi.nlm.nih.gov/pubmed/21298085>.
- C. Gontan, E. M. Achame, J. Demmers, T. S. Barakat, E. Rentmeester, IJcken W. van, J. A. Grootegoed, and J. Gribnau. Rnf12 initiates x-chromosome inactivation by targeting rex1 for degradation. *Nature*, 485(7398):386–90, 2012. ISSN 1476-4687 (Electronic) 0028-0836 (Linking). doi: 10.1038/nature11070. URL <https://www.ncbi.nlm.nih.gov/pubmed/22596162>.
- J. Shin, M. C. Wallingford, J. Gallant, C. Marcho, B. Jiao, M. Byron, M. Bossenz, J. B. Lawrence, S. N. Jones, J. Mager, and I. Bach. Rlim is dispensable for x-chromosome inactivation in the mouse embryonic epiblast. *Nature*, 511(7507):86–9, 2014. ISSN 1476-4687 (Electronic) 0028-0836 (Linking). doi: 10.1038/nature13286. URL <https://www.ncbi.nlm.nih.gov/pubmed/24870238>.
- C. A. McHugh, C. K. Chen, A. Chow, C. F. Surka, C. Tran, P. McDonel, A. Pandya-Jones, M. Blanco, C. Burghard, A. Moradian, M. J. Sweredoski, A. A. Shishkin, J. Su, E. S. Lander, S. Hess, K. Plath, and M. Guttman. The xist lncrna interacts directly with sharp to silence transcription through hdac3. *Nature*, 521(7551):232–6, 2015. ISSN 1476-4687 (Electronic) 0028-0836 (Linking). doi: 10.1038/nature14443. URL <https://www.ncbi.nlm.nih.gov/pubmed/25915022>.
- A. Minajigi, J. Froberg, C. Wei, H. Sunwoo, B. Kesner, D. Colognori, D. Lessing, B. Payer, M. Boukhali, W. Haas, and J. T. Lee. Chromosomes. a comprehensive xist interactome reveals cohesin repulsion and an rna-directed chromosome conformation. *Science*, 349(6245), 2015. ISSN 1095-9203 (Electronic) 0036-8075 (Linking). doi: 10.1126/science.aab2276aab227610.1126/science.aab2276. URL <https://www.ncbi.nlm.nih.gov/pubmed/26089354>.

- C. Chu, Q. C. Zhang, S. T. da Rocha, R. A. Flynn, M. Bharadwaj, J. M. Calabrese, T. Magnuson, E. Heard, and H. Y. Chang. Systematic discovery of xist rna binding proteins. *Cell*, 161(2):404–16, 2015. ISSN 1097-4172 (Electronic) 0092-8674 (Linking). doi: 10.1016/j.cell.2015.03.025. URL <https://www.ncbi.nlm.nih.gov/pubmed/25843628>.
- F. Dossin, I. Pinheiro, J. J. Zyllicz, J. Roensch, S. Collombet, A. Le Saux, T. Chelmicki, M. Attia, V. Kapoor, Y. Zhan, F. Dingli, D. Loew, T. Mercher, J. Dekker, and E. Heard. Spen integrates transcriptional and epigenetic control of x-inactivation. *Nature*, 578(7795):455–460, 2020. doi: 10.1038/s41586-020-1974-9. URL <https://www.ncbi.nlm.nih.gov/pubmed/32025035>.
- J. Zhao, B. K. Sun, J. A. Erwin, J. J. Song, and J. T. Lee. Polycomb proteins targeted by a short repeat rna to the mouse x chromosome. *Science*, 322(5902):750–6, 2008. ISSN 1095-9203 (Electronic) 0036-8075 (Linking). doi: 10.1126/science.1163045. URL <https://www.ncbi.nlm.nih.gov/pubmed/18974356>.
- M. Almeida, G. Pintacuda, O. Masui, Y. Koseki, M. Gdula, A. Cerase, D. Brown, A. Mould, C. Innocent, M. Nakayama, L. Schermelleh, T. B. Nesterova, H. Koseki, and N. Brockdorff. Pcgf3/5-prc1 initiates polycomb recruitment in x chromosome inactivation. *Science*, 356(6342):1081–1084, 2017. ISSN 1095-9203 (Electronic) 0036-8075 (Linking). doi: 10.1126/science.aal2512. URL <https://www.ncbi.nlm.nih.gov/pubmed/28596365>.
- N. Brockdorff. Polycomb complexes in x chromosome inactivation. *Philos Trans R Soc Lond B Biol Sci*, 372(1733), 2017. ISSN 1471-2970 (Electronic) 0962-8436 (Linking). doi: 10.1098/rstb.2017.0021. URL <https://www.ncbi.nlm.nih.gov/pubmed/28947664>.
- S. Kalantry, K. C. Mills, D. Yee, A. P. Otte, B. Panning, and T. Magnuson. The polycomb group protein eed protects the inactive x-chromosome from differentiation-induced reactivation. *Nat Cell Biol*, 8(2):195–202, 2006. ISSN 1465-7392 (Print) 1465-7392 (Linking). doi: 10.1038/ncb1351. URL <https://www.ncbi.nlm.nih.gov/pubmed/16415857>.

- A. Keniry, L. J. Gearing, N. Jansz, J. Liu, A. Z. Holik, P. F. Hickey, S. A. Kinkel, D. L. Moore, K. Breslin, K. Chen, R. Liu, C. Phillips, M. Pakusch, C. Biben, J. M. Sheridan, B. T. Kile, C. Carmichael, M. E. Ritchie, D. J. Hilton, and M. E. Blewitt. Setdb1-mediated h3k9 methylation is enriched on the inactive x and plays a role in its epigenetic silencing. *Epigenetics Chromatin*, 9:16, 2016. ISSN 1756-8935 (Print) 1756-8935 (Linking). doi: 10.1186/s13072-016-0064-6. URL <https://www.ncbi.nlm.nih.gov/pubmed/27195021>.
- R. S. Hansen, T. K. Canfield, A. D. Fjeld, and S. M. Gartler. Role of late replication timing in the silencing of x-linked genes. *Hum Mol Genet*, 5(9):1345–53, 1996. ISSN 0964-6906 (Print) 0964-6906 (Linking). doi: 10.1093/hmg/5.9.1345. URL <https://www.ncbi.nlm.nih.gov/pubmed/8872476>.
- C. Costanzi and J. R. Pehrson. Histone macroh2a1 is concentrated in the inactive x chromosome of female mammals. *Nature*, 393(6685):599–601, 1998. ISSN 0028-0836 (Print) 0028-0836 (Linking). doi: 10.1038/31275. URL <https://www.ncbi.nlm.nih.gov/pubmed/9634239>.
- T. Sado, M. H. Fenner, S. S. Tan, P. Tam, T. Shioda, and E. Li. X inactivation in the mouse embryo deficient for dnmt1: distinct effect of hypomethylation on imprinted and random x inactivation. *Dev Biol*, 225(2):294–303, 2000. ISSN 0012-1606 (Print) 0012-1606 (Linking). doi: 10.1006/dbio.2000.9823. URL <https://www.ncbi.nlm.nih.gov/pubmed/10985851>.
- T. Pollex and E. Heard. Nuclear positioning and pairing of x-chromosome inactivation centers are not primary determinants during initiation of random x-inactivation. *Nat Genet*, 51(2):285–295, 2019. ISSN 1546-1718 (Electronic) 1061-4036 (Linking). doi: 10.1038/s41588-018-0305-7. URL <https://www.ncbi.nlm.nih.gov/pubmed/30643252>.
- C. K. Chen, M. Blanco, C. Jackson, E. Aznauryan, N. Ollikainen, C. Surka, A. Chow, A. Cerase, P. McDonel, and M. Guttman. Xist recruits the x chromosome to the nuclear lamina to enable chromosome-wide silencing. *Science*, 354(6311):468–472, 2016a. ISSN 1095-9203 (Electronic) 0036-8075 (Linking).

- doi: 10.1126/science.aae0047. URL <https://www.ncbi.nlm.nih.gov/pubmed/27492478>.
- L. F. Zhang, K. D. Huynh, and J. T. Lee. Perinucleolar targeting of the inactive x during s phase: evidence for a role in the maintenance of silencing. *Cell*, 129(4):693–706, 2007. ISSN 0092-8674 (Print) 0092-8674 (Linking). doi: 10.1016/j.cell.2007.03.036. URL <https://www.ncbi.nlm.nih.gov/pubmed/17512404>.
- E. P. Nora, B. R. Lajoie, E. G. Schulz, L. Giorgetti, I. Okamoto, N. Servant, T. Piolot, N. L. van Berkum, J. Meisig, J. Sedat, J. Gribnau, E. Barillot, N. Bluthgen, J. Dekker, and E. Heard. Spatial partitioning of the regulatory landscape of the x-inactivation centre. *Nature*, 485(7398):381–5, 2012. ISSN 1476-4687 (Electronic) 0028-0836 (Linking). doi: 10.1038/nature11049. URL <https://www.ncbi.nlm.nih.gov/pubmed/22495304>.
- L. Giorgetti, B. R. Lajoie, A. C. Carter, M. Attia, Y. Zhan, J. Xu, C. J. Chen, N. Kaplan, H. Y. Chang, E. Heard, and J. Dekker. Structural organization of the inactive x chromosome in the mouse. *Nature*, 535(7613):575–9, 2016. ISSN 1476-4687 (Electronic) 0028-0836 (Linking). doi: 10.1038/nature18589. URL <https://www.ncbi.nlm.nih.gov/pubmed/27437574>.
- E. Splinter, E. de Wit, E. P. Nora, P. Klous, H. J. van de Werken, Y. Zhu, L. J. Kaaij, W. van Ijcken, J. Gribnau, E. Heard, and W. de Laat. The inactive x chromosome adopts a unique three-dimensional conformation that is dependent on xist rna. *Genes Dev*, 25(13):1371–83, 2011. ISSN 1549-5477 (Electronic) 0890-9369 (Linking). doi: 10.1101/gad.633311. URL <https://www.ncbi.nlm.nih.gov/pubmed/21690198>.
- S. S. Rao, M. H. Huntley, N. C. Durand, E. K. Stamenova, I. D. Bochkov, J. T. Robinson, A. L. Sanborn, I. Machol, A. D. Omer, E. S. Lander, and E. L. Aiden. A 3d map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*, 159(7):1665–80, 2014. ISSN 1097-4172 (Electronic) 0092-8674 (Linking). doi: 10.1016/j.cell.2014.11.021. URL <https://www.ncbi.nlm.nih.gov/pubmed/25497547>.

- X. Deng, W. Ma, V. Ramani, A. Hill, F. Yang, F. Ay, J. B. Berletch, C. A. Blau, J. Shendure, Z. Duan, W. S. Noble, and C. M. Disteche. Bipartite structure of the inactive mouse x chromosome. *Genome Biol*, 16:152, 2015. ISSN 1474-760X (Electronic) 1474-7596 (Linking). doi: 10.1186/s13059-015-0728-8. URL <https://www.ncbi.nlm.nih.gov/pubmed/26248554>.
- T. Jegu, E. Aeby, and J. T. Lee. The x chromosome in space. *Nat Rev Genet*, 2017. doi: 10.1038/nrg.2017.17. URL <https://www.ncbi.nlm.nih.gov/pubmed/28479596>.
- C. Y. Wang, T. Jegu, H. P. Chu, H. J. Oh, and J. T. Lee. Smchd1 merges chromosome compartments and assists formation of super-structures on the inactive x. *Cell*, 174(2):406–421 e25, 2018. ISSN 1097-4172 (Electronic) 0092-8674 (Linking). doi: 10.1016/j.cell.2018.05.007. URL <https://www.ncbi.nlm.nih.gov/pubmed/29887375>.
- C. Y. Wang, D. Colognori, H. Sunwoo, D. Wang, and J. T. Lee. Prc1 collaborates with smchd1 to fold the x-chromosome and spread xist rna between chromosome compartments. *Nat Commun*, 10(1):2950, 2019. ISSN 2041-1723 (Electronic) 2041-1723 (Linking). doi: 10.1038/s41467-019-10755-3. URL <https://www.ncbi.nlm.nih.gov/pubmed/31270318>.
- M. R. Gdula, T. B. Nesterova, G. Pintacuda, J. Godwin, Y. Zhan, H. Ozadam, M. McClellan, D. Moralli, F. Krueger, C. M. Green, W. Reik, S. Kriaucionis, E. Heard, J. Dekker, and N. Brockdorff. The non-canonical smc protein smchd1 antagonises tad formation and compartmentalisation on the inactive x chromosome. *Nat Commun*, 10(1):30, 2019. ISSN 2041-1723 (Electronic) 2041-1723 (Linking). doi: 10.1038/s41467-018-07907-2. URL <https://www.ncbi.nlm.nih.gov/pubmed/30604745>.
- M. Blewitt and E. Whitelaw. The use of mouse models to study epigenetics. *Cold Spring Harb Perspect Biol*, 5(11):a017939, 2013. ISSN 1943-0264 (Electronic) 1943-0264 (Linking). doi: 10.1101/cshperspect.a017939. URL <https://www.ncbi.nlm.nih.gov/pubmed/24186070>.

- A. Inoue, L. Jiang, F. Lu, and Y. Zhang. Genomic imprinting of xist by maternal h3k27me3. *Genes Dev*, 31(19):1927–1932, 2017a. ISSN 1549-5477 (Electronic) 0890-9369 (Linking). doi: 10.1101/gad.304113.117. URL <https://www.ncbi.nlm.nih.gov/pubmed/29089420>.
- I. Okamoto and E. Heard. The dynamics of imprinted x inactivation during preimplantation development in mice. *Cytogenet Genome Res*, 113(1-4):318–24, 2006. ISSN 1424-859X (Electronic) 1424-8581 (Linking). doi: 10.1159/000090848. URL <https://www.ncbi.nlm.nih.gov/pubmed/16575196>.
- T. Sado and A. C. Ferguson-Smith. Imprinted x inactivation and reprogramming in the preimplantation mouse embryo. *Hum Mol Genet*, 14 Spec No 1:R59–64, 2005. ISSN 0964-6906 (Print) 0964-6906 (Linking). doi: 10.1093/hmg/ddi117. URL <https://www.ncbi.nlm.nih.gov/pubmed/15809274>.
- W. Mak, T. B. Nesterova, M. de Napoles, R. Appanah, S. Yamanaka, A. P. Otte, and N. Brockdorff. Reactivation of the paternal x chromosome in early mouse embryos. *Science*, 303(5658):666–9, 2004. ISSN 1095-9203 (Electronic) 0036-8075 (Linking). doi: 10.1126/science.1092674. URL <https://www.ncbi.nlm.nih.gov/pubmed/14752160>.
- M. Sugimoto and K. Abe. X chromosome reactivation initiates in nascent primordial germ cells in mice. *PLoS Genet*, 3(7):e116, 2007. ISSN 1553-7404 (Electronic) 1553-7390 (Linking). doi: 10.1371/journal.pgen.0030116. URL <https://www.ncbi.nlm.nih.gov/pubmed/17676999>.
- M. A. Surani, S. C. Barton, and M. L. Norris. Development of reconstituted mouse eggs suggests imprinting of the genome during gametogenesis. *Nature*, 308(5959):548–50, 1984. ISSN 0028-0836 (Print) 0028-0836 (Linking). doi: 10.1038/308548a0. URL <https://www.ncbi.nlm.nih.gov/pubmed/6709062>.
- James McGrath and Davor Solter. Completion of mouse embryogenesis requires both the maternal and paternal genomes. *Cell*, 37(1):179–183, 1984. ISSN 00928674. doi: 10.1016/0092-8674(84)90313-1.

- S. C. Barton, M. A. Surani, and M. L. Norris. Role of paternal and maternal genomes in mouse development. *Nature*, 311(5984):374–6, 1984. ISSN 0028-0836 (Print) 0028-0836 (Linking). doi: 10.1038/311374a0. URL <https://www.ncbi.nlm.nih.gov/pubmed/6482961>.
- J. A. Thomson and D. Solter. The developmental fate of androgenetic, parthenogenetic, and gynogenetic cells in chimeric gastrulating mouse embryos. *Genes Dev*, 2(10):1344–51, 1988. ISSN 0890-9369 (Print) 0890-9369 (Linking). doi: 10.1101/gad.2.10.1344. URL <https://www.ncbi.nlm.nih.gov/pubmed/3203909>.
- M. Azim Surani, Rashmi Kothary, Nicholas D. Allen, Prim B. Singh, Reinald Fundele, Anne C. Ferguson-Smith, and Sheila C. Barton. Genome imprinting and development in the mouse. *Development*, 108(Supplement):89, 1990. URL <http://dev.biologists.org/content/108/Supplement/89.abstract>.
- A. C. Ferguson-Smith. Genomic imprinting: the emergence of an epigenetic paradigm. *Nat Rev Genet*, 12(8):565–75, 2011. ISSN 1471-0064 (Electronic) 1471-0056 (Linking). doi: 10.1038/nrg3032. URL <https://www.ncbi.nlm.nih.gov/pubmed/21765458>.
- B. M. Cattanach and M. Kirk. Differential activity of maternally and paternally derived chromosome regions in mice. *Nature*, 315(6019):496–8, 1985. ISSN 0028-0836 (Print) 0028-0836 (Linking). doi: 10.1038/315496a0. URL <https://www.ncbi.nlm.nih.gov/pubmed/4000278>.
- Bruce M. Cattanach and Colin V. Beechey. Autosomal and x-chromosome imprinting. *Development*, 108(Supplement):63, 1990. URL <http://dev.biologists.org/content/108/Supplement/63.abstract>.
- M. S. Bartolomei, S. Zemel, and S. M. Tilghman. Parental imprinting of the mouse h19 gene. *Nature*, 351(6322):153–5, 1991. ISSN 0028-0836 (Print) 0028-0836 (Linking). doi: 10.1038/351153a0. URL <https://www.ncbi.nlm.nih.gov/pubmed/1709450>.
- A. C. Ferguson-Smith, B. M. Cattanach, S. C. Barton, C. V. Beechey, and M. A. Surani. Embryological and molecular investigations of parental imprinting on

- mouse chromosome 7. *Nature*, 351(6328):667–70, 1991. ISSN 0028-0836 (Print) 0028-0836 (Linking). doi: 10.1038/351667a0. URL <https://www.ncbi.nlm.nih.gov/pubmed/2052093>.
- Thomas M. DeChiara, Elizabeth J. Robertson, and Argiris Efstratiadis. Parental imprinting of the mouse insulin-like growth factor ii gene. *Cell*, 64(4):849–859, 1991. ISSN 00928674. doi: 10.1016/0092-8674(91)90513-x.
- D. P. Barlow, R. Stoger, B. G. Herrmann, K. Saito, and N. Schweifer. The mouse insulin-like growth factor type-2 receptor is imprinted and closely linked to the tme locus. *Nature*, 349(6304):84–7, 1991. ISSN 0028-0836 (Print) 0028-0836 (Linking). doi: 10.1038/349084a0. URL <https://www.ncbi.nlm.nih.gov/pubmed/1845916>.
- D. P. Barlow and M. S. Bartolomei. Genomic imprinting in mammals. *Cold Spring Harb Perspect Biol*, 6(2), 2014. ISSN 1943-0264 (Electronic) 1943-0264 (Linking). doi: 10.1101/cshperspect.a018382. URL <https://www.ncbi.nlm.nih.gov/pubmed/24492710>.
- S. Gigante, Q. Gouil, A. Lucattini, A. Keniry, T. Beck, M. Tinning, L. Gordon, C. Woodruff, T. P. Speed, M. E. Blewitt, and M. E. Ritchie. Using long-read sequencing to detect imprinted dna methylation. *Nucleic Acids Res*, 47(8):e46, 2019. ISSN 1362-4962 (Electronic) 0305-1048 (Linking). doi: 10.1093/nar/gkz107. URL <https://www.ncbi.nlm.nih.gov/pubmed/30793194>.
- B. Neumann, P. Kubicka, and D. P. Barlow. Characteristics of imprinted genes. *Nat Genet*, 9(1):12–3, 1995. ISSN 1061-4036 (Print) 1061-4036 (Linking). doi: 10.1038/ng0195-12. URL <https://www.ncbi.nlm.nih.gov/pubmed/7704015>.
- M. Paulsen, O. El-Maarri, S. Engemann, M. Stroedicke, O. Franck, K. Davies, R. Reinhardt, W. Reik, and J. Walter. Sequence conservation and variability of imprinting in the beckwith-wiedemann syndrome gene cluster in human and mouse. *Hum Mol Genet*, 9(12):1829–41, 2000. ISSN 0964-6906 (Print) 0964-6906 (Linking). doi: 10.1093/hmg/9.12.1829. URL <https://www.ncbi.nlm.nih.gov/pubmed/10915772>.

- W. Reik, A. Collick, M. L. Norris, S. C. Barton, and M. A. Surani. Genomic imprinting determines methylation of parental alleles in transgenic mice. *Nature*, 328(6127):248–51, 1987. ISSN 0028-0836 (Print) 0028-0836 (Linking). doi: 10.1038/328248a0. URL <https://www.ncbi.nlm.nih.gov/pubmed/3600805>.
- C. Sapienza, A. C. Peterson, J. Rossant, and R. Balling. Degree of methylation of transgenes is dependent on gamete of origin. *Nature*, 328(6127):251–4, 1987. ISSN 0028-0836 (Print) 0028-0836 (Linking). doi: 10.1038/328251a0. URL <https://www.ncbi.nlm.nih.gov/pubmed/3600806>.
- Judith L. Swain, Timothy A. Stewart, and Philip Leder. Parental legacy determines methylation and expression of an autosomal transgene: A molecular mechanism for parental imprinting. *Cell*, 50(5):719–727, 1987. ISSN 00928674. doi: 10.1016/0092-8674(87)90330-8.
- F. Sleutels, R. Zwart, and D. P. Barlow. The non-coding air rna is required for silencing autosomal imprinted genes. *Nature*, 415(6873):810–3, 2002. ISSN 0028-0836 (Print) 0028-0836 (Linking). doi: 10.1038/415810a. URL <https://www.ncbi.nlm.nih.gov/pubmed/11845212>.
- D. Mancini-Dinardo, S. J. Steele, J. M. Levorse, R. S. Ingram, and S. M. Tilghman. Elongation of the *kcnqlot1* transcript is required for genomic imprinting of neighboring genes. *Genes Dev*, 20(10):1268–82, 2006. ISSN 0890-9369 (Print) 0890-9369 (Linking). doi: 10.1101/gad.1416906. URL <https://www.ncbi.nlm.nih.gov/pubmed/16702402>.
- A. Inoue, L. Jiang, F. Lu, T. Suzuki, and Y. Zhang. Maternal h3k27me3 controls dna methylation-independent imprinting. *Nature*, 547(7664):419–424, 2017b. ISSN 1476-4687 (Electronic) 0028-0836 (Linking). doi: 10.1038/nature23262. URL <https://www.ncbi.nlm.nih.gov/pubmed/28723896>.
- C. W. Hanna and G. Kelsey. Genomic imprinting beyond dna methylation: a role for maternal histones. *Genome Biol*, 18(1):177, 2017. ISSN 1474-760X (Electronic) 1474-7596 (Linking). doi: 10.1186/s13059-017-1317-9. URL <https://www.ncbi.nlm.nih.gov/pubmed/28927436>.

- H. Zheng, B. Huang, B. Zhang, Y. Xiang, Z. Du, Q. Xu, Y. Li, Q. Wang, J. Ma, X. Peng, F. Xu, and W. Xie. Resetting epigenetic memory by reprogramming of histone modifications in mammals. *Mol Cell*, 63(6):1066–79, 2016. ISSN 1097-4164 (Electronic) 1097-2765 (Linking). doi: 10.1016/j.molcel.2016.08.032. URL <https://www.ncbi.nlm.nih.gov/pubmed/27635762>.
- Tom Moore and Walter Mills. Evolutionary theories of imprinting— enough already! In Jon F. Wilkins, editor, *Genomic Imprinting*, pages 116–122. Springer New York, New York, NY, 2008. ISBN 978-0-387-77576-0. doi: 10.1007/978-0-387-77576-0_9. URL https://doi.org/10.1007/978-0-387-77576-0_9.
- T. Moore. Genomic imprinting in mammalian development: a parental tug-of-war. *Trends in Genetics*, 7(2):45–49, 1991. ISSN 01689525. doi: 10.1016/0168-9525(91)90230-n.
- R. D. Nicholls, J. H. Knoll, M. G. Butler, S. Karam, and M. Lalande. Genetic imprinting suggested by maternal heterodisomy in nondeletion prader-willi syndrome. *Nature*, 342(6247):281–5, 1989. ISSN 0028-0836 (Print) 0028-0836 (Linking). doi: 10.1038/342281a0. URL <https://www.ncbi.nlm.nih.gov/pubmed/2812027>.
- A. J. Ping, A. E. Reeve, D. J. Law, M. R. Young, M. Boehnke, and A. P. Feinberg. Genetic linkage of beckwith-wiedemann syndrome to 11p15. *American journal of human genetics*, 44(5):720–723, 1989. ISSN 0002-9297 1537-6605. URL <https://pubmed.ncbi.nlm.nih.gov/2565083https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1715646/>.
- A. Koufos, P. Grundy, K. Morgan, K. A. Aleck, T. Hadro, B. C. Lampkin, A. Kalbakji, and W. K. Cavenee. Familial wiedemann-beckwith syndrome and a second wilms tumor locus both map to 11p15.5. *Am J Hum Genet*, 44(5):711–9, 1989. ISSN 0002-9297 (Print) 0002-9297.

- J. Peters. The role of genomic imprinting in biology and disease: an expanding view. *Nat Rev Genet*, 15(8):517–30, 2014. ISSN 1471-0064 (Electronic) 1471-0056 (Linking). doi: 10.1038/nrg3766. URL <https://www.ncbi.nlm.nih.gov/pubmed/24958438>.
- M. P. Adam, H. H. Ardinger, R. A. Pagon, S. E. Wallace, L. J. H. Bean, K. Stephens, and A. Amemiya. *GeneReviews(®)[Internet]*. University of Washington, Seattle Copyright © 1993-2020, University of Washington, Seattle. GeneReviews is a registered trademark of the University of Washington, Seattle. All rights reserved., Seattle (WA), 1993-2020. Adam, Margaret P Ardinger, Holly H Pagon, Roberta A Wallace, Stephanie E Bean, Lora JH Stephens, Karen Amemiya, Anne Review Book NBK1116 [bookaccession].
- D. Monk, D. J. G. Mackay, T. Eggermann, E. R. Maher, and A. Riccio. Genomic imprinting disorders: lessons on how genome, epigenome and environment interact. *Nat Rev Genet*, 20(4):235–248, 2019. ISSN 1471-0064 (Electronic) 1471-0056 (Linking). doi: 10.1038/s41576-018-0092-0. URL <https://www.ncbi.nlm.nih.gov/pubmed/30647469>.
- F. Court, A. Martin-Trujillo, V. Romanelli, I. Garin, I. Iglesias-Platas, I. Salafsky, M. Guitart, G. Perez de Nanclares, P. Lapunzina, and D. Monk. Genome-wide allelic methylation analysis reveals disease-specific susceptibility to multiple methylation defects in imprinting syndromes. *Hum Mutat*, 34(4):595–602, 2013. ISSN 1098-1004 (Electronic) 1059-7794 (Linking). doi: 10.1002/humu.22276. URL <https://www.ncbi.nlm.nih.gov/pubmed/23335487>.
- Heinrich Ernst Ziegler. Experimentelle studien über die zelltheilung. *Archiv für Entwicklungsmechanik der Organismen*, 6(2):249–293, 1898. ISSN 0949-944X 1432-041X. doi: 10.1007/bf02152958.
- H. Redfield. The maternal inheritance of a sex-limited lethal effect in drosophila melanogaster. *Genetics*, 11(5):482–502, 1926. ISSN 0016-6731 (Print) 0016-6731. Redfield, H Journal Article Genetics. 1926 Sep;11(5):482-502.

- F. L. Marlow. Developmental biology. In *Maternal Control of Development in Vertebrates: My Mother Made Me Do It!* Morgan & Claypool Life Sciences Copyright © 2010 by Morgan & Claypool Life Sciences., San Rafael (CA), 2010. doi: 10.4199/c00023ed1v01y201012deb005.
- D. Bourc'his, G. L. Xu, C. S. Lin, B. Bollman, and T. H. Bestor. Dnmt3l and the establishment of maternal genomic imprints. *Science*, 294(5551):2536–9, 2001. ISSN 0036-8075 (Print) 0036-8075 (Linking). doi: 10.1126/science.1065848. URL <https://www.ncbi.nlm.nih.gov/pubmed/11719692>.
- Bernhard Payer, Mitinori Saitou, Sheila C. Barton, Rosemary Thresher, John P. C. Dixon, Dirk Zahn, William H. Colledge, Mark B. L. Carlton, Toru Nakano, and M. Azim Surani. stella is a maternal effect gene required for normal early development in mice. *Current Biology*, 13(23):2110–2117, 2003. ISSN 09609822. doi: 10.1016/j.cub.2003.11.026.
- X. Li, M. Ito, F. Zhou, N. Youngson, X. Zuo, P. Leder, and A. C. Ferguson-Smith. A maternal-zygotic effect gene, zfp57, maintains both maternal and paternal imprints. *Dev Cell*, 15(4):547–57, 2008. ISSN 1878-1551 (Electronic) 1534-5807 (Linking). doi: 10.1016/j.devcel.2008.08.014. URL <https://www.ncbi.nlm.nih.gov/pubmed/18854139>.
- J. Shin, M. Bossenz, Y. Chung, H. Ma, M. Byron, N. Taniguchi-Ishigaki, X. Zhu, B. Jiao, L. L. Hall, M. R. Green, S. N. Jones, I. Hermans-Borgmeyer, J. B. Lawrence, and I. Bach. Maternal rnf12/rlim is required for imprinted x-chromosome inactivation in mice. *Nature*, 467(7318):977–81, 2010. ISSN 1476-4687 (Electronic) 0028-0836 (Linking). doi: 10.1038/nature09457. URL <https://www.ncbi.nlm.nih.gov/pubmed/20962847>.
- D. M. Messerschmidt, W. de Vries, M. Ito, D. Solter, A. Ferguson-Smith, and B. B. Knowles. Trim28 is required for epigenetic stability during mouse oocyte to embryo transition. *Science*, 335(6075):1499–502, 2012. ISSN 1095-9203 (Electronic) 0036-8075 (Linking). doi: 10.1126/science.1216154. URL <https://www.ncbi.nlm.nih.gov/pubmed/22442485>.

- S. Mahadevan, V. Sathappan, B. Utama, I. Lorenzo, K. Kaskar, and I. B. Van den Veyver. Maternally expressed nlrp2 links the subcortical maternal complex (scmc) to fertility, embryogenesis and epigenetic reprogramming. *Sci Rep*, 7: 44667, 2017. ISSN 2045-2322 (Electronic) 2045-2322 (Linking). doi: 10.1038/srep44667. URL <https://www.ncbi.nlm.nih.gov/pubmed/28317850>.
- M. Sanchez-Delgado, A. Riccio, T. Eggermann, E. R. Maher, P. Lapunzina, D. Mackay, and D. Monk. Causes and consequences of multi-locus imprinting disturbances in humans. *Trends Genet*, 32(7):444–455, 2016. ISSN 0168-9525 (Print) 0168-9525 (Linking). doi: 10.1016/j.tig.2016.05.001. URL <https://www.ncbi.nlm.nih.gov/pubmed/27235113>.
- L. E. Docherty, F. I. Rezwan, R. L. Poole, C. L. Turner, E. Kivuva, E. R. Maher, S. F. Smithson, J. P. Hamilton-Shield, M. Patalan, M. Gizewska, J. Peregud-Pogorzelski, J. Beygo, K. Buiting, B. Horsthemke, L. Soellner, M. Begemann, T. Eggermann, E. Baple, S. Mansour, I. K. Temple, and D. J. Mackay. Mutations in nlrp5 are associated with reproductive wastage and multilocus imprinting disorders in humans. *Nat Commun*, 6:8086, 2015. ISSN 2041-1723 (Electronic) 2041-1723 (Linking). doi: 10.1038/ncomms9086. URL <https://www.ncbi.nlm.nih.gov/pubmed/26323243>.
- L. Fontana, M. F. Bedeschi, S. Maitz, A. Cereda, C. Fare, S. Motta, A. Seresini, P. D’Ursi, A. Orro, V. Pecile, M. Calvello, A. Selicorni, F. Lalatta, D. Milani, S. M. Sirchia, M. Miozzo, and S. Tabano. Characterization of multi-locus imprinting disturbances and underlying genetic defects in patients with chromosome 11p15.5 related imprinting disorders. *Epigenetics*, 13(9):897–909, 2018. ISSN 1559-2308 (Electronic) 1559-2294 (Linking). doi: 10.1080/15592294.2018.1514230. URL <https://www.ncbi.nlm.nih.gov/pubmed/30221575>.
- M. E. Blewitt, N. K. Vickaryous, S. J. Hemley, A. Ashe, T. J. Bruxner, J. I. Preis, R. Arkell, and E. Whitelaw. An n-ethyl-n-nitrosourea screen for genes involved in variegation in the mouse. *Proc Natl Acad Sci U S A*, 102(21):7629–34, 2005. ISSN 0027-8424 (Print) 0027-8424 (Linking). doi: 10.1073/pnas.0409375102. URL <https://www.ncbi.nlm.nih.gov/pubmed/15890782>.

- Jost I. Preis, Meredith Downes, Nathan A. Oates, John E. J. Rasko, and Emma Whitelaw. Sensitive flow cytometric analysis reveals a novel type of parent-of-origin effect in the mouse genome. *Current Biology*, 13(11):955–959, 2003. ISSN 09609822. doi: 10.1016/s0960-9822(03)00335-x.
- A. Ashe, D. K. Morgan, N. C. Whitelaw, T. J. Bruxner, N. K. Vickaryous, L. L. Cox, N. C. Butterfield, C. Wicking, M. E. Blewitt, S. J. Wilkins, G. J. Anderson, T. C. Cox, and E. Whitelaw. A genome-wide screen for modifiers of trans-gene variegation identifies genes with critical roles in development. *Genome Biol*, 9(12):R182, 2008. ISSN 1474-760X (Electronic) 1474-7596 (Linking). doi: 10.1186/gb-2008-9-12-r182. URL <https://www.ncbi.nlm.nih.gov/pubmed/19099580>.
- L. Daxinger, S. K. Harten, H. Oey, T. Epp, L. Isbel, E. Huang, N. Whitelaw, A. Apedaile, A. Sorolla, J. Yong, V. Bharti, J. Sutton, A. Ashe, Z. Pang, N. Wallace, D. J. Gerhardt, M. E. Blewitt, J. A. Jeddloh, and E. Whitelaw. An enu mutagenesis screen identifies novel and known genes involved in epigenetic processes in the mouse. *Genome Biol*, 14(9):R96, 2013. ISSN 1474-760X (Electronic) 1474-7596 (Linking). doi: 10.1186/gb-2013-14-9-r96. URL <https://www.ncbi.nlm.nih.gov/pubmed/24025402>.
- S. Chong, N. Vickaryous, A. Ashe, N. Zamudio, N. Youngson, S. Hemley, T. Stopka, A. Skoultchi, J. Matthews, H. S. Scott, D. de Kretser, M. O'Bryan, M. Blewitt, and E. Whitelaw. Modifiers of epigenetic reprogramming show paternal effects in the mouse. *Nat Genet*, 39(5):614–22, 2007. ISSN 1061-4036 (Print) 1061-4036 (Linking). doi: 10.1038/ng2031. URL <https://www.ncbi.nlm.nih.gov/pubmed/17450140>.
- N. C. Whitelaw, S. Chong, D. K. Morgan, C. Nestor, T. J. Bruxner, A. Ashe, E. Lambley, R. Meehan, and E. Whitelaw. Reduced levels of two modifiers of epigenetic gene silencing, dnmt3a and trim28, cause increased phenotypic noise. *Genome Biol*, 11(11):R111, 2010. ISSN 1474-760X (Electronic) 1474-7596 (Linking). doi: 10.1186/gb-2010-11-11-r111. URL <https://www.ncbi.nlm.nih.gov/pubmed/21092094>.

- N. A. Youngson, T. Epp, A. R. Roberts, L. Daxinger, A. Ashe, E. Huang, K. L. Lester, S. K. Harten, G. F. Kay, T. Cox, J. M. Matthews, S. Chong, and E. Whitelaw. No evidence for cumulative effects in a *dnmt3b* hypomorph across multiple generations. *Mamm Genome*, 24(5-6):206–17, 2013. ISSN 1432-1777 (Electronic) 0938-8990 (Linking). doi: 10.1007/s00335-013-9451-5. URL <https://www.ncbi.nlm.nih.gov/pubmed/23636699>.
- A. Sorolla, M. R. Tallack, H. Oey, S. K. Harten, L. C. Daxinger, G. W. Magor, A. N. Combes, M. Ilsley, E. Whitelaw, and A. C. Perkins. Identification of novel hypomorphic and null mutations in *klf1* derived from a genetic screen for modifiers of alpha-globin transgene variegation. *Genomics*, 105(2):116–22, 2015. ISSN 1089-8646 (Electronic) 0888-7543 (Linking). doi: 10.1016/j.ygeno.2014.09.013. URL <https://www.ncbi.nlm.nih.gov/pubmed/25451176>.
- S. K. Harten, T. J. Bruxner, V. Bharti, M. Blewitt, T. M. Nguyen, E. Whitelaw, and T. Epp. The first mouse mutants of *d14abb1e* (*fam208a*) show that it is critical for early development. *Mamm Genome*, 25(7-8):293–303, 2014. ISSN 1432-1777 (Electronic) 0938-8990 (Linking). doi: 10.1007/s00335-014-9516-0. URL <https://www.ncbi.nlm.nih.gov/pubmed/24781204>.
- R. S. Nozawa, K. Nagao, K. T. Igami, S. Shibata, N. Shirai, N. Nozaki, T. Sado, H. Kimura, and C. Obuse. Human inactive x chromosome is compacted through a *prc2*-independent *smchd1-hbix1* pathway. *Nature Structural & Molecular Biology*, 20(5):566–+, 2013. ISSN 1545-9993. doi: 10.1038/nsmb.2532. URL [GotoISI://WOS:000318617000009](https://www.ncbi.nlm.nih.gov/pubmed/23636699).
- A. V. Gendrel, A. Apedaile, H. Coker, A. Termanis, I. Zvetkova, J. Godwin, Y. A. Tang, D. Huntley, G. Montana, S. Taylor, E. Giannoulatou, E. Heard, I. Stancheva, and N. Brockdorff. *Smchd1*-dependent and -independent pathways determine developmental dynamics of *cpG* island methylation on the inactive x chromosome. *Dev Cell*, 23(2):265–79, 2012. ISSN 1878-1551 (Electronic) 1534-5807 (Linking). doi: 10.1016/j.devcel.2012.06.011. URL <https://www.ncbi.nlm.nih.gov/pubmed/22841499>.

- M. Ganji, I. A. Shaltiel, S. Bisht, E. Kim, A. Kalichava, C. H. Haering, and C. Dekker. Real-time imaging of dna loop extrusion by condensin. *Science*, 360(6384):102–105, 2018. ISSN 1095-9203 (Electronic) 0036-8075 (Linking). doi: 10.1126/science.aar7831. URL <https://www.ncbi.nlm.nih.gov/pubmed/29472443>.
- K. Chen, R. C. Dobson, I. S. Lucet, S. N. Young, F. G. Pearce, M. E. Blewitt, and J. M. Murphy. The epigenetic regulator smchd1 contains a functional ghkl-type atpase domain. *Biochem J*, 473(12):1733–44, 2016b. ISSN 1470-8728 (Electronic) 0264-6021 (Linking). doi: 10.1042/BCJ20160189. URL <https://www.ncbi.nlm.nih.gov/pubmed/27059856>.
- K. Chen, P. E. Czabotar, M. E. Blewitt, and J. M. Murphy. The hinge domain of the epigenetic repressor smchd1 adopts an unconventional homodimeric configuration. *Biochem J*, 473(6):733–42, 2016c. ISSN 1470-8728 (Electronic) 0264-6021 (Linking). doi: 10.1042/BJ20151049. URL <https://www.ncbi.nlm.nih.gov/pubmed/26733688>.
- L. C. Pedersen, K. Inoue, S. Kim, L. Perera, and N. D. Shaw. A ubiquitin-like domain is required for stabilizing the n-terminal atpase module of human smchd1. *Commun Biol*, 2:255, 2019. ISSN 2399-3642 (Electronic) 2399-3642 (Linking). doi: 10.1038/s42003-019-0499-y. URL <https://www.ncbi.nlm.nih.gov/pubmed/31312724>.
- Kelan Chen, Richard W. Birkinshaw, Alexandra D. Gurzau, Iromi Wanigasuriya, Ruoyun Wang, Megan Iminoff, Jarrod J. Sandow, Samuel N. Young, Patrick J. Hennessy, Tracy A. Willson, Denise A. Heckmann, Andrew I. Webb, Marnie E. Blewitt, Peter E. Czabotar, and James M. Murphy. Crystal structure of the hinge domain of smchd1 reveals its dimerization mode and nucleic acid-binding residues. *Science Signaling*, 13(636), 2020. ISSN 1945-0877 1937-9145. doi: 10.1126/scisignal.aaz5599.
- M. S. van Ruiten and B. D. Rowland. Smc complexes: Universal dna looping machines with distinct regulators. *Trends Genet*, 34(6):477–487, 2018. ISSN

- 0168-9525 (Print) 0168-9525 (Linking). doi: 10.1016/j.tig.2018.03.003. URL <https://www.ncbi.nlm.nih.gov/pubmed/29606284>.
- A. Lammens, A. Schele, and K. P. Hopfner. Structural biochemistry of atp-driven dimerization and dna-stimulated activation of smc atpases. *Curr Biol*, 14(19): 1778–82, 2004. ISSN 0960-9822 (Print) 0960-9822 (Linking). doi: 10.1016/j.cub.2004.09.044. URL <https://www.ncbi.nlm.nih.gov/pubmed/15458651>.
- R. Dutta and M. Inouye. Ghkl, an emergent atpase/kinase superfamily. *Trends in Biochemical Sciences*, 25(1):24–28, 2000. ISSN 09680004. doi: 10.1016/s0968-0004(99)01503-0.
- N. J. Brideau, H. Coker, A. V. Gendrel, C. A. Siebert, K. Bezstarosti, J. Demmers, R. A. Poot, T. B. Nesterova, and N. Brockdorff. Independent mechanisms target smchd1 to trimethylated histone h3 lysine 9-modified chromatin and the inactive x chromosome. *Mol Cell Biol*, 35(23):4053–68, 2015. ISSN 1098-5549 (Electronic) 0270-7306 (Linking). doi: 10.1128/MCB.00432-15. URL <https://www.ncbi.nlm.nih.gov/pubmed/26391951>.
- R. J. Lemmers, J. J. Goeman, P. J. van der Vliet, M. P. van Nieuwenhuizen, J. Balog, M. Vos-Versteeg, P. Camano, M. A. Ramos Arroyo, I. Jerico, M. T. Rogers, D. G. Miller, M. Upadhyaya, J. J. Verschuuren, A. Lopez de Munain Arregui, B. G. van Engelen, G. W. Padberg, S. Sacconi, R. Tawil, S. J. Tapscott, B. Bakker, and S. M. van der Maarel. Inter-individual differences in cpg methylation at d4z4 correlate with clinical variability in fshd1 and fshd2. *Hum Mol Genet*, 24(3):659–69, 2015. ISSN 1460-2083 (Electronic) 0964-6906 (Linking). doi: 10.1093/hmg/ddu486. URL <https://www.ncbi.nlm.nih.gov/pubmed/25256356>.
- J. C. Deenen, H. Arnts, S. M. van der Maarel, G. W. Padberg, J. J. Verschuuren, E. Bakker, S. S. Weinreich, A. L. Verbeek, and B. G. van Engelen. Population-based incidence and prevalence of facioscapulohumeral dystrophy. *Neurology*, 83(12):1056–9, 2014. ISSN 1526-632X (Electronic) 0028-3878 (Linking). doi: 10.1212/WNL.0000000000000797. URL <https://www.ncbi.nlm.nih.gov/pubmed/25122204>.

- G. W. Padberg, P. W. Lunt, M. Koch, and M. Fardeau. Diagnostic criteria for facioscapulohumeral muscular dystrophy. *Neuromuscular Disorders*, 1(4):231–234, 1991. ISSN 09608966. doi: 10.1016/0960-8966(91)90094-9.
- J. Hamel and R. Tawil. Facioscapulohumeral muscular dystrophy: Update on pathogenesis and future treatments. *Neurotherapeutics*, 2018. ISSN 1878-7479 (Electronic) 1878-7479 (Linking). doi: 10.1007/s13311-018-00675-3. URL <https://www.ncbi.nlm.nih.gov/pubmed/30361930>.
- K. Nguyen, F. Puppo, S. Roche, M. C. Gaillard, C. Chaix, A. Lagarde, M. Pierret, C. Vovan, S. Olschwang, E. Salort-Campana, S. Attarian, M. Bartoli, R. Bernard, F. Magdinier, and N. Levy. Molecular combing reveals complex 4q35 rearrangements in facioscapulohumeral dystrophy. *Hum Mutat*, 2017. ISSN 1098-1004 (Electronic) 1059-7794 (Linking). doi: 10.1002/humu.23304. URL <https://www.ncbi.nlm.nih.gov/pubmed/28744936>.
- S. Sacconi, L. Salviati, and C. Desnuelle. Facioscapulohumeral muscular dystrophy. *Biochim Biophys Acta*, 1852(4):607–14, 2015. ISSN 0006-3002 (Print) 0006-3002 (Linking). doi: 10.1016/j.bbadis.2014.05.021. URL <https://www.ncbi.nlm.nih.gov/pubmed/24882751>.
- S. Sacconi, R. J. Lemmers, J. Balog, P. J. van der Vliet, P. Lahaut, M. P. van Nieuwenhuizen, K. R. Straasheijm, R. D. Debipersad, M. Vos-Versteeg, L. Salviati, A. Casarin, E. Pegoraro, R. Tawil, E. Bakker, S. J. Tapscott, C. Desnuelle, and S. M. van der Maarel. The fshd2 gene smchd1 is a modifier of disease severity in families affected by fshd1. *Am J Hum Genet*, 93(4):744–51, 2013. ISSN 1537-6605 (Electronic) 0002-9297 (Linking). doi: 10.1016/j.ajhg.2013.08.004. URL <https://www.ncbi.nlm.nih.gov/pubmed/24075187>.
- A. D. Gurzau, K. Chen, S. Xue, W. Dai, I. S. Lucet, T. T. N. Ly, B. Reversade, M. E. Blewitt, and J. M. Murphy. Fshd2- and bams-associated mutations confer opposing effects on smchd1 function. *J Biol Chem*, 2018. ISSN 1083-351X (Electronic) 0021-9258 (Linking). doi: 10.1074/jbc.RA118.003104. URL <https://www.ncbi.nlm.nih.gov/pubmed/29748383>.

- N. Jansz, K. Chen, J. M. Murphy, and M. E. Blewitt. The epigenetic regulator *smchd1* in development and disease. *Trends Genet*, 33(4):233–243, 2017. ISSN 0168-9525 (Print) 0168-9525 (Linking). doi: 10.1016/j.tig.2017.01.007. URL <https://www.ncbi.nlm.nih.gov/pubmed/28222895>.
- K. U. Wagner, R. J. Wall, L. St-Onge, P. Gruss, A. Wynshaw-Boris, L. Garrett, M. Li, P. A. Furth, and L. Hennighausen. Cre-mediated gene deletion in the mammary gland. *Nucleic Acids Research*, 25(21):4323–30, 1997. ISSN 0305-1048 (Print) 0305-1048 (Linking). doi: 10.1093/nar/25.21.4323. URL <https://www.ncbi.nlm.nih.gov/pubmed/9336464>.
- J. C. de Greef, Y. D. Krom, B. den Hamer, L. Snider, Y. Hiramuki, R. F. P. van den Akker, K. Breslin, M. Pakusch, D. C. F. Salvatori, B. Slutter, R. Tawil, M. E. Blewitt, S. J. Tapscott, and S. M. van der Maarel. *Smchd1* haploinsufficiency exacerbates the phenotype of a transgenic FSHD1 mouse model. *Human Molecular Genetics*, 27(4):716–731, 2018. ISSN 1460-2083 (Electronic) 0964-6906 (Linking). doi: 10.1093/hmg/ddx437. URL <https://www.ncbi.nlm.nih.gov/pubmed/29281018>.
- Mark Lewandoski, Karen Montzka Wassarman, and Gail R. Martin. Zp3-cre, a transgenic mouse line for the activation or inactivation of loxp-flanked target genes specifically in the female germ line. *Current Biology*, 7(2):148–151, 1997. ISSN 09609822. doi: 10.1016/s0960-9822(06)00059-5.
- H. S. Leong, K. L. Chen, Y. F. Hu, S. Lee, J. Corbin, M. Pakusch, J. M. Murphy, I. J. Majewski, G. K. Smyth, W. S. Alexander, D. J. Hilton, and M. E. Blewitt. Epigenetic regulator *Smchd1* functions as a tumor suppressor. *Cancer Research*, 73(5):1591–1599, 2013. ISSN 0008-5472. doi: 10.1158/0008-5472.Can-12-3019. URL [GotoISI://WOS:000315741500015](https://www.ncbi.nlm.nih.gov/pubmed/23711115).
- A. K. Ludwig, K. De Miroschedji, T. R. Doeppner, V. Borger, J. Ruesing, V. Rebmann, S. Durst, S. Jansen, M. Bremer, E. Behrmann, B. B. Singer, H. Jastrow, J. D. Kuhlmann, F. El Magraoui, H. E. Meyer, D. M. Hermann,

- B. Opalka, S. Raunser, M. Eppe, P. A. Horn, and B. Giebel. Precipitation with polyethylene glycol followed by washing and pelleting by ultracentrifugation enriches extracellular vesicles from tissue culture supernatants in small and large scales. *J Extracell Vesicles*, 7(1):1528109, 2018. ISSN 2001-3078 (Print) 2001-3078 (Linking). doi: 10.1080/20013078.2018.1528109. URL <https://www.ncbi.nlm.nih.gov/pubmed/30357008>.
- Iromi Wanigasuriya, Quentin Gouil, Sarah A. Kinkel, Andrés Tapia del Fierro, Tamara Beck, Ellise E. A. Roper, Kelsey Breslin, Jessica Stringer, Karla Hutt, Heather J. Lee, Andrew Keniry, Matthew E. Ritchie, and Marnie E. Blewitt. Smchd1 is a maternal effect gene required for autosomal imprinting. 2020. doi: 10.1101/2020.01.20.913376. URL <https://www.biorxiv.org/content/10.1101/2020.01.20.913376v2>.
- M. Borensztein, L. Syx, K. Ancelin, P. Diabangouaya, C. Picard, T. Liu, J. B. Liang, I. Vassilev, R. Galupa, N. Servant, E. Barillot, A. Surani, C. J. Chen, and E. Heard. Xist-dependent imprinted X inactivation and the early developmental consequences of its failure. *Nature Structural & Molecular Biology*, 24(3):226–233, 2017. ISSN 1545-9985 (Electronic) 1545-9985 (Linking). doi: 10.1038/nsmb.3365. URL <https://www.ncbi.nlm.nih.gov/pubmed/28134930>.
- J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J. Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, and A. Cardona. Fiji: an open-source platform for biological-image analysis. *Nature Methods*, 9(7):676–82, 2012. ISSN 1548-7105 (Electronic) 1548-7091 (Linking). doi: 10.1038/nmeth.2019. URL <https://www.ncbi.nlm.nih.gov/pubmed/22743772>.
- C. Harris, M. Cloutier, M. Trotter, M. Hinten, S. Gayen, Z. Du, W. Xie, and S. Kalantry. Conversion of random x-inactivation to imprinted x-inactivation by maternal prc2. *Elife*, 8, 2019. ISSN 2050-084X (Electronic) 2050-084X (Linking). doi: 10.7554/eLife.44258. URL <https://www.ncbi.nlm.nih.gov/pubmed/30938678>.

- M. J. Levesque, P. Ginart, Y. Wei, and A. Raj. Visualizing snvs to quantify allele-specific expression in single cells. *Nat Methods*, 10(9):865–7, 2013. ISSN 1548-7105 (Electronic) 1548-7091 (Linking). doi: 10.1038/nmeth.2589. URL <https://www.ncbi.nlm.nih.gov/pubmed/23913259>.
- Anton Wutz and Rudolf Jaenisch. A shift from reversible to irreversible x inactivation is triggered during es cell differentiation. *Molecular Cell*, 5(4):695–705, 2000. ISSN 10972765. doi: 10.1016/s1097-2765(00)80248-8.
- M. Hinten, E. Maclary, S. Gayen, C. Harris, and S. Kalantry. Visualizing long noncoding rnas on chromatin. *Methods Mol Biol*, 1402:147–164, 2016. ISSN 1940-6029 (Electronic) 1064-3745 (Linking). doi: 10.1007/978-1-4939-3378-5_12. URL <https://www.ncbi.nlm.nih.gov/pubmed/26721489>.
- B. C. Chen, W. R. Legant, K. Wang, L. Shao, D. E. Milkie, M. W. Davidson, C. Janetopoulos, X. S. Wu, 3rd Hammer, J. A., Z. Liu, B. P. English, Y. Mimori-Kiyosue, D. P. Romero, A. T. Ritter, J. Lippincott-Schwartz, L. Fritz-Laylin, R. D. Mullins, D. M. Mitchell, J. N. Bembenek, A. C. Reymann, R. Bohme, S. W. Grill, J. T. Wang, G. Seydoux, U. S. Tulu, D. P. Kiehart, and E. Betzig. Lattice light-sheet microscopy: imaging molecules to embryos at high spatiotemporal resolution. *Science*, 346(6208):1257998, 2014. ISSN 1095-9203 (Electronic) 0036-8075 (Linking). doi: 10.1126/science.1257998. URL <https://www.ncbi.nlm.nih.gov/pubmed/25342811>.
- K. A. Alexander, X. Wang, M. Shibata, A. G. Clark, and M. J. Garcia-Garcia. TRIM28 controls genomic imprinting through distinct mechanisms during and after early genome-wide reprogramming. *Cell Reports*, 13(6):1194–1205, 2015. ISSN 2211-1247 (Electronic). doi: 10.1016/j.celrep.2015.09.078. URL <https://www.ncbi.nlm.nih.gov/pubmed/26527006>.
- H. Hiura, Y. Obata, J. Komiyama, M. Shirai, and T. Kono. Oocyte growth-dependent progression of maternal imprinting in mice. *Genes Cells*, 11(4):353–61, 2006. ISSN 1356-9597 (Print) 1356-9597 (Linking). doi: 10.1111/j.1365-2443.2006.00943.x. URL <https://www.ncbi.nlm.nih.gov/pubmed/16611239>.

- U. Midic, K. A. Vincent, K. Wang, A. Lokken, A. L. Severance, A. Ralston, J. G. Knott, and K. E. Latham. Novel key roles for structural maintenance of chromosome flexible domain containing 1 (Smchd1) during preimplantation mouse development. *Molecular Reproductive Development*, 85(7):635–648, 2018. ISSN 1098-2795 (Electronic) 1040-452X (Linking). doi: 10.1002/mrd.23001. URL <https://www.ncbi.nlm.nih.gov/pubmed/29900695>.
- M. L. Ruebel, K. A. Vincent, P. Z. Schall, K. Wang, and K. E. Latham. SM-CHD1 terminates the first embryonic genome activation event in mouse two-cell embryos and contributes to a transcriptionally repressive state. *American Journal of Physiology and Cell Physiology*, 317(4):C655–C664, 2019. ISSN 1522-1563 (Electronic) 0363-6143 (Linking). doi: 10.1152/ajpcell.00116.2019. URL <https://www.ncbi.nlm.nih.gov/pubmed/31365290>.
- D. Andergassen, M. Muckenhuber, P. C. Bammer, T. M. Kulinski, H. C. Theussl, T. Shimizu, J. M. Penninger, F. M. Pauler, and Q. J. Hudson. The airn lncrna does not require any dna elements within its locus to silence distant imprinted genes. *PLoS Genet*, 15(7):e1008268, 2019. ISSN 1553-7404 (Electronic) 1553-7390 (Linking). doi: 10.1371/journal.pgen.1008268. URL <https://www.ncbi.nlm.nih.gov/pubmed/31329595>.
- X. Wang, P. D. Soloway, and A. G. Clark. A survey for novel imprinted genes in the mouse placenta by mrna-seq. *Genetics*, 189(1):109–22, 2011. ISSN 1943-2631 (Electronic) 0016-6731 (Linking). doi: 10.1534/genetics.111.130088. URL <https://www.ncbi.nlm.nih.gov/pubmed/21705755>.
- T. Nagano, J. A. Mitchell, L. A. Sanz, F. M. Pauler, A. C. Ferguson-Smith, R. Feil, and P. Fraser. The air noncoding rna epigenetically silences transcription by targeting g9a to chromatin. *Science*, 322(5908):1717–20, 2008. ISSN 1095-9203 (Electronic) 0036-8075 (Linking). doi: 10.1126/science.1163802. URL <https://www.ncbi.nlm.nih.gov/pubmed/18988810>.
- L. Redrup, M. R. Branco, E. R. Perdeaux, C. Krueger, A. Lewis, F. Santos, T. Nagano, B. S. Cobb, P. Fraser, and W. Reik. The long noncoding rna

- kenq1ot1 organises a lineage-specific nuclear domain for epigenetic gene silencing. *Development*, 136(4):525–30, 2009. ISSN 0950-1991 (Print) 0950-1991 (Linking). doi: 10.1242/dev.031328. URL <https://www.ncbi.nlm.nih.gov/pubmed/19144718>.
- Annabelle Lewis, Kohzoh Mitsuya, David Umlauf, Paul Smith, Wendy Dean, Joern Walter, Michael Higgins, Robert Feil, and Wolf Reik. Imprinting on distal chromosome 7 in the placenta involves repressive histone methylation independent of dna methylation. *Nature Genetics*, 36(12):1291–1295, 2004. ISSN 1546-1718. doi: 10.1038/ng1468. URL <https://doi.org/10.1038/ng1468>.
- David Umlauf, Yuji Goto, Ru Cao, Frédérique Cerqueira, Alexandre Wagschal, Yi Zhang, and Robert Feil. Imprinting along the kenq1 domain on mouse chromosome 7 involves repressive histone methylation and recruitment of polycomb group complexes. *Nature Genetics*, 36(12):1296–1300, 2004. ISSN 1546-1718. doi: 10.1038/ng1467. URL <https://doi.org/10.1038/ng1467>.
- I. Okamoto, C. Patrat, D. Thepot, N. Peynot, P. Fauque, N. Daniel, P. Diabangouaya, J. P. Wolf, J. P. Renard, V. Duranthon, and E. Heard. Eutherian mammals use diverse strategies to initiate x-chromosome inactivation during development. *Nature*, 472(7343):370–4, 2011. ISSN 1476-4687 (Electronic) 0028-0836 (Linking). doi: 10.1038/nature09872. URL <https://www.ncbi.nlm.nih.gov/pubmed/21471966>.
- T. Tada, Y. Obata, M. Tada, Y. Goto, N. Nakatsuji, S. Tan, T. Kono, and N. Takagi. Imprint switching for non-random x-chromosome inactivation during mouse oocyte growth. *Development*, 127(14):3101–5, 2000. ISSN 0950-1991 (Print) 0950-1991.
- H. Chiba, R. Hirasawa, M. Kaneda, Y. Amakawa, E. Li, T. Sado, and H. Sasaki. De novo dna methylation independent establishment of maternal imprint on x chromosome in mouse oocytes. *Genesis*, 46(12):spc one, 2008. ISSN 1526-968X (Electronic) 1526-954X (Linking). doi: 10.1002/dvg.20496. URL <https://www.ncbi.nlm.nih.gov/pubmed/19105216>.

- D. P. Norris, N. Brockdorff, and S. Rastan. Methylation status of cpg-rich islands on active and inactive mouse x chromosomes. *Mamm Genome*, 1(2):78–83, 1991. ISSN 0938-8990 (Print) 0938-8990 (Linking). doi: 10.1007/BF02443782. URL <https://www.ncbi.nlm.nih.gov/pubmed/1799791>.
- M. Grant, M. Zuccotti, and M. Monk. Methylation of cpg sites of two x-linked genes coincides with x-inactivation in the female mouse embryo but not in the germ line. *Nat Genet*, 2(2):161–6, 1992. ISSN 1061-4036 (Print) 1061-4036 (Linking). doi: 10.1038/ng1092-161. URL <https://www.ncbi.nlm.nih.gov/pubmed/1303266>.
- S. McGraw, C. C. Oakes, J. Martel, M. C. Cirio, P. de Zeeuw, W. Mak, C. Plass, M. S. Bartolomei, J. R. Chaillet, and J. M. Trasler. Loss of dnmt1o disrupts imprinted x chromosome inactivation and accentuates placental defects in females. *PLoS Genet*, 9(11):e1003873, 2013. ISSN 1553-7404 (Electronic) 1553-7390 (Linking). doi: 10.1371/journal.pgen.1003873. URL <https://www.ncbi.nlm.nih.gov/pubmed/24278026>.
- L. Prokopuk, J. M. Stringer, C. R. White, Rham Vossen, S. J. White, A. S. A. Cohen, W. T. Gibson, and P. S. Western. Loss of maternal eed results in postnatal overgrowth. *Clin Epigenetics*, 10(1):95, 2018. ISSN 1868-7083 (Electronic) 1868-7075 (Linking). doi: 10.1186/s13148-018-0526-8. URL <https://www.ncbi.nlm.nih.gov/pubmed/30005706>.
- S. Kalantry, S. Purushothaman, R. B. Bowen, J. Starmer, and T. Magnuson. Evidence of xist rna-independent initiation of mouse imprinted x-chromosome inactivation. *Nature*, 460(7255):647–51, 2009. ISSN 1476-4687 (Electronic) 0028-0836 (Linking). doi: 10.1038/nature08161. URL <https://www.ncbi.nlm.nih.gov/pubmed/19571810>.
- E. Abbe. Beiträge zur theorie des mikroskops und der mikroskopischen wahrnehmung. *Archiv für Mikroskopische Anatomie*, 9(1):413–468, 1873. ISSN 0176-7364. doi: 10.1007/bf02956173.

- T.D. Allen. *Introduction to Electron Microscopy for Biologists*. Elsevier Science, 2008. ISBN 9780080888163. URL <https://books.google.com.au/books?id=P3d1Pvou8wC>.
- Rachel Won. Eyes on super-resolution. *Nature Photonics*, 3(7):368–369, 2009. ISSN 1749-4885 1749-4893. doi: 10.1038/nphoton.2009.103.
- L. Schermelleh, R. Heintzmann, and H. Leonhardt. A guide to super-resolution fluorescence microscopy. *J Cell Biol*, 190(2):165–75, 2010. ISSN 1540-8140 (Electronic) 0021-9525 (Linking). doi: 10.1083/jcb.201002018. URL <https://www.ncbi.nlm.nih.gov/pubmed/20643879>.
- H. C. Ishikawa-Ankerhold, R. Ankerhold, and G. P. C. Drummen. Advanced fluorescence microscopy techniques-frap, flip, flap, fret and flim. *Molecules*, 17(4):4047–4132, 2012. ISSN 1420-3049. doi: 10.3390/molecules17044047. URL [<GotoISI>://WOS:000303309900034](http://WOS:000303309900034).
- D. Axelrod, D. E. Koppel, J. Schlessinger, E. Elson, and W. W. Webb. Mobility measurement by analysis of fluorescence photobleaching recovery kinetics. *Biophysical Journal*, 16(9):1055–1069, 1976. ISSN 00063495. doi: 10.1016/s0006-3495(76)85755-4.
- M. E. van Royen, P. Farla, K. A. Mattern, B. Geverts, J. Trapman, and A. B. Houtsmuller. Fluorescence recovery after photobleaching (frap) to study nuclear protein dynamics in living cells. *Methods Mol Biol*, 464:363–85, 2009. ISSN 1940-6029 (Electronic) 1064-3745 (Linking). doi: 10.1007/978-1-60327-461-6_20. URL <https://www.ncbi.nlm.nih.gov/pubmed/18951195>.
- M. Kang, M. Andreani, and A. K. Kenworthy. Validation of normalizations, scaling, and photofading corrections for frap data analysis. *PLoS One*, 10(5):e0127966, 2015. ISSN 1932-6203 (Electronic) 1932-6203 (Linking). doi: 10.1371/journal.pone.0127966. URL <https://www.ncbi.nlm.nih.gov/pubmed/26017223>.

- S. T. Hess, T. P. Girirajan, and M. D. Mason. Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. *Biophys J*, 91(11):4258–72, 2006. ISSN 0006-3495 (Print) 0006-3495 (Linking). doi: 10.1529/biophysj.106.091116. URL <https://www.ncbi.nlm.nih.gov/pubmed/16980368>.
- E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess. Imaging intracellular fluorescent proteins at nanometer resolution. *Science*, 313(5793):1642–5, 2006. ISSN 1095-9203 (Electronic) 0036-8075 (Linking). doi: 10.1126/science.1127344. URL <https://www.ncbi.nlm.nih.gov/pubmed/16902090>.
- M. Heilemann, S. van de Linde, M. Schuttpelz, R. Kasper, B. Seefeldt, A. Mukherjee, P. Tinnefeld, and M. Sauer. Subdiffraction-resolution fluorescence imaging with conventional fluorescent probes. *Angew Chem Int Ed Engl*, 47(33):6172–6, 2008. ISSN 1521-3773 (Electronic) 1433-7851 (Linking). doi: 10.1002/anie.200802376. URL <https://www.ncbi.nlm.nih.gov/pubmed/18646237>.
- D. R. Whelan and T. D. Bell. Image artifacts in single molecule localization microscopy: why optimization of sample preparation protocols matters. *Sci Rep*, 5:7924, 2015. doi: 10.1038/srep07924. URL <https://www.ncbi.nlm.nih.gov/pubmed/25603780>.
- A. Jimenez, K. Friedl, and C. Leterrier. About samples, giving examples: Optimized single molecule localization microscopy. *Methods*, 174:100–114, 2020. ISSN 1095-9130 (Electronic) 1046-2023 (Linking). doi: 10.1016/j.ymeth.2019.05.008. URL <https://www.ncbi.nlm.nih.gov/pubmed/31078795>.
- B. L. Sprague and J. G. McNally. Frap analysis of binding: proper and fitting. *Trends Cell Biol*, 15(2):84–91, 2005. ISSN 0962-8924 (Print) 0962-8924 (Linking). doi: 10.1016/j.tcb.2004.12.001. URL <https://www.ncbi.nlm.nih.gov/pubmed/15695095>.
- T. Aoto, N. Saitoh, Y. Sakamoto, S. Watanabe, and M. Nakao. Polycomb group protein-associated chromatin is reproduced in post-mitotic g1 phase and is required for s phase progression. *J Biol Chem*, 283(27):18905–15, 2008. ISSN

- 0021-9258 (Print) 0021-9258 (Linking). doi: 10.1074/jbc.M709322200. URL <https://www.ncbi.nlm.nih.gov/pubmed/18453536>.
- Alexandra Gurzau. Functional and structural characterisation of the epigenetic regulator, smchd1. unpublished thesis, 2020.
- J. D. P. Rhodes, J. H. I. Haarhuis, J. B. Grimm, B. D. Rowland, L. D. Lavis, and K. A. Nasmyth. Cohesin can remain associated with chromosomes during dna replication. *Cell Rep*, 20(12):2749–2755, 2017. ISSN 2211-1247 (Electronic). doi: 10.1016/j.celrep.2017.08.092. URL <https://www.ncbi.nlm.nih.gov/pubmed/28930671>.
- D. Virant, B. Traenkle, J. Maier, P. D. Kaiser, M. Bodenhofer, C. Schmees, I. Vojnovic, B. Pisak-Lukats, U. Endesfelder, and U. Rothbauer. A peptide tag-specific nanobody enables high-quality labeling for dstorm imaging. *Nat Commun*, 9(1):930, 2018. ISSN 2041-1723 (Electronic) 2041-1723 (Linking). doi: 10.1038/s41467-018-03191-2. URL <https://www.ncbi.nlm.nih.gov/pubmed/29500346>.
- S. Schmid and T. Hugel. Controlling protein function by fine-tuning conformational flexibility. *Elife*, 9, 2020. ISSN 2050-084X (Electronic) 2050-084X (Linking). doi: 10.7554/eLife.57180. URL <https://www.ncbi.nlm.nih.gov/pubmed/32697684>.
- J. A. West, M. Mito, S. Kurosaka, T. Takumi, C. Tanegashima, T. Chujo, K. Yanaka, R. E. Kingston, T. Hirose, C. Bond, A. Fox, and S. Nakagawa. Structural, super-resolution microscopy analysis of paraspeckle nuclear body organization. *J Cell Biol*, 214(7):817–30, 2016. ISSN 1540-8140 (Electronic) 0021-9525 (Linking). doi: 10.1083/jcb.201601071. URL <https://www.ncbi.nlm.nih.gov/pubmed/27646274>.
- K. McDole, L. Guignard, F. Amat, A. Berger, G. Malandain, L. A. Royer, S. C. Turaga, K. Branson, and P. J. Keller. In toto imaging and reconstruction of

post-implantation mouse development at the single-cell level. *Cell*, 175(3):859–876 e33, 2018. ISSN 1097-4172 (Electronic) 0092-8674 (Linking). doi: 10.1016/j.cell.2018.09.031. URL <https://www.ncbi.nlm.nih.gov/pubmed/30318151>.