Deciphering the role of BORIS/CTCFL in melanoma

by

Sanne Marlijn Janssen

Faculty of Medicine
Department of Pathology
McGill University
Montréal, Quebec, Canada
December 2018

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy

“Our greatest weakness lies in giving up.
The most certain way to succeed is always to try just one more time.”
Thomas A. Edison

“De Aanhouder Wint”

TABLE OF CONTENTS

ABSTRACT

Brother of Regulator of Imprinted Sites (BORIS) is a DNA binding protein that is expressed in the testis and becomes aberrantly expressed in different types of cancer, hence the designation of BORIS as a cancer testis antigen. Aberrant BORIS expression in cancer is linked to alterations in gene expression and epigenetic modifications that impact key biological processes. In melanoma, BORIS expression is higher in metastasis compared to primary tumors, indicating a role for BORIS in melanoma progression. Currently, BORIS functions in melanoma remain to be elucidated. Here we set out to investigate the effect of BORIS expression on biological processes involved in melanoma progression.

In the first part of this thesis we extend previous work from the laboratory that demonstrated altered gene expression from the X-chromosome upon modified BORIS expression. This is important, as differences in gene expression from the X-chromosome likely contribute to the superior survival of female melanoma patients compared to males. By exploiting single nucleotide polymorphisms in genes located on the X-chromosome we reveal for the first time that ectopic BORIS expression can alter gene expression in an allele-specific manner. More specifically, we show that expression of *ZBTB33*, a transcriptional regulator involved in cancer cell migration and invasion, is increased from the inactive X-chromosome upon BORIS expression.

To address BORIS' involvement in biological processes and pathways in melanoma, we investigated BORIS-induced transcriptional changes and BORIS-interacting proteins using RNA-sequencing and mass spectrometry. We demonstrate that ectopic BORIS expression in melanoma cells promotes phenotype switching from a proliferative to invasive state at the transcriptional and phenotypic level. Using *in silico* analysis, we identify potential transcriptional regulators that contribute to this switch. Both the BORIS-induced gene expression changes and newly identified BORIS protein partners point to a role for BORIS in the DNA damage response. This finding is supported by our observations that BORIS increases DNA damage, activates the DNA damage response, enhances the sensitivity to DNA damage, and slows the repair process, suggesting that BORIS promotes genomic instability.

In the final part of this thesis, we used chromatin immunoprecipitation, bisulfite sequencing and the assay for transposase accessible chromatin (ATAC) to elucidate mechanisms underlying the BORIS-altered gene expression. We demonstrate for six genes that BORIS-induced transcriptional changes are associated with BORIS-DNA binding in the promoter region. For one of these genes, *PAX6*, we observe that this is not linked to DNA demethylation at the BORIS binding site. Furthermore, we successfully generated ATAC-sequencing data that will be used to assess the effect of BORIS expression on chromatin accessibility.

Overall, the work presented in this thesis brings forth new insight into the effect of BORIS expression in melanoma and provides evidence for previously unknown roles of BORIS in melanoma progression.

RÉSUMÉ

La molécule BORIS (Brother of Regulator of Imprinted Sites) est une protéine de liaison à l'ADN qui est présente dans les testicules mais qui est exprimée d'une manière aberrante dans plusieurs types de cancer, d'où la désignation de BORIS comme un antigène du cancer testiculaire. La présence aberrante de BORIS dans certains cancers altère l'expression des gènes et engendre des modifications épigénétiques qui impactent des processus biologiques clés. Dans le cas des mélanomes, l'expression de BORIS est plus élevée dans les métastases que dans la tumeur primaire, ce qui indique que BORIS joue un rôle dans la progression des mélanomes. Actuellement, la fonction de BORIS dans les mélanomes reste à élucider. Ici, nous étudions les effets de l'expression de BORIS sur les processus biologiques de la progression des mélanomes.

Dans la première partie de cette thèse, nous utilisons des résultats antérieurs de notre laboratoire, qui avaient démontré que des modifications dans l'expression de BORIS altérent l'expression génique du chromosome X. Ceci est important, car les différences d'expression génique du chromosome X contribuent probablement à la survie supérieure des femmes atteintes de mélanome comparativement aux hommes. En exploitant des polymorphismes mononucléotidiques dans des gènes localisés sur le chromosome X, nous révélons pour la première fois que l'expression ectopique de BORIS peut altérer l'expression génique d'une manière allèle-spécifique. Plus précisément, nous démontrons que l'expression de *ZBTB33*, un régulateur transcriptionnel impliqué dans la migration et l'invasion de cellules cancéreuses, est augmentée lors de l'expression de BORIS à partir de l'allèle situè sur le chromosome X inactif.

Afin de cibler l'implication de BORIS dans les processus et les voies biologiques des mélanomes, nous avons évalué les changements transcriptionnels induits par BORIS ainsi que les protéines interagissant avec BORIS en utilisant le séquençage d'ARN et la spectrométrie de masse. Nous démontrons que l'expression ectopique de BORIS dans les cellules de mélanome favorise le passage d'un état prolifératif à un état invasif au niveau transcriptionnel et phénotypique. Grâce à une analyse *in silico,* nous identifions de potentiels régulateurs transcriptionnels qui contribuent à ce passage. Les changements d'expression génique induits par BORIS ainsi que par ses partenaires protéiques nouvellement identifiés suggèrent un rôle de

BORIS dans la réponse aux dommages de l'ADN. Ceci est étayé par nos observations que BORIS augmente les dommages à l'ADN, active la réponse aux dommages de l'ADN, rehausse la sensibilité aux dommages de l'ADN et ralentit le processus de réparation, ce qui suggère que BORIS favorise l'instabilité génomique.

Dans la dernière partie de cette thèse, nous utilisons l'immunoprécipitation de la chromatine, le séquençage au bisulfite et l'étude de la chromatine accessible aux transposases (ATAC) pour élucider les mécanismes sous-jacents à l'expression génique altérée par BORIS. Nous démontrons que pour six gènes, les changements transcriptionnels induits par BORIS sont associés à sa liaison à l'ADN au niveau de la région promotrice. Pour un de ces gènes, *PAX6*, nous observons que son expression aberrante n'est pas liée à la déméthylation d'ADN au site de liaison de l'ADN. De plus, nous générons des données de séquençage d'ATAC qui seront utilisées pour évaluer les effets de l'expression de BORIS sur l'accessibilité à la chromatine.

ACKNOWLEDGEMENTS

First and foremost I want to thank my supervisor, Dr. Alan Spatz, for giving me the opportunity to pursue my research career in his laboratory. You provided me with a supportive environment for my PhD studies and the freedom to develop my own research ideas. It was a steep learning curve, but it allowed me to grow as a researcher and personally. In addition, I want to thank Dr. Léon van Kempen for sharing his scientific knowledge, his help with shaping the projects and providing valuable advice during the years of co-supervision. I also want to thank the Department of Pathology for their help and support during my studies.

I wish to thank the members of my advisory committee, Dr. Michael Witcher and Dr. Jason Karamchandani, for their guidance and suggestions throughout my PhD.

I want to thank our collaborators at the National Institute for Biotechnology in the Negev in Israel, Dr. Eitan Rubin and Roy Moscona, for the RNA-seq and ATAC-seq data analysis, and the team of Dr. Borchers at the UVic Genome BC Proteomics Centre for their help with the mass spectrometry experiment.

This work would not have been possible without the help from all the wonderful current and past lab members. A big thanks to Mounib, Marie-Noël, Andreas, Qasim, Reem, Maria, Anke, Kishanda, Jonathan, Margaret, June, Jasmin, Fanny, Samuel L, Samuel O, Romain and Vincent, as well as the volunteers/summer students Mohammad, Farhana, Anna, Sophia, Meri, Greta and Vivienne. A special thanks to Mounib and Marie-Noël, you were there for all these years and helped me through the good and bad times. Besides the invaluable mentorship in the lab, you were also always there for help and advice regarding all aspects of life. To me, you are my "parents abroad" who helped me grow both personally and career-wise, something I will always be grateful for.

A HUGE thanks to my family! Moving across an ocean is never easy, but with the unconditional support of my parents and brother I managed to build a wonderful life here. Thank you for giving me the opportunity to chase my dreams! Pap, mam en Bart, kei bedankt!

Andy, no words can truly express how thankful I am for your love and patience, especially during the difficult times and late nights. Coming home to the smell of delicious food is what made working late much more bearable! You have not only helped me get through the tough parts of my PhD, but also supported me during my knee surgery and the difficulties of the recovery afterwards. For that, I will always be grateful! I also want to thank you for proofreading abstracts and applications, bringing me to Camp, teaching me how to ski, and taking me on beautiful adventures. Andy, thank you so much for everything you have done for me. In addition, I want to thank your parents for their support and advice.

I want to thank the lovely people I met at the LDI for a great work environment, the always interesting lunch discussions as well as fun times outside the lab: Jasmin, Fred, Chris (an additional thanks for helping me out with flow experiments and organizing so many fun events), Edgar, Fernando, Steph, Maria, Tamiko, Barbara, Mona, Andre, René, Vincent, Jackie, Saina, James, Claudia, Luca (thanks for driving me to the ER after busting my knee!), Elaheh, Martin, Zoya, Mike, Alicia, Jesse, and Arjuna.

A big thanks to my friends and family across the world!
Op Nederlandse bodem, vele malen dank aan iedereen die mij over de afgelopen jaren niet is vergeten en in allerlei vormen steun hebben gestuurd voor beide de studie en het (vrij) koude weer: de families Janssen, Wijnhoven en Akkermans, de Meiden (en hun mannen), de Matties, de Schaijkse crew, Vicky, Rebecca, Lotte en Sylvie (bedankt voor al jou advies).
I want to thank the people I met in Canada for the fun times and outdoors adventures: Fernanda, Oli, Balqis, Eimear, Brandy, Arzoo, Bruno, Felix, Nathasha, Jake, Steve, Doug, Rene, Pierre-Marie, Hawley, Ayesha, Mark, Ric, Jida, Nathaniel, Dea, Augustin, Thais and the kind people from badminton. A special thanks to Eimear and Jida for translating the abstract into French.

I was fortunate to receive financial support for my PhD project from the MICRTP and Banque Nationale Scholarships as well as travel support to present my work at national and international meetings from the Melanoma Summit, EORTC-melanoma group, McGill Pathology department and CEEHRC Network.

PREFACE

This doctoral thesis is written in traditional format in accordance with the *Guidelines for Thesis Preparation* from McGill University. It consists of the following four chapters:

Chapter 1 provides a general overview of the literature and introduces concepts and ideas that led to the doctoral thesis rationale and aims.

Chapter 2 contains a detailed description of the materials and methods used to obtain the results described in Chapter 3.

Chapter 3 is divided into four subsections (3.1-3.4), each with its own background and objectives, a detailed description of the results, and the figures and tables that support the reported results. **Chapter 3.2** contains data that is included in a manuscript in preparation.

The first section of **chapter 4** consists of a summary and final model based on the obtained results. The following sections provide a comprehensive discussion, integrating all the results and suggesting possible future avenues to explore. The last section of chapter 4 contains the final conclusion.

The references for all chapters are listed at the end of the thesis

AUTHOR CONTRIBUTIONS

The projects included in this thesis were conceptualized by me, Sanne Janssen, together with my supervisor Dr. Alan Spatz, and Dr. Léon van Kempen. This thesis was written by me with the help of Dr. Alan Spatz, Dr. Léon van Kempen, Dr. Andreas Papadakis and Dr. Mounib Elchebly. The experiments described in this thesis were designed, coordinated, performed and analyzed by me with the help of Dr. Alan Spatz, Dr. Léon van Kempen and Dr. Mounib Elchebly, except when stated otherwise below. The manuscript derived from chapter 3.2 was written by me, with the help from my supervisor and the co-authors.

Chapter 3.1 - In this chapter, DNA-FISH and RNA-FISH was performed by Dr. Anke van Rijk. The expression plasmids HK-EV, HK-BORIS and mBORIS-3xF were cloned by Dr. Mounib Elchebly. Experiments related to the identification of single nucleotide polymorphisms in the MM057 cell line were performed with the assistance of Mohammad Amraei. Handling of the C57BL/6J and FVB/NJ mice was done by members of the Koromilas laboratory. Mouse embryonic fibroblast cell lines were established with the help of Dr. Mounib Elchebly. Sophia Borchers assisted me with genotyping the mouse embryonic fibroblasts and identification of the *Xist* SNP in the clones. The immunoblots for mouse BORIS expression were performed by Dr. Fanny Beltzung and me.

Chapter 3.2 - In this chapter, Dr. Andreas Papadakis assisted with lentivirus production and infection of the MM057 cell line. The RNA-seq libraries and sequencing were performed by The Centre for Applied Genomics at the SickKids hospital in Toronto, Canada. RNA-seq data analysis (differential gene expression and gene set enrichment analysis) was performed by Roy Moscona under the supervision of Dr. Eitan Rubin. Dr. Mounib Elchebly assisted with the migration and invasion assays.

Chapter 3.3 - In this chapter, mass spectrometry samples were prepared by Dr. Mounib Elchebly. Mass spectrometry and analysis were done at the University of Victoria Proteomics Centre with Dr. Christoph Borchers. Dr. Mounib Elchebly assisted with immunofluorescence for UBF. Flow cytometry ModFit LT™ analysis was done with the help of Christian Young. The

HaloTag® pull-down and Co-immunoprecipitation experiments were performed by Dr. Mounib Elchebly.

Chapter 3.4 – In this chapter, ChIP-qPCR was performed with the help of Dr. Mounib Elchebly. Sophia Borchers assisted with bisulfite conversion experiments. ATAC-seq data analysis was performed by Roy Moscona under the supervision of Dr. Eitan Rubin.

I have made contributions to a published abstract that was not included in this thesis:
S. Lagabrielle, **S. Janssen**, M. Elchebly, F. Blaudou, A. Spatz, L. van Kempen. (2015) PAX6 est un modulateur de l'activité transcriptionnelle du récepteur aux androgènes dans le cancer de prostate. Prog. Urol. 25(13):735.

ORIGINAL CONTRIBUTIONS TO KNOWLEDGE

1) We provide the first evidence that BORIS protein is expressed in metastatic melanoma cell lines and demonstrate that BORIS increases allele-specific expression of the X-chromosome-linked gene *ZBTB33* from the inactivated X-chromosome in female melanoma cells. In addition, we generated a mouse embryonic fibroblast cell model with a known X-chromosome inactivation status based on a single nucleotide polymorphism in *XIST*.

2) We were the first to perform RNA-seq on melanoma cells with altered BORIS expression and reveal large-scale differential gene expression. This prompted us to use the RNA-seq dataset for the identification of biological processes disrupted due to BORIS expression. We show enrichment of migration and invasion-related pathways and processes among differentially expressed genes, which led to the first report that BORIS expression contributes to a switch from a proliferative to invasive state in melanoma cells.

3) Using mass spectrometry, we were able to identify 79 new putative BORIS-interacting proteins. We were the first to provide evidence for BORIS-BORIS homodimer formation and to demonstrate an interaction between BORIS and members of the histone chaperone complex FACT. The BORIS-interacting proteins suggested a role for BORIS in DNA damage repair. We are the first to report increased DNA damage and activation of the DNA damage response upon BORIS expression in melanoma cells. In addition, we provide preliminary data that BORIS-expressing cells are more prone to DNA damage-inducing agents.

4) We were able to identify six new BORIS target genes that are upregulated upon BORIS expression and are bound by BORIS in their promoter region. For one of these genes, *PAX6*, we provide the first evidence that the BORIS binding site in the promoter region is not methylated in melanoma cells. In addition, we are the first to assess the effect of BORIS expression on chromatin accessibility in melanoma cells.

Research communications

Manuscript currently under review:

Janssen SM, Moscona R., Elchebly M., Papadakis AI, van Kempen LCL, Rubin E, Spatz A. BORIS/CTCFL promotes a switch from a proliferative towards an invasive phenotype in melanoma cells. (BioRxiv doi: 10.1101/560474)

Published abstract

Janssen SM.; van Kempen LC.; Elchebly M.; M'Boutchou M.; Spatz A. (2017) Impact of the Epigenetic Modulator BORIS on Sensitivity of Melanoma Cells to UV-induced DNA Damage. Journal of Clinical Oncology, DOI: 10.1200/JCO.2017.35.15_suppl.e21077 ASCO Annual Meeting, Chicago, IL, USA

Oral presentations

- Lady Davis Institute Elevator Talks Montréal, QC, Canada. BORIS-mediated Invasion of Melanoma Cells. Janssen S.M.; Moscona R.; Elchebly M.; van Kempen L.C.; Rubin E.; Spatz A. March 2019. **Oral presentation prize**
- Lady Davis Institute Cancer Seminar, Montréal, QC, Canada. BORIS/CTCFL Mediates Melanoma Phenotype Switching. Janssen S.M.; Moscona R.; Elchebly M.; van Kempen L.C.; Rubin E.; Spatz A. Jan 2019
- Finlayson Research Day, Montréal, QC, Canada. The Role of BORIS in DNA Damage Sensitivity. Janssen S.M.; van Kempen L.C.; Elchebly M.; M'Boutchou M.; Spatz A. June 2017. **Oral presentation prize**
- EORTC - EGAM meeting, Brussels, Belgium. Impact of the Epigenetic Modulator BORIS on Sensitivity of Melanoma Cells to UV-Induced DNA Damage. Janssen S.M.; van Kempen L.C.; Elchebly M.; M'Boutchou M.; Spatz A. March 2017
- Lady Davis Institute Cancer Seminar, Montréal, QC, Canada. Examining the Role of the Epigenetic Factor BORIS in the Nucleolus of Melanoma Cells. Janssen S.M.; van Kempen L.C.; Elchebly M.; Spatz A. March 2017

- 2nd Annual Summit on Practical and Emerging Trends in Melanoma, Pasadena, CA, USA. BORIS is an epigenetic modifier of melanoma virulence. Janssen S.M.; van Kempen L.C.; Elchebly M.; Ohayon S.; Spatz A. Sept. 2015
- Finlayson Research Day, Montréal, QC, Canada. Chromosome-wide and single gene expression regulation by the epigenetic factor BORIS in cancer. Janssen S.M.; Elchebly M.; van Rijk A.; Ohayon S.; Spatz A.; van Kempen L.C. June 2015
- Joint Histoseminar - Université de Montréal and McGill University, Montréal, QC, Canada. The role of BORIS as an epigenetic modulator of gene expression in cancer. Janssen S.M.; Elchebly M.; Ohayon S.; van Rijk A.; Spatz A.; van Kempen L.C. April 2015
- Lady Davis Institute Cancer Seminar, Montréal, QC, Canada. BORIS is a critical pro-tumorigenic regulator of the epigenetic landscape. Janssen S.M.; Ohayon S.; van Rijk A.; Elchebly M.; Spatz A.; van Kempen L.C. April 2014

Poster presentations

- Keystone Symposia: Epigenetics and Human Disease, Banff, AB, Canada. BORIS/CTCFL Promotes a Transcriptional Switch from a Proliferative to Invasive State in Melanoma Cells. Janssen S.M.; Moscona R.; Elchebly M.; van Kempen L.C.; Rubin E.; Spatz A. March 2019
- 5th Canadian Conference on Epigenetics, Estérel, QC, Canada. BORIS Promotes a Switch from a Proliferative to Invasive Gene Signature in Melanoma Cells. Janssen S.M.; Moscona R.; Elchebly M.; van Kempen L.C.; Rubin E.; Spatz A. Sept. 2018. **Poster presentation prize**
- Finlayson Research Day, Montréal, QC, Canada. Investigating the Effect of BORIS/CTCFL Expression on DNA Damage and Repair in Melanoma. Janssen S.M.; Elchebly M.; van Kempen L.C.; Spatz A. June 2018
- Lady Davis Institute Scientific Retreat, Montréal, QC, Canada. The Role of the Epigenetic Factor BORIS in the DNA Damage Response. Janssen S.M.; Elchebly M.; van Kempen L.C.; Spatz A. May 2018
- Keystone Symposia: Chromatin Architecture and Chromosome Organization, Whistler, BC, Canada. Deciphering the Role of BORIS in the DNA Damage Response. Janssen S.M.; van Kempen L.C.; Elchebly M.; Spatz A. March 2018

- Annual McGill Biomedical Graduate Conference, Montréal, QC, Canada. The Effect of BORIS on DNA Damage and Repair in Melanoma. Janssen S.M.; Elchebly M.; van Kempen L.C.; Spatz A. March 2018. **Poster presentation prize**
- Genomics Exchange Community (geXc) Symposium, Montréal, QC, Canada. Deciphering the Role of the Epigenetic Factor BORIS in DNA Damage Sensitivity in Cancer. Janssen S.M.; van Kempen L.C.; Elchebly M.; Spatz A. Dec. 2017. **Poster presentation prize**
- EACR Symposium Protecting the Code: Epigenetic Impacts on Genome Stability, Berlin, Germany. Impact of the Epigenetic Factor BORIS on DNA Damage Sensitivity in Cancer. Janssen S.M.; van Kempen L.C.; Elchebly M.; Spatz A. Oct. 2017
- Lady Davis Institute Scientific Retreat, Montréal, QC, Canada. Examining the Effect of BORIS on DNA Damage Sensitivity in Cancer. Janssen S.M.; van Kempen L.C.; Elchebly M.; M'Boutchou M.; Spatz A. May 2017
- Annual McGill Biomedical Graduate Conference, Montréal, QC, Canada. Deciphering the Role of the Epigenetic Factor BORIS in DNA Damage Sensitivity of Melanoma Cells. Janssen S.M.; van Kempen L.C.; Elchebly M.; M'Boutchou M.; Spatz A. March 2017
- Lady Davis Institute Scientific Retreat, Montréal, QC, Canada. Analysis of aberrant gene reactivation from the inactive X-chromosome by the epigenetic modulator BORIS. Janssen S.M.; van Kempen L.C.; Elchebly M.; M'Boutchou M.; Spatz A. June 2016.
- Finlayson Research Day, Montréal, QC, Canada. Elucidating the role of the epigenetic regulator BORIS in aberrant gene reactivation from the inactive X-chromosome. Janssen S.M.; van Kempen L.C.; Elchebly M.; M'Boutchou M.; Spatz A. June 2016
- Lady Davis Institute Scientific Retreat, Montréal, QC, Canada. BORIS, an epigenetic regulator of PAX6 expression in cancer. Janssen S.M.; Elchebly M.; Spatz A.; van Kempen L.C. June 2015
- Finlayson Research Day, Montréal, QC, Canada. BORIS, an epigenetic modulator in cancer, regulates gene expression from the X-chromosome in melanoma. Janssen S.M.; Ohayon S.; Elchebly M.; van Rijk A.; Spatz A.; van Kempen L.C. June 2014. **Poster presentation prize**
- Lady Davis Institute Scientific Retreat, Montréal, QC, Canada. Gene expression regulation by the epigenetic modulator BORIS in cancer. Janssen S.M.; Ohayon S.; van Rijk A.; Elchebly M.; van Kempen L.C.: Spatz A. June 2014

LIST OF ABBREVIATIONS

1x CTSes	Individual CTSes
2xCTSes	Clustered CTSes
5-aza-dC	5-aza-2'deoxy-cytidine
AP-1	Activator Protein 1
ATAC-seq	Assay for transposase accessible chromatin-sequencing
ATM	Ataxia telangiectasia-mutated
BORIS	Brother of regulator of imprinted sites
BORneg/pos	MM057 infected with BORIS-6xH without/with dox
BRAF	B-Raf proto-oncogene
BRDT	Bromodomain testis associated
CDH1	Epithelial (E)-cadherin
CDH2	Neuronal (N)-cadherin
CDKN2A / CDK	Cyclin dependent kinase inhibitor 2A
cDNA	Complementary DNA
ChIA-PET	Chromatin interaction analysis by paired-end tags
ChIP / IP	Chromatin immunoprecipitation
CPDB	Consensuspath database
CpG	Cytosine-phosphate-Guanine
CPM	Counts per million
CSD	Chronically sun damaged
CT45	Cancer testis family 45
CTAs	Cancer testis antigens
CTCF	CCCTC binding factor
CTCF&BORIS sites	2x CTCF target sites occupied by both BORIS and CTCF
CTCFL	CTCF-Like
CTSes	CTCF target sites
DAPI	4',6-diamidino-2-phenylindole
DDR	DNA damage response
DEGs	Differentially expressed genes

DNA	Deoxyribonucleic acid
DNase-seq	Deoxyribonuclease I-hypersensitivity site-sequencing
DNMT	DNA Methyltransferase
Dox	Doxycycline
DSB	Double strand break
ECM	Extracellular matrix
eIF	Eukaryotic initiation factor
EMT / EMT-TFs	Epithelial-mesenchymal transition transcription factors
ENCODE	Encyclopedia of DNA elements
EV-3xF	pInducer20-empty vector-three C-terminal FLAG®-tags
EV-6xH	pInducer20-empty vector-six C-terminal His-tags
EVneg/pos	MM057 infected with EV-6xH without/with dox
FACS	Fluorescence activated cell sorting
FACT	Facilitates Chromatin Transcription
FAIRE-seq	Formaldehyde-assisted isolation of regulatory elements-sequencing
FDR	False discovery rate
FISH	Fluorescent *in situ* hybridization
GAGE / XAGE	G / X antigen
gDNA	Genomic DNA
GFP	Green fluorescent protein
GIF-BORIS	pLenti-GFP-IRES-Flag-tag-full-length human BORIS
GIF-CTCF	pLenti-GFP-IRES-Flag-tag-full-length human CTCF
GIF-EV	pLenti-GFP-IRES-Flag-tag-empty vector
GSEA	Gene set enrichment analysis
hBORIS	Human BORIS
hBORIS-3xF	pInducer20-full-length human BORIS-three C-terminal FLAG®-tags
hBORIS-6xH	pInducer20 full-length human BORIS-six C-terminal His-tags
HDAC	Histone deacetylase
HK-BORIS	pLenti-His-tag-Katushka-full-length human BORIS
HK-EV	pLenti-His-tag-Katushka-empty vector

HT	HaloTag®
HT-BORIS	pHaloTag®-full-length human BORIS
HT-CTCF	pHaloTag®-full-length human CTCF
HT-EV	pHaloTag®-empty vector
HT-PD	HaloTag® pull-down
IB	Immunoblot
ICR	Imprinting control region
IF	Immunofluorescence
Igf2	Insulin growth factor 2
IRES	Internal ribosome entry site
log_2FC	log2 fold-change
MAGE / BAGE	(B) Melanoma antigen
MAPK	Mitogen-activated protein kinase
mBORIS	Mouse BORIS
mBORIS-3xF	pInducer20-full-length mouse BORIS-three C-terminal FLAG®-tags
mBORIS-6xH	pInducer20 full-length mouse BORIS-six C-terminal His-tags
MEFs	Mouse embryonic fibroblasts
MGI	Mouse genome informatics
MIST	Molecular interaction search tool
MITF	Microphthalmia-associated transcription factor
MMPs	Matrix-metalloproteinases
mRNA	Messenger RNA
MS	Mass spectrometry
mtDNA	mitochondrial DNA
MTT	3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
NES	Normalized enrichment score
NFATC2	Nuclear factor of activated T cells 2
NFIB	Nuclear factor I B
Non-X-CTAs	Non-X-chromosome CTA genes
NRAS	Neuroblastoma RAS viral oncogene homolog

NuRD	Nucleosome Remodeling and Deacetylation
NY-EOS-1	New York squamous cell carcinoma 1
PANTHER	Protein analysis through evolutionary relationship
PARP	poly ADP ribose polymerase
PAX6	Paired-box 6
PDL1	Programmed death-ligand 1
PEI	Polyethylenimine
PI	Propidium iodide
PRSS50	Serine protease 50
qPCR	Quantitative PCR
rDNA	Ribosomal DNA
REViGO	REduce and Visualize Gene Ontology
RGP	Radial-growth phase
RING	Really interesting new gene
RNA	Ribonucleic acid
RNA-seq	RNA-sequencing
rRNA	Ribosomal RNA
SAGE	Sarcoma antigen
shRNA	Short hairpin RNA
SILAC	Stable isotope labeling of amino acids in cell culture
siRNA	Small interfering RNA
SNPs	Single nucleotide polymorphisms
SPARC	Secreted protein acidic and rich cystein
SSC	Saline sodium citrate
SSRP1	Structure specific recognition protein 1
SSX	Synovial Sarcoma X
TAs	Tumor antigens
TCE	Total cell extract
TEAD	TEA domain
TGFB	Transforming growth factor beta
THBS2	Thrombospondin

TSA	Trichostatin A
TSS	Transcription start site
UBF	Upstream binding factor
UTR	Untranslated region
UV	Ultraviolet
VGP	Vertical-growth phase
WNT/Ca2+	Wnt/calcium pathway
X-chromosome-linked	X-linked
X-CTAs	X-chromosome CTA genes
Xa	Active X-chromosome
XCI	X-chromosome inactivation
Xi	Inactive X-chromosome
XIST	X-inactive specific transcript
ZEB	Zinc finger E-box-binding homeobox
ZF	Zinc finger
γH2AX	Gamma Histone 2A.X

LIST OF FIGURES

Chapter 1 - Literature review and introduction

Chapter 3.1 - BORIS Expression Alters X-linked Gene Expression in a XaXi Melanoma Cell Line

Chapter 3.2 - BORIS Expression Alters the Transcriptome of Melanoma Cells

Chapter 3.3 - Identification of BORIS Protein Partners

LIST OF TABLES

CHAPTER 1

-

LITERATURE REVIEW AND INTRODUCTION

1.1 CUTANEOUS MELANOMA

This section provides background on melanoma, outlines the steps involved in melanoma progression and describes the phenotype switching model.

1.1.1 Melanoma

Melanoma is a form of skin cancer that originates from melanin-producing cells called melanocytes. Melanocytes reside in the basal layer of the epidermis, hair follicles, mucosal surfaces, meninges and in the choridal layer of the eye, and are derived from neural crest stem cells [1]. Melanocytes are responsible for the production of melanin pigment, which they transfer to their neighboring keratinocytes that use it as a shield to protect from ultraviolet (UV) radiation [2]. Normally, cutaneous melanocytes divide very infrequently, less than twice a year [3]. However, upon certain environmental and genetic influences, cutaneous melanocytes can start to proliferate uncontrollably and undergo malignant transformation, giving rise to cutaneous melanoma [4]. Melanoma is among the most aggressive cancers due to its tendency to spread to vital parts of the body early on and is therefore the most deadly type of skin cancer. Yearly, 232,100 new cases of melanoma are diagnosed, resulting in about 55,000 deaths worldwide. The incidence of melanoma is rising, especially in fair-skinned populations. Melanoma incidence rates vary widely by country and the highest rates are reported in Australia and New Zealand [5]. In Canada about 4,000 males and 3,300 females are diagnosed with melanoma each year [6].

1.1.2 Risk factors for cutaneous melanoma

Known risk factors for melanoma include UV radiation via sun exposure and indoor tanning, pigmentation and the tendency to freckle (skin type), a high number of nevi (commonly known as skin moles), and a family history of melanoma [5, 7-9].

Ultraviolet radiation

The main risk factor for melanoma is exposure to UV radiation, mainly from sunlight. UV radiation can be divided into three wavelengths: UVA, UVB and UVC. Both UVA and UVB can pass the ozone layer and penetrate the skin. While UVA has a longer wavelength, it is less energetic than UVB and therefore less damaging to the skin. Extensive exposure to UVB is known to result in deoxyribonucleic acid (DNA) damage in the form of C>T transitions at

dipyrimidine sites, one of the melanoma UV signature mutations. Another UV signature mutation is the G>T transition as a result of UVA-induced oxidative damage [10]. These transitions can result in the formation of pyrimidine dimers via photochemical reactions. Two types of pyrimidine dimers exist, cyclobutane adducts and 6,4-photoproducts, that alter the double stranded DNA structure by partial unwinding and abnormal base-pairing. Up to a certain level, UV-induced lesions can be repaired by either nucleotide excision repair or the enzyme photolyase via photoreactivation. However, excessive and/or chronic UV radiation results in an accumulation of mistakes during DNA repair, which can result in melanoma. In addition, genetic defects in nucleotide excision repair are the cause for a rare human disease known as xeroderma pigmentosum that leads to a 2,000 fold higher incidence in melanoma [11].

Over the past decade, next generation sequencing studies looking at melanoma have clearly demonstrated the effect of UVB radiation on C>T transitions and revealed that melanoma has a particularly high mutational load [12-14]. The main genetic drivers of melanomas are B-Raf proto-oncogene (*BRAF*), neurofibromin 1, and neuroblastoma RAS viral oncogene homolog (*NRAS*). Interestingly, mutations in these genes are usually not UV signature mutations [10, 12]. Nevertheless, it is hypothesized that UV radiation does contribute to mutations in these genes. One possible mechanism is UV irradiation-induced production of reactive oxygen species formation that damages the DNA. Another proposed mechanism is the introduction of mutations by the error-prone DNA polymerase upon UV radiation [4].

Importantly, not all cutaneous melanomas found on sun-exposed skin are the same and these melanomas can be categorized into two distrinct subtypes, chronically sun damaged (CSD) melanomas and non-CSD melanomas. While, CSD melanomas arise in areas with the highest levels of sun exposure, have a high mutation burden and a late age of onset, non-CSD melanomas occur in younger individuals who have undergone intense, infrequent UV exposure. The different types of UV exposure in these two subtypes are associated with distinct mutations. Non-CSD melanomas mainly display the well-known $BRAF^{V600E}$ mutation and CSD melanomas are characterized by mutations in *NRAS*, neurofibrin 1, and the KIT proto-oncogene receptor tyrosine kinase as well as $BRAF^{nonV600E}$ mutations [4, 15].

Family history

About 10% of cutaneous melanomas affect family members, demonstrating the existence of melanoma-susceptibility genes. One gene that is often mutated within melanoma families is Cyclin dependent kinase inhibitor 2A (*CDKN2A*). Through alternative splicing of the *CDKN2A* transcript encodes for two tumor-suppressor proteins, $p14^{ARF}$ and $p16^{INK4A}$ that promote cell cycle progression when mutated. In addition, mutations in *CDK4*, Retinoblastoma 1 and Microphthalmia-associated transcription factor (*MITF*) were identified in families as high-risk alleles for melanoma development. Germline mutations in human Telomerase Reverse Transcriptase and Protection Of Telomeres 1 are also associated with familial melanoma [15-21].

1.1.3 Melanoma progression

The malignant transformation of normal melanocytes into melanoma is a complex, multi-step process that can evolve into different melanoma subtypes depending on the precursor lesion, gene mutations and stage of transformation [4]. Five distinct stages of melanoma have been described based on their location and stage of progression. These stages are: 1) common nevus, 2) dysplastic nevus, 3) radial-growth phase (RGP) melanoma, 4) vertical-growth phase (VGP) melanoma, and 5) metastatic melanoma, schematically depicted in (**Fig. 1**). The first step towards metastatic melanoma involves the proliferation of melanocytes, resulting in the formation of a nevus that is considered a benign tumor. When this nevus develops, sometimes morphologically abnormal cells are present, deeming the lesion a dysplastic nevus. This transformation is associated with clinical changes, including asymmetry, irregular border, change in color or larger diameter (ABCD rule [22]). Upon acquiring environmental, genetic and epigenetic changes, these dysplastic nevi can become a melanoma *in situ* (the RGP). At this stage the melanoma remains entirely confined to the epidermis and can still be surgically removed, resulting in a nearly 100% survival rate. During the VGP, cells invade the epidermis and dermis into the subcutaneous tissue from where they can spread to distant sites and establish metastasis. Of importance to note, metastatic melanoma does not only evolve from VGP, but can also originate from primary lesions [4, 23].

Based on these different steps melanomas are clinically classified in stages according to guidelines set by the American Joint Committee on Cancer [24]. Staging follows the TNM

staging system, which takes into account the tumor (T) thickness (Breslow-depth) and if the epidermis on top of the melanoma has ruptured (ulceration), the presence or absence as well as the number of regional lymph node metastasis (N), and the presence or absence and site of distant metastasis (M). Stage 0 represents melanoma *in situ*, meaning that the tumor is very small and confined to the upper layer of the epidermis. Stage I and II melanomas are still localized (no positive regional lymph nodes), though the tumor is larger and ulceration might be present. When a melanoma has passed stage II, the tumor is not considered localized anymore. At stage III a regional metastasis was identified in >1 lymph node. Further progression of disease by spreading of the tumor to distant metastatic sites represents stage IV melanoma (**Fig. 1**). For a thorough overview of the melanoma sub-classification, please refer to [24].

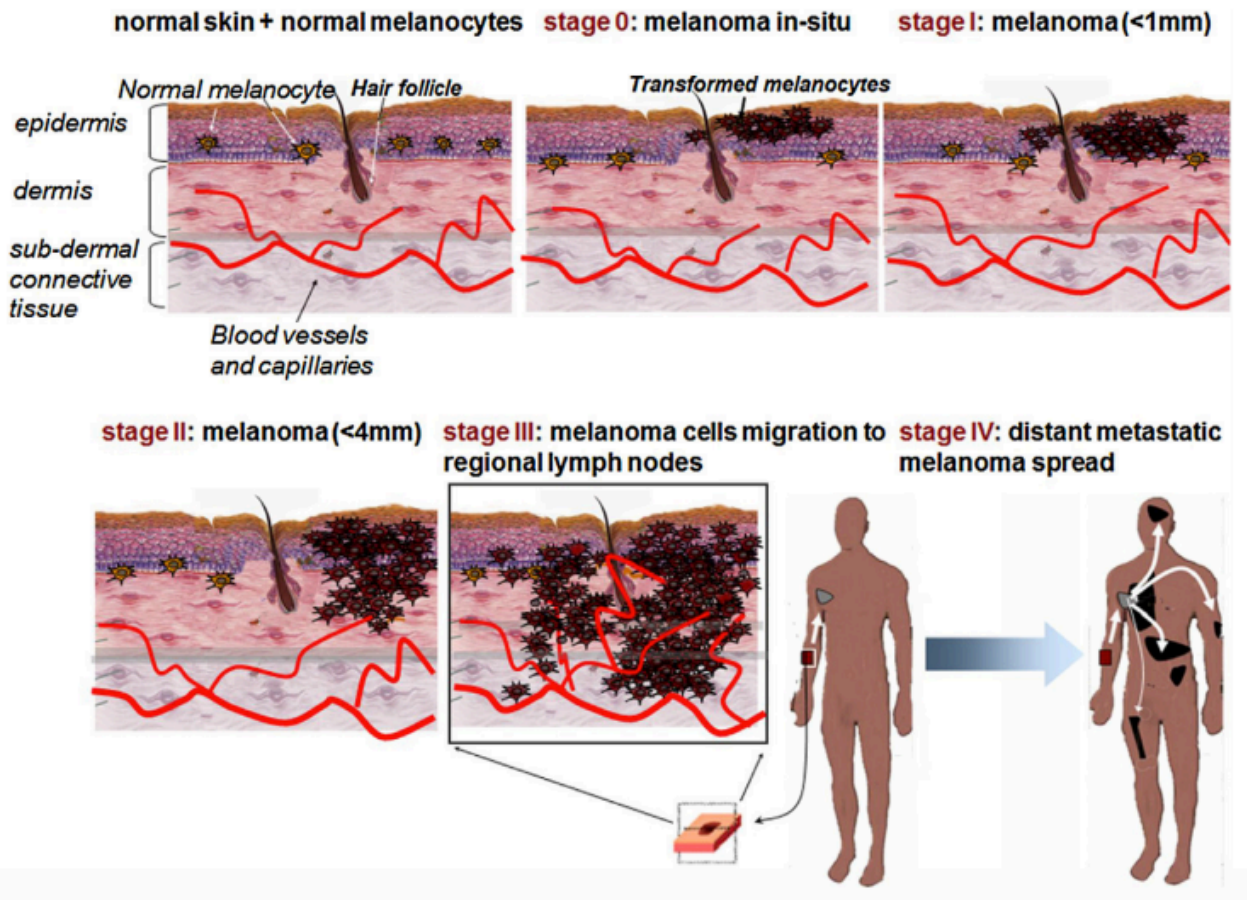


Figure 1. A schematic representation of melanoma progression with stages of disease.

The first step in melanoma progression involves the transformation and proliferation of melanocytes to a melanoma *in situ* (radial growth phase). The melanoma cells then progress to the vertical growth phase during which the cells invade the epidermis and dermis into the subcutaneous tissue (stage I and II). Once the melanoma cells enter the circulation the melanoma is considered metastatic. Stage III melanoma is characterized by migration to a regional lymph node, while metastases at distant sites is considered stage IV melanoma. Figure from Pitcovski *et al.* [25], authorized by Elsevier and Copyright Clearance Centre (license number 4471531007531).

1.1.4 Phenotype switching

As mentioned above, the progression of melanoma from the primary site to regional lymph nodes and distant sites requires cells to migrate and invade into subcutaneous tissue from where they can enter the circulation. To achieve this, the melanoma cells need to undergo morphological changes. One cellular process that plays a very important role in the cell's ability to spread is epithelial-mesenchymal transition (EMT). EMT, first described in embryonic development, is a process during which cells change in morphology and functionality from an epithelial to a mesenchymal phenotype. Over the past decades, many studies have focused on EMT in the context of tumor progression. EMT is involved in the initial steps of metastasis by changing the adhesive properties of the cell, including loss of cell-cell adhesion, altered polarity and reorganization of the cytoskeleton, allowing increased cell motility. This provides tumor cells with the invasive ability to access the circulatory system and migrate to a distant site to establish metastases. During this process, cells gain mesenchymal characteristics via the downregulation of epithelial cadherin (E-cadherin, *CDH1*) and gain in neuronal cadherin (N-cadherin, *CDH2*) expression, known as the "cadherin switch" [26-28]. Induction of EMT is controlled by a complex network of EMT-inducing transcription factors (EMT-TFs), including SNAIL, SLUG, TWIST, Zinc finger E-box-Binding homeobox (ZEB)1 and ZEB2. While there is strong evidence supporting involvement of EMT in metastatic spread, the level at which EMT is required for invasiveness and metastasis remains a topic of debate [29, 30].

Phenotype switching model

Melanoma cells originate from melanocytes, which are derived from neural crest cells, and are therefore not from epithelial origin. Nevertheless, a process similar to EMT was observed in melanoma progression whereby cells transition towards a de-differentiated, more aggressive phenotype [31]. The observation in metastatic melanomas that both proliferative and invasive cells are present in metastatic melanoma together with the finding that cells can switch between the proliferative and invasive state, both at the transcriptional and phenotypic level [32], led to the hypothesis that metastatic melanoma can arise via two pathways: 1) the acquisition of pro-metastatic mutations that promote invasiveness, which is considered an irreversible event, and/or 2) through phenotype switching whereby microenvironment-driven changes to a cell's phenotype lead to an increased invasive potential. Once the invasive cell enters a different

microenvironment, stimuli from this environment can again change the phenotype of the cell into a more proliferative state. This pathway towards metastasis is proposed to be a reversible event [33]. Importantly, this reversibility between a proliferative and invasive state would not only contribute to the initial spread of tumor cells from the primary tumor, but can also foster metastasis arising from a previous metastasis (**Fig. 2**) [32].

Similar to 'typical' EMT, studies in melanoma tissues and cell lines have demonstrated that expression levels of E-cadherin and N-cadherin play an important role in phenotype switching. Cell motility and metastatic potential were increased by E-cadherin downregulation and N-cadherin upregulation [31]. On the other hand, some 'typical' EMT-TFs like SNAIL and ZEB were shown to behave rather differently during melanoma progression. Paradoxically, SNAIL2 and ZEB2 are expressed in benign melanocytes and lost during progression to a metastatic lesion, while TWIST and ZEB1 expression are increased [34]. In another study it was shown that the increase in ZEB1 expression resulted in decreased E-cadherin expression [35]. The EMT-TF SLUG, whose expression depends on Secreted Protein Acidic and Rich in Cystein (SPARC) [36], promotes melanoma invasion via transcriptional activation of ZEB1 [37].

Another EMT-TF that is normally restricted to the melanocytic lineage and plays a crucial role in melanoma progression is MITF. MITF is involved in multiple steps during melanocyte development and plays a crucial role in melanoma plasticity between invasive, proliferative and differentiated phenotypes [38, 39]. MITF was shown to be required for melanocyte development in mice [40] and drives expression of melanocyte differentiation genes, which led to MITF being named the "master regulator" of melanocytes [41]. MITF activity is modulated at transcriptional, post-transcriptional, post-translational, and epigenetic levels as well as via microenvironmental signals. In addition, MITF does not act alone as a transcription factor and many co-regulators of MITF-target genes have been described [38, 39]. The relation between MITF activity and phenotype switching appeared to be rather complex, which led to the proposal of the so-called "MITF rheostat" [33, 42]. According to this model absence of MITF activity causes senescence or cell death, low-level activity promotes an invasive stem cell-like phenotype and mid-level activity promotes proliferation [39]. This model is supported by downregulation of MITF in invasive cells compared to proliferative cells [32]. One mechanism that appears to be an

important regulator of MITF expression during phenotype switching is the switch in ZEB expression [34, 35]. These findings indicate a key role for MITF in phenotype switching and support the importance of an MITF low state for invasion. However, in a recent review, Arozarena and Wellbrock pose the interesting question if de-differentiation towards neural-crest-like cells, which are MITF negative, is really required or if de-differentiation to melanoblast-like cells, which are MITF positive, is sufficient for migration and invasion [43]. Future studies are needed to address this question.

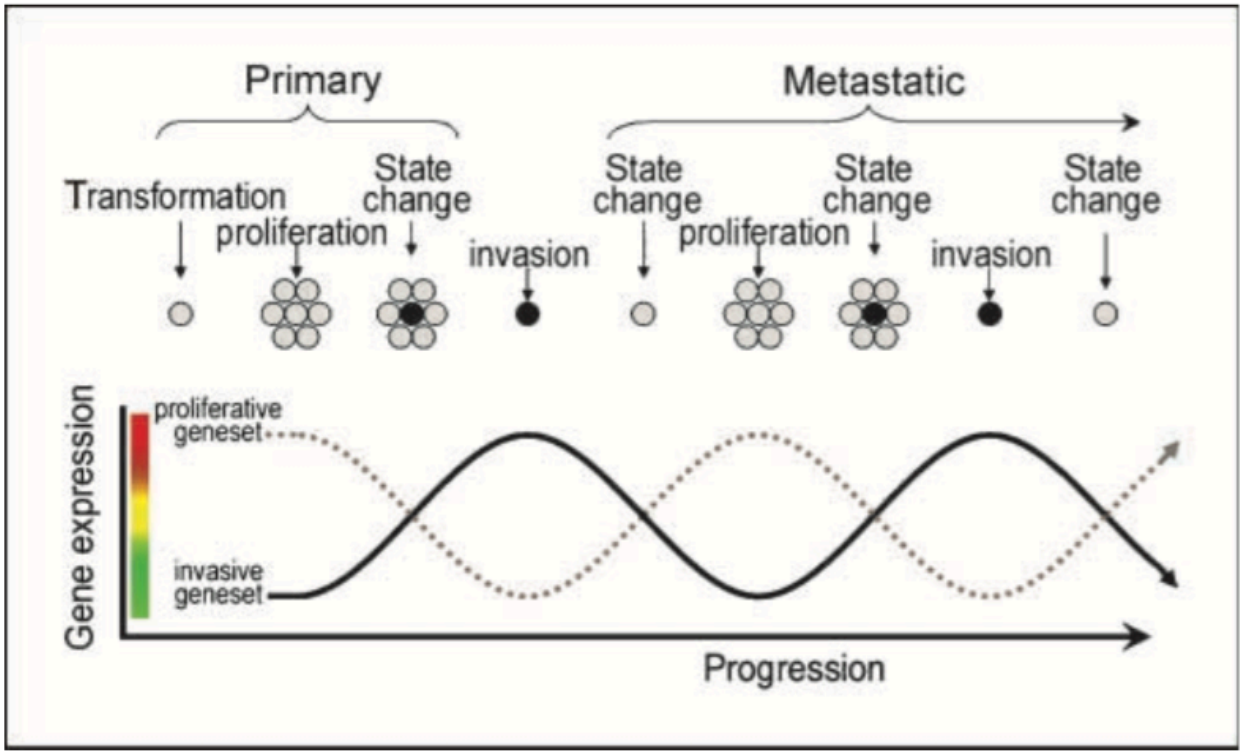


Figure 2. A schematic overview of the phenotype switching model.

The primary tumor is formed upon transformation followed by proliferation of melanoma cells that harbor a proliferative geneset. Various signals, likely triggered by microenvironmental changes, can lead to a change in cell state from proliferative to invasive, which is accompanied by a switch to an invasive geneset. The invasive cell has the ability to reach a distal site where, under suitable conditions, the cell can revert back to a proliferative state and form a metastasis. More metastases can arise by repetition of the switch between phenotypes and their transcriptional signatures. Figure from Hoek *et al.* [32], authorized by American Association for Cancer Research and Copyright Clearance Centre (license number 4471520158222).

Pathways in phenotype switching

Phenotype switching is not only induced via transcriptional reprogramming, but can also be promoted via signaling pathways, including the mitogen activated protein kinase (MAPK), transforming growth factor (TGF)-beta and Wnt pathways.

The main signaling pathway affected in melanoma is the MAPK pathway with approximately 50% of melanoma patients carrying a *BRAF* mutation, most commonly the *BRAF*^{V600E}, and 25% a *NRAS* mutation [14]. It is therefore not surprising that over-activation of BRAF is linked to invasion, cell motility and the expression of mesenchymal proteins. One way through which *BRAF*^{V600E} achieves this is via activation of Nuclear Factor-κB, which promotes the expression of matrix-metalloproteinases (MMPs) that increase the cell's migratory capacity [44, 45]. Another wat is via MEK/ERK that are reported to regulate proteases that play a role in matrix degradation and cell adhesion, promoting invasion [46]. In addition, activation of the MAPK pathway is associated with the induction of SNAIL and SLUG [47] as well as TWIST and ZEB1 [34].

TGF-beta is one of the most extensively studied factors involved in melanoma progression. Many roles have been described for TGF-beta, including extracellular matrix (ECM) remodeling via synthesis of MMPs, MITF downregulation, and induction of SNAIL and SLUG, all contributing to phenotype switching [47, 48]. In addition, there is an extensive interplay between TGF-beta family members and the ZEB transcription factors [49], resulting in cooperation of ZEB1 with GLI zinc finger protein 2 to repress E-cadherin expression [50]. Furthermore, melanoma cell lines with an invasive transcriptional state demonstrated high levels of TGF-beta expression and signaling [33, 51]. These findings point to an important role of the TGF-beta family in phenotype switching.

Another signaling pathway that plays an important role in melanoma progression and phenotype switching is the Wnt pathway. Wnt proteins can signal via three distinct pathways: 1) the canonical Wnt pathway that involves beta-catenin, and the non-canonical 2) planar cell polarity pathway, and 3) calcium (WNT/Ca^{2+}) pathway [52]. The role of the canonical Wnt pathway in phenotype switching remains unclear due to conflicting findings. For example, Wnt activation

was linked to downregulation of E-cadherin via upregulation of SNAIL, promoting melanoma invasion [53, 54]. On the other hand, high level and nuclear localization of beta-catenin, a key downstream component of the canonical pathway, were associated with improved survival and suppresses invasion [55, 56]. Furthermore, Wnt/beta-catenin and MAPK signaling cooperate to activate MITF and upregulate expression of differentiation antigens, promoting proliferation [57]. Besides canonical signaling, the WNT/Ca^{2+} pathway is involved in phenotype switching as well, mainly by promoting invasion via Wnt5A signaling. Activation of Wnt5A reduces MITF expression, resulting in the downregulation of differentiation antigens [58]. In addition Wnt5a signaling leads to an invasive phenotype via increased migratory capabilities and upregulation of Vimentin and SNAIL [52].

The proliferative and invasive transcriptional state

It was the development of technologies for large-scale gene expression analysis that provided the opportunity to investigate the transcriptional landscape underpinning phenotypic behaviors of melanoma cells. This has led to the identification of two transcriptional states that correspond to the two major different functional phenotypes and are referred to as the 'proliferative' and 'invasive' state [32, 51, 59]. Interestingly, no correlation was observed between *BRAF/NRAS* mutation status and the gene expression pattern related to either state. While melanoma cells belonging to the proliferative gene expression cohort demonstrated inhibition of proliferation upon TGF-beta treatment, the highly motile cells with an invasive gene expression signature were resistant to TGF-beta. In addition, they displayed TGF-beta-mediated upregulation of ECM remodeling factors and inhibitors of canonical Wnt signaling [51]. Further studies *in vivo* revealed that cells with a proliferative or invasive gene signature had a corresponding phenotypic effect upon injection in mice. Surprisingly, immunohistochemistry of the tumors suggested the presence of both proliferative and invasive cell types, which led to the proposal that melanoma cells can undergo a transcriptional switch [32]. The importance of the above defined transcriptional states was highlighted by the finding that about 86% of tested melanoma cell lines and cultures could be assigned to either phenotypic state [60]. Recently, with the use of ribonucleic acid-sequencing (RNA-seq), the transcriptional signature for proliferative and invasive melanoma cell lines from Hoek *et al.* [51] was confirmed and expanded using both publicly available data and in-house generated datasets. In their study, Verfaillie et *al.* revealed

that cell lines with an invasive phenotype could be identified based on their chromatin state, which demonstrated widespread open and active chromatin at regulatory regions that are active in cells with a mesenchymal regulatory program (skin fibroblasts). Next, they set out to find transcriptional regulators for each state and identified SRY-Box 10 and MITF as regulators of the proliferative state, and Activator Protein 1 (AP-1) and the TEA domain (TEAD) family as regulators of the invasive state [59]. In another study, the same group confirmed the role for MITF as regulator of the proliferative state and identified Nuclear Factor of Activated T Cells 2 (NFATC2) and Nuclear Factor I B (NFIB) as additional regulators of the invasive transcriptional state [61].

1.1.5 Melanoma outcome

Melanoma outcome is directly linked to the progression and stage of disease, with a worse outcome for stage III and IV patients. Therefore it is not surprising that the location and thickness of the tumor, absence or presence of ulceration and the mitotic rate all have an effect on melanoma survival. In addition, the age and sex of the patient play a role in the disease outcome. Interestingly, various studies consistently report a superior survival for women with melanoma, even after adjusting for multiple confounding factors like tumor thickness, age and anatomic site [62-65]. Even stage-specific analysis of survival differences between males and females revealed a survival benefit for women [66, 67]. Despite these clear differences in outcome between men and women, the underlying reason for this effect remains poorly understood. Multiple biological and behavioral differences have been proposed, including differences in UV radiation, sex-specific hormone levels, immune homeostasis, oxidative stress response, and health maintenance [64, 68, 69].

Overall, melanoma has a high relative 5-year survival with 92% for females and 85% for males in Canada [6]. The low overall mortality rate (1.9% in males and 1.2% in females in Canada, all stages together [6]) can be mainly attributed to the ability to cure melanoma via surgical removal if the melanoma is detected at stages I-III. While the importance of early detection is well-appreciated, the first clinical signs of melanoma are often not diagnosed till the stage of metastasis (stage IV) [25]. Over the past decade, treatment options for metastatic melanoma have greatly enhanced melanoma outcome. Vemurafenib, an inhibitors targeting the *BRAF*^{V600E}

mutation, demonstrated improved rates of overall and progression-free survival compared to chemotherapy with dacarbazine [70]. Unfortunately, the majority of patients developed therapeutic resistance after 6 to 8 months of treatment [71, 72]. Melanoma is considered highly immunogenic and is therefore seen as a prime target for immunotherapies including, monoclonal antibodies, adoptive T cell therapy and cancer vaccines [25]. One approach aims to overcome the tumor's ability to suppress antitumor immune responses by inhibiting negative regulators of the immune response. This has led to the development and approval of monoclonal antibodies known as immune checkpoint inhibitors, which have revolutionized the treatment and outcome of metastatic melanoma [73-75]. Another approach attempts to stimulate a cytotoxic T cell response to attack the tumor cells based on the expression of tumor-specific antigens. This will be further elaborated on in section 1.2. Despite these advances in treatment options, in Canada, melanoma still leads to about 1250 deaths yearly (3.1 deaths per 100,000) and is the most common cancer-related death among young adults (15-29 years) [6].

1.2 CANCER TESTIS ANTIGENS

This section provides background on tumor antigens and focusses on the expression, regulation and biological functions of a subset of tumor antigens called cancer testis antigens, which is discussed in the context of melanoma based on available literature.

1.2.1 Tumor antigens

Tumor antigens (TAs) are proteins or other molecules that are present on cancer cells, but absent or expressed at a low level on normal cells, and can therefore be used to elicit an immune response targeting cancer cells. It was the observation that melanoma patients can generate T cells with antitumor reactivity that elicit an immune response against their own tumor, which sparked the investigations that led to the identification of TAs [76, 77]. Nowadays, it is known that many tumor types produce antigens that can be released into the bloodstream or remain on the cancer cell surface. TAs can originate from viral proteins in cancer, release of antigens that are normally sequestered in the cell's organelles or membrane, and expression of proteins that are normally not present or at very low level. In addition, TAs include epitopes generated from mutated genes, like oncogenes and tumor suppressor genes initially involved in tumor development, that produce mutated proteins not present in normal cells. Our immune system can detect these non-self antigens and elicit a response to eradicate them. Due to their cancer specificity, targeting TAs is considered a promising approach for cancer immunotherapy, including cancer vaccines [78-80].

Tumor antigens in melanoma

Melanoma-associated antigens were among the first TAs to be identified in humans [81-83]. Some TAs demonstrated increased expression with melanoma progression and could therefore be used as immunological markers [84-86]. Based on tissue distribution and molecular function, melanoma-associated antigens can be categorized into several classes, including differentiation antigens, overexpressed antigens, neoantigens, and cancer testis antigens (CTAs) [25, 87].

Differentiation antigens are genes involved in different stages of melanocyte/melanoma differentiation [87]. These antigens are considered to be aberrantly expressed self-antigens, as they are expressed by normal melanocytes as well as by malignant melanoma cells, but not by

other cell types. The most well-known melanocyte-specific differentiation antigens include Premelanosome protein [88], Melan-A (*MLANA*) [89], Tyrosinase [90], Tyrosinase-related protein 1 [91], and Tyrosinase-related protein 2 [92]. Overexpressed antigens are antigens that are expressed in both melanocytes and other normal cells, but at a much lower level compared to melanoma cells. In some cases aberrant expression is due to alterations in expression regulation, like mutations in the promoter region [17]. This category includes human telomerase reverse transcriptase [93] and survivin [94]. Neoantigens result from mutations in tumor cells (tumor-specific mutations) that produce new antigenic peptides and are therefore considered non-self antigens. Due to the high mutational load in melanoma [12] multiple neo-antigens have been identified, including beta-catenin [95], CDK4 [96], and BRAF [97]. For an extensive list of neoantigens identified in melanoma please see Lu and Robbins [98]. Another group of TAs expressed in melanoma are the CTAs. As this thesis focuses on a specific CTA, CTAs will further elaborated on in the following section.

1.2.2 Cancer testis antigens

CTAs are a group of TAs that are normally expressed in male germ cells and become aberrantly re-activated in cancer. This unique expression pattern together with their immunogenicity makes CTAs attractive biomarkers and/or targets for immunotherapy [99]. To date >200 CTAs have been identified. Known CTAs and CTA families include, Melanoma Antigen (*MAGE*), B Melanoma Antigen (*BAGE*), G Antigen (*GAGE*), Sarcoma Antigen (*SAGE*), X Antigen (*XAGE*), BORIS, Synovial Sarcoma X (*SSX*), CTA family 45 (*CT45*), New York Esophageal Squamous Cell Carcinoma 1 (*NY-ESO-1*) and Serine Protease 50 (*PRSS50/TSP50*) [100, 101]. CTA genes can be broadly divided into two groups: X-chromosome CTA genes (X-CTAs), which are encoded from the X-chromosome, and non-X-chromosome CTA genes (non-X-CTAs) located on autosomes and the Y-chromosome in males [101]. Interestingly, more than half of all CTAs are X-CTAs, covering about 10% of all genes mapped to the X-chromosome [102, 103]. On the X-chromosome, CTAs can usually be found in clusters as inverted or direct repeats [103], while non-X-CTAs are distributed throughout the genome as mostly single copies [101]. Another difference between X-CTAs and non-X-CTAs is their expression pattern in normal testis: X-CTAs are expressed on proliferating germ cells and non-X-CTAs on spermatocytes during germ cell differentiation [101]. Melanomas frequently express X-CTAs with about 40-80% of the

samples being positive for *MAGE-A1*, 57-76% for *MAGEA3*, 17-42% for *NY-ESO-1*, 35% for *SSX-2* [100], and 14-28% for *BAGE* [104]. This is one of the reasons why most clinical trials investigating the use of CTAs in the context of immunotherapy have been performed in melanoma [100, 105]. X-CTAs are often co-expressed in tumors [101]. In agreement, it was found that the large majority of melanomas (~65%) express three or more X-CTAs [106]. Importantly, expression of X-CTAs in tumors is heterogeneous [107], meaning that different cells within a tumor can express different CTAs, which is likely due to the different cell types that make up a tumor. Less is known about expression patterns of non-X-CTAs [101]. CTAs become expressed via deregulated epigenetic mechanisms and their expression has been linked to various processes in tumor cells, which will be discussed in more detail below.

1.2.3 Regulation of cancer testis antigen expression

The observation that different CTAs can be co-expressed in a tumor suggests a common regulatory mechanism driving their expression [108, 109]. However, transfection studies using *MAGE-A1* promoter constructs in *MAGE-A1* positive and negative cell lines indicated that the presence of transcription factors alone is not sufficient to drive expression of this CTA [110, 111]. These findings all pointed to the involvement of other mechanisms of expression regulation. Studies over the past two decades have revealed that the unique expression pattern of CTAs in both germ cell and cancer cells is mainly driven by epigenetic regulatory mechanisms, including DNA methylation and histone modifications [104, 112].

DNA methylation

DNA methylation of CTA promoters is observed in normal somatic tissues for all CTAs studied thus far, with DNA demethylation resulting in CTA activation in germ cells during spermatogenesis [100, 111]. Weber *et al.* were the first to demonstrate activation of a CTA, *MAGE-A1*, in melanoma cell lines upon treatment with the DNA methyltransferase (DNMT) 1 inhibitor 5-aza-2'deoxy-cytidine (5-aza-dC) [113]. This finding was confirmed by a study from De Smet *et al.* who in addition revealed that methylation of a CpG dinucleotide in the *MAGE-1* promoter inhibits transcription factor binding. Furthermore, they showed an inverse correlation between the overall level of CpG methylation and *MAGE-1* expression in 20 melanoma cell lines [110]. Expression regulation via DNA methylation has also been reported for other members of

the MAGE family [114-116]. DNA (de)methylation is not restricted to CTAs of the MAGE family. In prostate cell lines various CTAs were activated upon treatment with 5-aza-dC. This effect appeared to be cell line specific and in some cases associated with androgen dependency. Examination of *SSX-2* and *NYSAR35* expression in prostate tissue cultures in the presence or absence of 5-aza-dC revealed an increase in *SSX-2* expression in all samples, whereas *NYSAR35* was activated in 3 out of 9 cultures [117]. The activation of *SSX-2* upon treatment with 5-aza-dC was confirmed in prostate cancer cells and also observed in other types of cancer [114, 118]. Other CTAs that are activated upon global DNA demethylation, include *NY-ESO-1, LAGE-1*, Thioredoxin peroxidase 1, Bromodomain testis associated (*BRDT*) and Leucine zipper protein 4 [114, 119-121]. In addition to global DNA demethylation by the use of 5-aza-dC, the effect of a genetic knockout of DNTM1 and/or DNMT3b on CTA promoter methylation and induction has been investigated. While the knockout of either DNMT1 or 3 had a moderate to negligible effect in HCT116 colon carcinoma cells, the double knockout resulted in robust hypomethylation and increased expression of *MAGE-A1, NY-ESO-1* and *XAGE-1* [122]. A similar effect was observed for *BRDT, LAGE-1, PX1* and *NY-ESO-1* upon DNMT1/3b knockout in a non-small-cell lung carcinoma cell line [120]. Overall, these findings highlight the crucial role of DNA methylation in the regulation of CTA expression.

Histone modifications

Given that in some tumors global hypomethylation did not induce CTA expression, other forms of regulatory mechanisms must exist. Histone modifications were a likely candidate to play a role in the regulation of CTA expression, since DNA methylation and post translational modifications of histones are often interdependent. Indeed, studies using histone deacetylase inhibitors demonstrated upregulation of various CTAs, especially in combination with global DNA demethylation [100, 123]. While treatment with the histone deacetylase (HDAC) inhibitor trichostatin A (TSA) led to increased *MAGE-A1, -2, -3* or *-12* expression in some cancer cell lines, a combination of TSA with 5-aza-dC almost always resulted in robust activation of these CTAs. The increase in CTA expression was at least in part due to promoter demethylation induced by 5-aza-dC treatment, but not by TSA [124]. Similarly, *MAGE-A3* expression was enhanced 2-fold upon treatment with the HDAC inhibitor Depsipeptide FR901228 in lung and esophageal cancer cell lines, though in combination with 5-aza-dC expression was induced 57-

fold [116]. Initial evidence for CTA expression regulation by histone methylation changes came from studies in mouse embryonic stem cells lacking the histone methyltransferase G9a or GLP. In these cells *Mage-a* antigens were upregulated, which coincided with reduced H3K9me2 at the *Mage-a* promoter [125, 126]. However, knockdown of G9a and GLP or G9a targeting with a small molecule inhibitor in the human colorectal cancer cells did not induce *MAGE-A1* or *XAGE-1* expression despite reduced H3K9me2 at the promoters. Combining histone methyltransferase inhibition with TSA did not consistently induce CTA expression, suggesting that reduced H3K9me2 with increased histone acetylation is not sufficient for CTA activation. However, an increase in *MAGE-A1* and *XAGE-1* expression was observed with both G9a knockdown and DNMT inhibition, again demonstrating the interdependence of DNA methylation and histone modification in the regulation of CTA expression [127]. Similar results were obtained for induction of GAGE expression when combining 5-aza-dC with a HDAC and histone methyltransferase inhibitor in the breast cancer cell line MCF7. Interestingly, further investigations into the regulation of *GAGE* expression in three breast cancer cell lines with low, medium and high *GAGE* levels revealed involvement of different epigenetic mechanisms in each cell line [128]. Overall, it appears that in most cases CTAs are silenced by DNA methylation and repressive histone marks, and require DNA demethylation as well as histone acetylation and active methylation marks for their induction.

1.2.4 The role of cancer testis antigens in cancer

The biological functions of CTAs are not fully understood, but evidence suggests they are involved in proliferation, differentiation, transcriptional control, apoptosis and survival in germ cells and likely have similar effects in cancer [100]. It is surprising how little is known about the functional role of CTAs in cancer, considering their potential use towards immunotherapy. While non-X-CTAs have known roles in spermatogenesis and fertility, their roles in cancer cells remain to be defined. Regarding X-CTAs, most studies have focused on the MAGEs [101, 112, 129].

For example, MAGE-A1 can act as a transcriptional repressor via its interaction with Ski-interacting protein and recruitment of HDAC1 [130]. In melanoma cells, MAGE-A1 was shown to enhance migration and invasion through activation of c-JUN [131], a subunit of the transcriptional regulator AP-1. MAGE-A4 interacts with the oncoprotein gankyrin in

osteosarcoma cells and suppresses its oncogenic activity as shown in mice [132]. In a myeoloma cell line, MAGE-A3 was required for survival, as short hairpin RNA (shRNA)-mediated knockdown of MAGE-A resulted in apoptosis via increased Bax expression and reduced Survivin expression [133]. In addition, several MAGEs play a role in protein ubiquitination through their interaction with Really Interesting New Gene (RING) E3 ubiquitin ligases forming so-called MAGE-RING ligases. Via the formation of these complexes, MAGEs have been implicated in a wide range of biological functions, including transcription, metabolism, protein trafficking, and cell proliferation. For further details please see Feng *et al.* [134] and Lee and Potts [98]. Overall, these findings indicate that MAGE proteins are multifunctional regulators with a wide range of functional effects [99, 101].

Besides MAGEs, some studies have revealed biological functions for GAGE, SSX, NY-ESO-1 and BORIS. Members of the GAGE family were shown to confer resistance to apoptosis upon a variety of apoptotic stimuli and to promote proliferation and colony formation in HeLa cervival cancer cells [135]. The SSX family was linked to melanoma progression, since knockdown of SSX in a melanoma cell line reduced MMP-2 levels as well as the cell's ability to migrate [136]. Another CTA with a potential role in melanoma progression is NY-ESO-1. Expression of *NY-ESO-1* in primary melanomas was associated with a higher frequency of metastasis and *NY-ESO-1* is expressed at a higher level in metastasis compared to primary tumors [137]. A similar expression pattern with higher expression in metastatic compared to primary melanomas was observed for BORIS, which led to the proposal that BORIS expression is linked to melanoma progression [138]. The biological functions of BORIS in cancer will be described in detail in chapter 1.3.

1.3 BORIS/CTCFL

The first description of BORIS demonstrated that BORIS expression is restricted to the testis and becomes abnormally activated in cancer cells, which has led to the designation of BORIS as a CTA [139]. BORIS is the single paralogue of the multifunctional transcription factor CCCTC binding factor (CTCF), also known as the master weaver of the genome [140]. Our interest in BORIS was sparked by its similarity with CTCF [139] and its proposed role in melanoma progression [138]. This section provides background regarding BORIS' discovery, expression and regulation, protein structure and interacting partners, role as transcriptional regulator, functions in cancer and the development of a BORIS vaccine.

1.3.1 Discovery of BORIS

While analyzing testis nuclear extracts for proteins that bind to CTCF DNA binding sites, Loukinov *et al.* discovered BORIS [139]. BORIS stands for Brother Of the Regulator of Imprinted Sites, which references to CTCF's function in reading imprinting marks. Therefore, BORIS is also known as CTCF-like (CTCFL). The human *BORIS* gene contains eleven exons (10 coding: exon 2-11) and maps to chromosome 20q13, a chromosomal region that is often amplified in cancers. The *BORIS* gene encodes for a highly identical eleven zinc finger (ZF) domain compared to CTCF (>70% homology), which *in vitro* bind to the same DNA target sites. However, the N- and C-terminal parts of both proteins display very little similarity with each other (<18% homology; **Fig. 3**) [139, 141, 142]. Due to the high similarity between mammalian, but not avian BORIS and CTCF, Loukinov *et al.* suggested that BORIS and CTCF descended with divergence from a common ancestral gene [139]. Further studies in vertebrates proposed that BORIS derived from CTCF via a duplication event that occurred in a common ancestor of all amniotes (reptiles, birds and mammals). While CTCF orthologues show very high similarity with the human CTCF ZF domain, the conservation between the BORIS orthologues and human BORIS ZF's is much lower. It is suggested that BORIS' functions specialized during mammalian evolution, together with the evolution of genomic imprinting. This would also explain why BORIS expression in amphibians and reptiles can be detected in various somatic tissues, becomes restricted to the gonads in marsupials and cattle, and eventually is testis specific in humans and mice. [143]. Initially, BORIS appeared to be lost in avian species, though a small

108bp segment of DNA was discovered in the chicken genome that is homologous to BORIS's first ZF [139, 143].

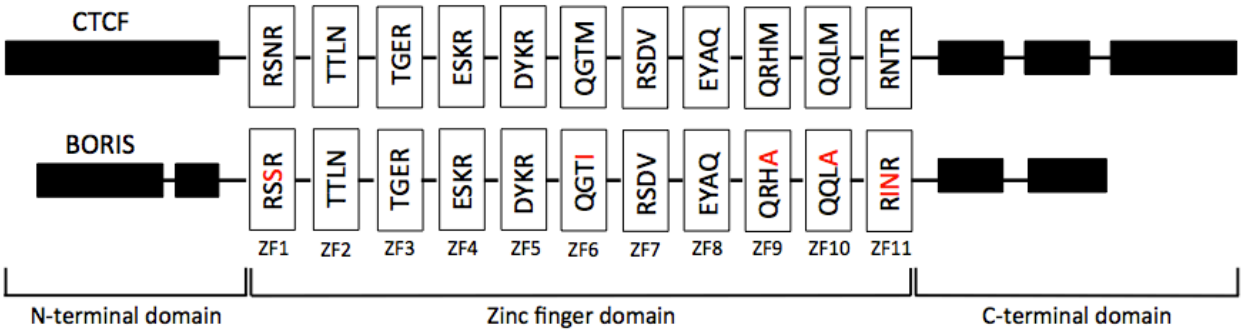


Figure 3. A schematic representation of the CTCF and BORIS protein.

CTCF and BORIS share a highly similar zinc finger domain (>70% homology), while the N- and C-terminal domains are divergent (<18% divergent). For both proteins the four essential amino acids involved in DNA binding are displayed for each zinc finger. The amino acids of the BORIS zinc fingers that differ from CTCF are highlighted in red. Adapted from Pugacheva *et al.* and Hore *et al.* [143, 144].

1.3.2 BORIS expression and regulation

Regulation of BORIS expression

Renaud *et al.* were the first to demonstrate that human *BORIS* transcripts can be expressed from at least three alternative promoters. Using 5' rapid amplification of complementary DNA (cDNA) ends, three transcription start sites were observed upstream of the first ATG designated A (-1447bp), B (-899bp) and C (-658bp). In the testis, *BORIS* is transcribed from all three promoters, while in the majority of cancer cell lines only promoters A and/or C are used. Upon discovery of the alternative promoters, five alternatively spliced 5'-untranslated regions (UTRs) were mapped that resulted in alternative *BORIS* transcripts [145]. Further studies revealed that the regulation of BORIS expression is an even more complex process by the discovery of 23 differentially expressed *BORIS* messenger RNA (mRNA) variants that have been functionally grouped into six isoform families (sf1-sf6). These gene isoforms encode specific N- and C-terminal translational frame shifts and contain various ZF combinations, giving rise to 17 different protein isoforms [146].

The *BORIS* promoter is negatively regulated by both DNA methylation and functional p53, while derepression of the *BORIS* promoter was observed upon reduced CTCF levels. This indicates an association between functional p53 and aberrant BORIS expression as well as haploinsufficiency of CTCF and BORIS activation [145]. The main factor driving BORIS expression appears to be promoter hypomethylation. Based on bisulfite sequencing, Hoffmann *et al*. demonstrated that the *BORIS* promoter is strongly methylated in normal tissues and cancer cell lines, partially methylated in testis and unmethylated in sperm cells. Methylation could be reduced by treatment with 5-aza-dC, resulting in increased BORIS expression [147]. While Hoffmann et al reported methylation of the BORIS promoter in cancer cell lines, many other studies demonstrated hypomethylation of the *BORIS* promoter in cancer cell lines as well as tumors, leading to aberrant BORIS expression [148, 149]. Interestingly, hypomethylation of the BORIS promoter is correlated with its own expression [147, 150]. In addition to methylation, BORIS expression can be increased in various cancer cell lines by treatment with the HDAC inhibitor Depsipeptide FK228 [151] or TSA [152].

BORIS expression in non-cancerous cells

Upon the discovery of BORIS, BORIS mRNA and protein expression in human and mice was only observed in the testis, more specifically in CTCF negative primary spermatocytes. In BORIS positive cells, approximately 10-30% of BORIS protein was observed in the nucleus and the remaining in the cytoplasm. Despite the majority of BORIS being present in the cytoplasm, it was shown using different experimental approaches that a significant portion of DNA bound BORIS to be localized in the nucleus. Interestingly, the expression of BORIS in CTCF negative cells has been associated with erasure of DNA methylation. On the other hand, CTCF expression in BORIS negative spermatids seems to co-occur with the establishment of methylation marks, suggesting that a BORIS-CTCF switch is linked to epigenetic reprogramming during male germ cell development [139, 141]. This has led to the designation of BORIS as an epigenetic modulator. The significance of BORIS expression in the testis was highlighted in BORIS knockout mice, which displayed defects in spermatogenesis due to a faulty meiosis resulting in infertility [153, 154]. Likely, BORIS's involvement in the expression of testis-specific mouse genes, including *Gal3st1*, *Stra8* and *Prss50* plays a role in the defect spermatogenesis [153]. While initially BORIS and CTCF expression was reported to occur in a mutually exclusive pattern [139], a more recent report demonstrated BORIS expression at stages of spermatogenesis where CTCF is present as well [153], suggesting a potential competition for shared binding sites in the testis.

The observation of restricted BORIS expression in the testis was recently challenged by various studies. Besides male germ cells [139, 153], BORIS was also detected in female germ cells (oocytes) [155]. Furthermore, expression of BORIS in many non-cancerous somatic tissues, including gut, kidney [156], skin [157], ovary [152, 156] and prostate [147, 156, 158] has been documented [149]. In a recent paper, Pugacheva *et al.* state that the observation of BORIS in somatic tissues is likely a false positive detection due to the lack of reliable commercial antibodies [144], hereby referring to the study from Jones *et al*. [156]. Currently, studies looking into BORIS expression in somatic tissues have only been performed in the context of skin cells. In epidermal cells, BORIS protein accumulated in nuclear and perinuclear foci and did not overlap with CTCF expression. Immunofluorescence (IF) in epidermal keratinocytes and keratinocyte cell cultures for BORIS and the nucleolar markers upstream binding factor (UBF)

and Fibrillarin revealed that BORIS localizes to the nucleoli. In addition, BORIS was found to co-localize with areas of nascent RNA and shown to play a role in pre-ribosomal RNA (rRNA) synthesis and overall transcription. Furthermore, BORIS expression was linked to cell cycle progression [157]. Overall, the expression of BORIS protein in somatic tissues and its role in these tissues remains to be further investigated.

BORIS expression in cancer

As mentioned before, the BORIS gene is located within a chromosomal region that is often amplified in cancer [141]. Furthermore, hypomethylation of the BORIS promoter has a high occurrence in cancer cells [148, 149]. Therefore, it is not surprising that BORIS expression in cancer is widely observed [149] (**Table I**). Hong *et al.* and Vatolin *et al.* were the first to report *BORIS* expression at the RNA level in lung cancer, melanoma, neuroblastoma, breast cancer, prostate cancer and colon cancer cell lines as well as breast, prostate and colon primary tumors [151, 159]. In addition, expression of *BORIS* has been reported in leukemic cell lines, ovarian, uterine, cervical, and endometrial cancer, glioblastoma, and esophageal and laryngeal squamous carcinoma [146, 150, 160-164]. Currently, two studies tried to tackle the expression of different *BORIS* RNA isoforms in cancer. Link *et al.* observed aberrant expression of select isoform families in epithelial ovarian cancer compared to normal ovary, which was associated with global hypomethylation in the tumors [152]. Novak *et al.* assessed *BORIS* mRNA variants containing exon 7 and lacking exon 7 in the context of laryngeal squamous cell carcinoma and reported a higher number of disease relapse with higher expression of the *BORIS* isoform containing exon 7 [165]. Expression of BORIS at the protein level was demonstrated in myelogenous leukemia, laryngeal squamous carcinoma, glioblastoma, hepatocellular carcinoma, neuroblastoma, melanoma, colorectal, breast, prostate, lung and esophageal squamous cancer [144, 151, 156, 158, 162, 164-168].

Regarding melanoma, *BORIS* expression was first reported by Vatolin *et al.* who demonstrated the presence of *BORIS* in 9 out of 10 melanoma cell lines using end point PCR (SK-MEL-5, MALME-3M, SK-MEL-2, SK-MEL-28, Godowns-MEL, 624.28-MEL, 928-MEL, 11395-MEL and A375) [159]. Renaud *et al.* confirmed the expression of *BORIS* in the SK-MEL-2 cell line using quantitative PCR (qPCR). In addition, they showed that *BORIS* was expressed from

promoters A (68%) and C (32%), despite methylation of promoter C in this cell line [145]. The most extensive analysis of *BORIS* expression in melanoma came from a study by Kholmanskikh *et al.* who demonstrated the presence of *BORIS* in 59% (19 out of 32) melanoma cell lines. Furthermore, they observed activation of *BORIS* in 16% (4 out of 25) of primary melanomas and 34% (13 out of 38) of metastatic melanomas. It was proposed that the discrepancy between BORIS expression in cell lines and tumor tissue could be due to the heterogeneous nature of tumor tissues. In addition, the authors speculate that higher frequency of BORIS activation in metastatic tumors compared to primary tumors indicates a role for BORIS in metastatic process [138].

BORIS mRNA versus protein expression

Jones *et al.* reported a lack of correlation between BORIS mRNA and protein expression in various cell lines and tissue types [156]. Ogunkolade *et al.* propose that this discrepancy could be due to a short half-life of the *BORIS* transcript and high BORIS protein stability [169]. However, the half-lifes determined for four *BORIS* isoforms expressed from promoters A and C revealed a relatively stable half-life between 4.7 and 7.9 hours in K562 cells. The half-life of *BORIS* expressed from promoter C did appear to be less stable. Similar half-lifes for some of these variants were also obtained in the Ovcar-8 cell line [145]. In contrast to Jones *et al.*, others did observe a correlation between BORIS mRNA and protein expression [158]. Currently, the half-life of BORIS protein isoforms remains unknown.

Table I. BORIS expression in cancer (cell lines and tissues).

Type of cancer	mRNA/Protein	References
Bladder	mRNA	[147]
Breast	mRNA/Protein	[145, 159, 170, 171]
Cervical	mRNA	[161]
Colon	mRNA/Protein	[145, 159]
Endometrial	mRNA	[160]
Esophageal	mRNA/Protein	[162]
Glioblastoma	mRNA/Protein	[164]
Head And Neck	mRNA	[172]
Hepatocellular	mRNA/Protein	[168, 173]
Kidney	mRNA	[145]
Laryngeal	mRNA	[165]
Leukemic	mRNA	[145, 146]
Lung	mRNA	[145, 151]
Melanoma	mRNA	[138, 145, 159]
Neuroblastoma	mRNA	[159]
Osteosarcoma	mRNA	[174]
Ovarian	mRNA	[145, 150, 152]
Prostate	mRNA/Protein	[147, 158, 159]
Uterine	mRNA	[163]

*Adapted from and according to Martin-Kleiner [148], de Necochea-Campion *et al.* [175] and Soltanian and Dehghani [149].

1.3.3 BORIS protein

BORIS protein structure

The human BORIS (hBORIS) protein counts 663 amino acids and consists of three main domains: a N-terminal domain and C-terminal domain that comprise about half of the protein, and an eleven ZF domain that is highly similar to CTCF and provides BORIS the ability to bind CTCF binding motifs [139, 141, 143] (**Fig. 3**). In an effort to address reported functional differences between BORIS and CTCF, Campbell *et al.* investigated the protein structure of BORIS' terminal domains [176]. Similarly to their findings for CTCF's terminal segments [177], they report disordered content in the N- and C-termini of BORIS. Using hydrodynamic properties and measuring circular dichroism they observed that the C-terminal fragment is predominantly unordered with algorithms estimating around 3.5% helix, 13.5% strand, 8% turns and 75% unordered content. Since purification of native preparations of the N-terminal fragment was unsuccessful computational methods were used, which predicted a disordered N-terminal fragment as well [176]. Unstructured proteins or protein regions are known to have multiple biological functions, a major one being molecular recognition, which includes the recognition of one protein by another protein that can facilitate protein structure conformation changes [178, 179]. Due to the lack of sequence homology between BORIS and CTCF, and the unstructured terminal extensions that likely function in molecular recognition, the authors speculated that the interaction with different regulatory proteins underlies the functional distinction between BORIS and CTCF [176]. This view is shared by Loukinov *et al.* and Necochea-Campion *et al.*, who proposed that specific cellular functions of BORIS and CTCF at specific locations is determined by their protein partners [139, 175].

Compared to the human protein, mouse BORIS (mBORIS) is very similar in size (636 amino acids [148]. Despite this similarity only 2 isoforms have been reported for mBORIS (Uniprot database [180]). The *mBORIS* gene is located on chromosome 2 within bands H3-H4, which is orthologous to the human 20q13 locus that contains the *hBORIS* gene [148]. BLAST alignment of the mBORIS and hBORIS protein sequence revealed 56% overall similarity and 82% identity between the zinc finger domains (**Fig. 4**). This indicates that hBORIS and mBORIS have a similar DNA binding domain and could bind to the same sites on the DNA. Due to the

differences in the N- and C-terminal domains, the function of mBORIS in human cells and vice versa is likely inadequate as a result of the inability to interact with their respective binding proteins.

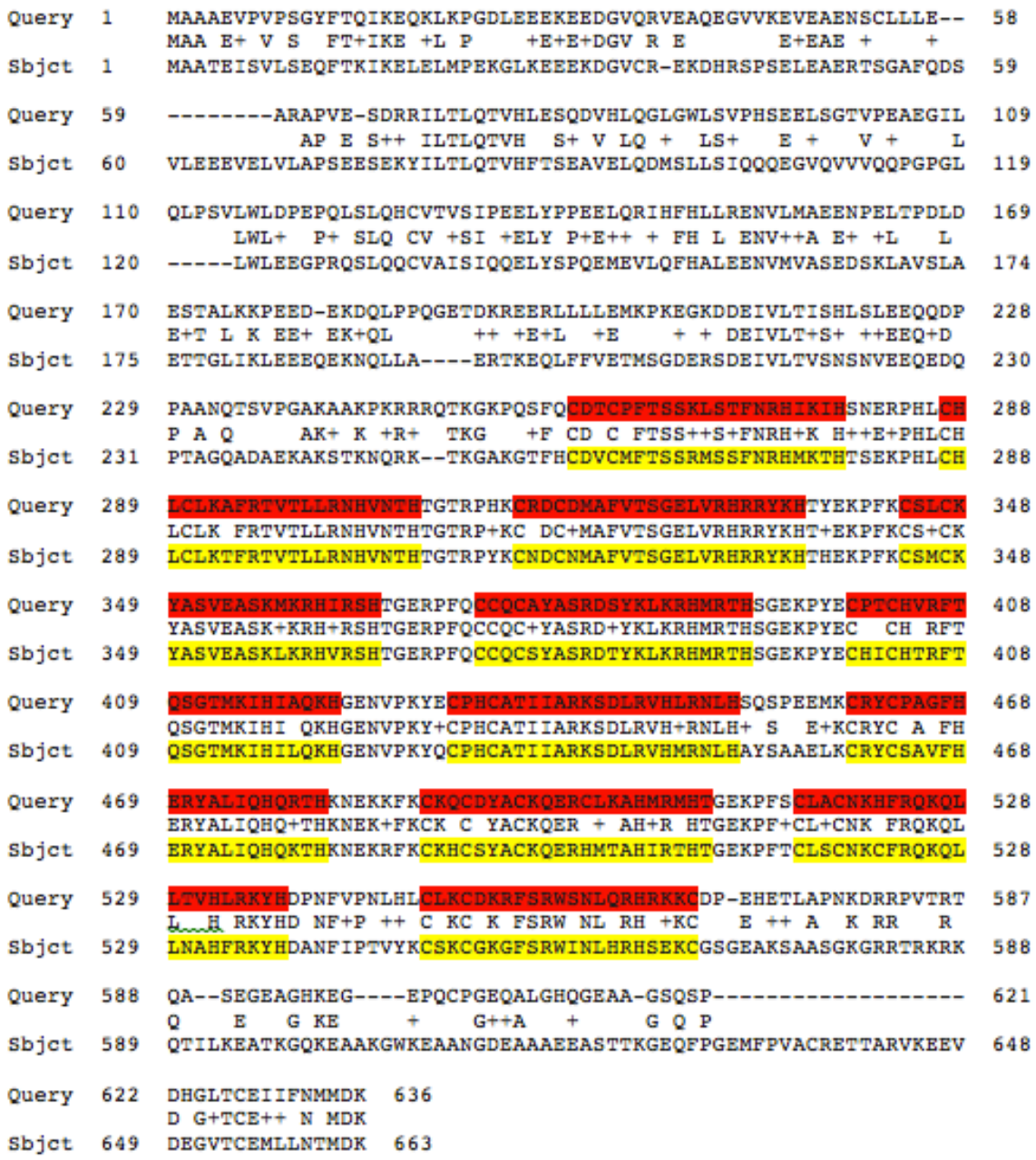


Figure 4. BLAST alignment of the mouse and human BORIS protein.

The Query represents mouse BORIS and the Subject (sbjct) human BORIS. Zinc finger domains are highlighted in red and yellow for mouse and human BORIS, respectively.

BORIS-interacting proteins

Currently, few studies have investigated BORIS protein partners. Most of the known BORIS-interacting proteins were identified in a study by Nguyen *et al.* who performed a yeast two-hybrid screen using multiple BORIS N-terminal and ZF domains as bait and a cDNA library of day 11.5 mouse embryos as prey. This approach identified 16 potential BORIS protein partners with five different N-terminal baits. These proteins are BAT3 (=BAG6), CHD8, ELF2, HCFC1, HCFC2, POGZ, MGA, TLK2, NFAT5, ZNF518, ATF7, CSTA, FSLZ1 (we are unable to identify this protein in online databases), SRCAP, MKL2 and FHL2. The remaining 8 baits, including all zinc finger domains did not reveal any potential protein partners [167]. In addition, other proteins reported to interact with BORIS are SET1A [167], CTCF [144, 181], TAF7L [181], Sp1 [182], UBF [183], PRMT7 and histones H1, H2A and H3 [184]. UBF is the only known common protein partner between BORIS and CTCF [183], which is in agreement with a recent *in silico* analysis that predicted very little overlap between BORIS and CTCF-interacting partners [185].

The different BORIS-interacting proteins are involved in various cellular processes, which indicates that BORIS is a multi-functional protein. Both, BAT3 and SET1A interact with BORIS to form a complex that is involved in H3K4me2. Binding of this complex at the promoter region of the pro-carcinogenic genes *MYC* and *BRCA1* promotes their expression [167]. TAF7L is a testis-specific transcriptional regulator that BORIS interacts with in human testis as shown by *in situ* proximity ligation assay. Mouse testis with heterozygous Taf7l knockout demonstrated reduced expression of the BORIS target gene *Prss50*, suggesting that either Tafl7 acts as a transcriptional repressor or inhibits transcriptional activation of *Prss50* by BORIS [181]. Using a yeast-two-hybrid screen with BORIS as bait and a mouse testis cDNA prey library Jelinc *et al.* identified PRMT7, a protein arginine methyltransferase, and demonstrated an interaction with the N-terminal region of BORIS. According to the authors, BORIS acts as an accessory protein for PRMT7 activity, which was supported by the finding that BORIS and PRMT7, together with Dnmt3a, Dnmt3b and Dnmt3L, induce methylation of the H19 ICR in *Xenopus* oocytes. In addition, oocyte microinjection with this combination of expression plasmids resulted in increased methylation of H4R3 [184]. While no direct functional readout was provided for the interaction between BORIS and the transcriptional activator Sp1, it was proposed that BORIS

recruits Sp1 to the promoter of *NY-ESO-1* and thereby promotes NY-ESO-1 expression [182]. Van de Nobelen *et al.* were the first to report an interaction between BORIS and the nucleolar ribosomal DNA (rDNA) transcription factor UBF. Both BORIS and CTCF interacted with the UBF dimerization domain and high mobility group box 1 via their C-terminal ZFs. The authors revealed that BORIS is able to bind the same rDNA region as UBF in K562 cells [183], though the effect of this finding on UBF function or rDNA transcription was not assessed. The localization of BORIS in the nucleoli of epidermal keratinocytes as well as various cancer cell lines prompted Rosa-Garrido *et al.* to further investigate the role of BORIS in rDNA transcription. They confirmed binding of BORIS to rDNA in HEK293T cells and demonstrated that BORIS knockdown using shRNAs leads to decreased pre-rRNA expression [157]. Interestingly, BORIS can also interact with CTCF, resulting in the formation of BORIS-CTCF dimers [144, 181]. Pugacheva *et al.* demonstrated that these dimers are present at CTCF target sites (CTSes) that are clustered (2xCTSes), but not at single/individual CTSes (1xCTSes) [144]. The functional effect of BORIS-DNA interactions is discussed in more detail in section 1.3.4 "BORIS DNA binding". The possibility of BORIS-BORIS homodimer formation was proposed by Jones *et al.* [156] and is supported by the finding of 2xCTSes [144]. However, to date, this interaction has not yet been experimentally validated. In a recent review, Soltanian and Dehghani emphasize the importance of identifying BORIS interacting partners to obtain a better understanding of the biological roles of BORIS in diverse cell types [149].

1.3.4 BORIS as a transcriptional regulator

BORIS-DNA binding

All the different ways through which BORIS can alter transcription rely on BORIS' ability to bind the DNA at specific binding motifs. Since BORIS and CTCF share a nearly identical ZF DNA binding domain, it was propose that the presence of both proteins, as observed in various tumors, leads to competition for shared DNA binding sites. It was predicted that this results in the replacement of CTCF by BORIS, thereby disrupting normal CTCF functions, leading to deregulation of various cellular processes that contribute to pathogenesis [139, 141]. In agreement with this hypothesis it was shown that in mouse embryonic stem cells 64% (3677 out of 5707) of BORIS binding sites overlap with CTCF binding sites. Interestingly, these analyses

revealed that BORIS mostly binds at CTCF sites near the promoter region of genes [153], supporting the importance of BORIS in gene expression regulation. Overall, CTCF binding was reduced at 1100 target sites in the presence of BORIS. Displacement of CTCF by BORIS was observed at specific sites, including the promoters of Stra8 and Prss50, in both mouse embryonic stem cells and cells isolated from testis [153], indicating a competition between BORIS and CTCF-DNA binding. The observation that BORIS and CTCF can form heterodimers [144, 181] raised the question if BORIS and CTCF indeed bind competitively or if they cooperate at DNA binding sites [144].

Recent findings have greatly contributed to our understanding of BORIS-DNA interactions, especially in the context of CTCF binding sites [144, 186]. Pugacheva *et al.* demonstrated that BORIS and CTCF binding patterns were very similar across different cancer cell lines despite their distinct tissues of origin. While BORIS and CTCF bind DNA with a similar affinity, only ~29-38% of CTCF binding sites were also occupied by BORIS (CTCF&BORIS sites) [144], which is much lower compared to the observation by Sleutels *et al.* mentioned above [153]. Interestingly, chromatin immunoprecipitation (ChIP)-re-ChIP assays in K562 and Delta47 cells revealed that these CTCF&BORIS sites can be co-occupied by both proteins simultaneously. This occurs at regions containing 2xCTSes [144], which were shown previously to exist in the *PRSS50* promoter [187]. It was proposed that 2xCTSes in BORIS negative cells are co-occupied by CTCF homodimers [144]. In agreement with the finding by Sleutels *et al.* that BORIS mostly binds in promoter regions [153], 2xCTSes occupied by both BORIS and CTCF were associated with promoters and active enhancers in K562 cells. In addition, a recent *in silico* analysis observed enrichment of BORIS binding at transcriptional start sites (TSS), regions bound by transcription factors, and regions of open, active chromatin. Furthermore, they demonstrated lack of BORIS binding in the case of a thymine at the first or fifth location of the CTCF consensus sequence [188]. For a representation of the BORIS and CTCF consensus sequence see **figure 5**. It is important to note that 2xCTSes cannot be identified by ChIP-seq experiments, but only by electrophoretic mobility shift assay, transcription factor footprinting techniques and ChIP-re-ChIP assays [144]. Overall, these findings shed light on BORIS binding specificity, though further studies are needed to investigate the functional effect of BORIS binding at CTCF&BORIS sites as well as single BORIS binding sites.

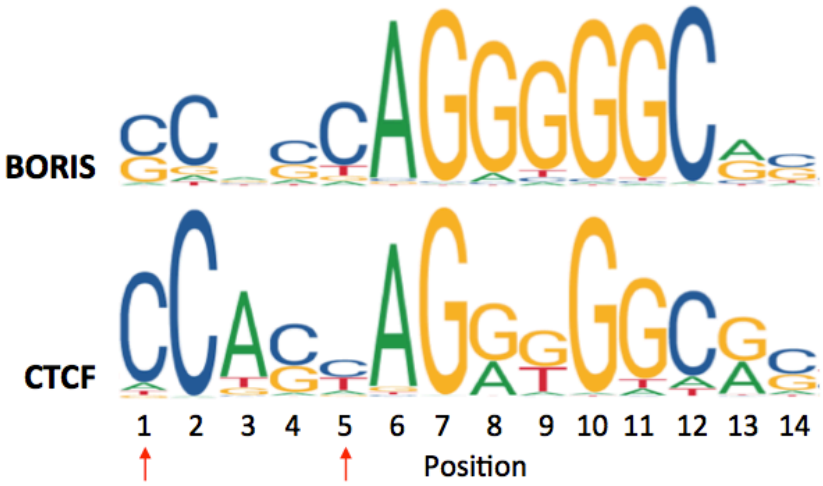


Figure 5. A comparison of the BORIS and CTCF consensus sequence.

The binding motifs of BORIS (matrix profile MA1102.1*) and CTCF (MA0139.1) were obtained from the Jaspar database of transcription factor binding profiles (7th release, 2018). Red arrows indicate the positions identified by Bergmaier *et al.* as crucial for BORIS DNA binding [188]. *The first position of the BORIS consensus sequence was not available in Jasper and therefore manually added according to the sequence identified by Sleutels *et al.* [153].

Transcriptional regulation

BORIS can directly change gene expression via its ability to bind in a gene promoter region and act as a transcriptional activator [151, 189-191] or recruit a transcriptional activator [182]. In addition, aberrant BORIS expression has been linked to various epigenetic changes that affect gene expression, including DNA demethylation [159, 189, 192] and activating histone modifications [189, 191-193]. The discovery of the BORIS protein in the testis demonstrated an association of BORIS expression with erasure of DNA methylation [139]. In cancer cells, changes in gene expression upon altered BORIS protein level also co-occur with DNA demethylation [189], especially at the promoter of CTAs [159, 192]. However, DNA demethylation is not always observed for BORIS-increased gene expression [151, 182, 192]. While BORIS can recruit chromatin modifiers to the DNA that can alter chromatin structure and accessibility [167, 184], it remains unknown how BORIS expression leads to changes in the DNA methylation status. In addition to changes in chromatin structure induced by chromatin remodelling proteins, BORIS is also implicated in chromatin looping [144, 186]. In the context of CTCF, Chromatin loops are DNA extrusions that form upon dimerization of two DNA bound CTCF proteins. Chromatin loops can have an effect on gene expression in various ways, including the separation of epigenetic marks within or outside a loop and increasing enhancer proximity to a gene promoter (for an extensive review read Holwerda and de Laat [194]). While it is likely that BORIS binding at 2xCTSes can alter CTCF-mediated chromatin looping, more studies are needed to experimentally validate this theory.

RNA expression

Currently, very few studies have investigated the effect of altered BORIS expression on genome wide gene expression in cancer cells [144, 153, 195, 196]. These include RNA-seq of the K562 myelogenous leukemia cell line upon BORIS knockdown [144, 195, 196] and of the MCF7 breast cancer cell line upon ectopic BORIS expression [144, 196]. In one study, both systems with altered BORIS expression resulted in differential gene expression of a large number of genes [144], indicating that BORIS expression can lead to genome-wide gene expression changes. Among the identified differentially expressed genes (DEGs) many were observed with a functional link to cancer [144, 195] as well as the inflammatory response and reproductive system disease [144]. Teplyakov *et al.* demonstrated that altered BORIS expression can change

the expression of non-coding RNAs (micro RNAs and Piwi-interacting RNAs) as well, though the differentially expressed non-coding RNAs were not linked to potential BORIS functions [195]. Besides mRNAs and non-coding RNAs, BORIS was implicated in the transcription of rRNA as discussed in section 1.3.3 "BORIS-interacting proteins".

CTA expression

Interestingly, BORIS is the first CTA found to regulate expression of other CTAs. Upon the discovery of BORIS, both Hong *et al.* and Vatolin *et al.* demonstrated activation of CTAs upon BORIS expression [151, 159]. Most BORIS-induced CTAs were identified by Vatolin *et al.* via transient expression of BORIS in NHDF cells, including LAGE-1, RAGE-2, Oct-3/4 and several members of the MAGE gene family [159]. Hong *et al.* showed that ectopic BORIS expression resulted in activation of the CTA gene NY-ESO-1, while BORIS downregulation reduced NY-ESO-1 promoter activity as assessed by luciferase assay in lung cancer cells. They revealed that BORIS mediates regulation of NY-ESO-1 via binding in the NY-ESO-1 promoter [151]. These findings were confirmed by other studies, which revealed even more CTAs that were activated upon BORIS expression [192, 197]. Since MAGE genes are among the CTAs activated in melanoma, Kholmanskikh *et al.* decided to investigate the relationship between BORIS and MAGE-A1 expression in melanoma. To their surprise, they observed no correlation between BORIS and MAGE-A1 expression in melanoma cell lines. In addition, ectopic BORIS expression did not lead to de-repression of MAGE-A1 expression, while treatment with the demethylating agent 5-aza-dC did. Furthermore, they demonstrated that MAGE-A1 expression is not increased upon ectopic BORIS expression in a sarcoma cell line, immortalized human keratinocytes or normal human fibroblasts [138]. To obtain a better insight into the effect of BORIS expression on CTA activation in melanoma further studies are needed.

Allele-specific expression

The observation that CTCF is involved in epigenetic regulation of imprinting and X-chromosome inactivation (XCI), which both result in allele-specific expression, led to the idea that BORIS might play a role in this process [151, 174]. It has been well established in the literature that binding of CTCF to specific DNA sites can block the interaction between regulatory elements, like enhancers (enhancer blocking), or block the spread of repressive

heterochromatin into regions of transcriptionally active chromatin and vice versa (boundary) [140, 198]. One of the main examples of CTCF dependent enhancer blocking is at the H19/Igf2 locus. At this locus the H19 and Insulin growth factor 2 (Igf2*)* genes are separated by a region that is differentially methylated in a parent specific manner. This region is also known as the H19 imprinting control region (ICR). Under normal conditions, the H19 ICR is methylated on the paternal allele and unmethylated on the maternal allele. In this scenario, CTCF can bind within the H19 ICR on the maternal allele and block enhancer interaction with the *Igf2* promoter, resulting in repression of *Igf2* and activation of *H19* expression on the maternal allele [198-201]. BORIS was shown to bind the H19 ICR in mouse testis, though the precise binding sequence was not investigated [184]. Nguyen *et al.* showed BORIS binding in the H19 ICR in HCT116 colon cancer cells and revealed preferential binding of BORIS to the methylated H19 ICR. However, knockdown of CTCF using small interfering RNA (siRNA) resulted in binding of BORIS to the unmethylated H19 ICR, suggesting that BORIS binding to the H19 ICR is methylation-independent [202]. This finding was challenged by Pugacheva *et al.* who observed binding of various BORIS isoforms only within the unmethylated H19 ICR [146]. Both studies suggest a competition between BORIS and CTCF for binding the unmethylated allele. While these findings suggest a role for BORIS as transcriptional regulator at the H19 ICR, the effect of BORIS expression on H19/Igf2 expression in cancer cells remains to be investigated

CTCF's role as insulator is especially crucial on the X-chromosome, specifically on the X-chromosome allele that becomes silenced during XCI. This allele is referred to as the inactive X-chromosome (Xi), also known as the Barr body, and is marked by expression of the non-coding RNA X-inactive specific transcript (*XIST)*, DNA methylation and repressive histone marks. XCI takes place to equalize expression of X-chromosome-linked (X-linked) genes between females (XaXi) and males (XaY). However, a subset of X-linked genes are located in regions of active chromatin and can escapes from the overall repressive chromatin environment of the Xi [203]. CTCF was shown to be involved at various steps during the process of XCI [204-210]. Furthermore, CTCF binding on the Xi is clustered at genes that are known to escape from the transcriptionally inactive status of the Xi [207, 211], indicating a role for CTCF as an insulator protein in maintaining boundaries on the Xi between active and repressive chromatin. In accordance with this idea, a study by Kemp *et al.* [212] showed that endometrial tumors with

heterozygous loss of CTCF display large-scale loss of DNA methylation on the X-chromosome, further demonstrating the importance of CTCF for maintenance of the inactive state of the Xi. Thus far, no studies have investigated the effect of BORIS expression on the transcriptional repressive state of the Xi.

1.3.5 BORIS functions in cancer

Various studies have attempted to elucidate the role of BORIS in cancer. This has revealed involvement of BORIS in a wide variety of cellular processes [142, 149], which is not surprising considering its large-scale effect on gene expression. These findings indicate that BORIS, like CTCF, is a multifunctional protein. Currently, BORIS has no clearly identified function in cancer as a result of inconsistent findings, which is possibly due to cell type specific and context dependent roles for BORIS.

Cell growth

Multiple studies have looked into the effect of BORIS expression on cell growth. Ectopic BORIS expression resulted in reduced colony formation and inhibition of proliferation in both primary cells and various cancer cell lines, which was similar to the effect observed for CTCF [213]. This contradicts the idea that BORIS may act opposite to CTCF's tumor suppressor function in cancer cells. Surprisingly, one study found that both ectopic BORIS expression and BORIS knockdown by shRNA resulted in reduced colony formation in HEK293T cells [157]. Smith *et al.* and Gaykalova *et al.* expressed a low level of BORIS using a tetracycline inducible expression plasmid and observed an increase in growth and colony formation, while this effect was not seen with high ectopic BORIS expression [189, 197].

Genomic stability

The role of BORIS in genomic stability is another topic on which a consensus has not yet been reached. On the one hand, ectopic expression of BORIS in both primary cells and cancer cell lines had no effect on the basal level of apoptosis and decreased apoptosis upon UV exposure, similar to findings for ectopic CTCF expression [213]. On the other hand, overexpression of BORIS provoked an accumulation of cells in the S-phase and was associated with polyploidy and cell size increase, indicating genomic instability [157].

Epithelial-mesenchymal transition

BORIS has been linked to the process of EMT in breast and neuroblastoma cell lines [214, 215]. Again, in these different cell types, the effect of BORIS expression influences EMT in opposite direction. More specifically, Alberti *et al.* demonstrated reduced expression of epithelial markers and increased expression of mesenchymal markers as well as increased migration upon BORIS silencing in MCF7 cells [214]. Conversely, Garikipati *et al.* revealed downregulation of mesenchymal markers upon BORIS knockdown, while these markers were upregulated upon ectopic BORIS expression in isolated cancer stem cells from a MYCN amplified neuroblastoma cell line. In addition, they showed reduced migration and invasion with BORIS silencing and increased migration and invasion with BORIS overexpression [215].

1.3.6 BORIS vaccine

Given the initial observation of BORIS as a CTA, Loukinov *et al.* were the first to assess the potential of a BORIS vaccine for cancer treatment. To prevent occurrence of the tumorigenic properties associated with wild-type BORIS, a truncated mBORIS lacking the DNA-binding ZF domain was generated. A combination of DNA- and adenoviral-based mBORIS vaccine in the presence or absence of CD80 molecular adjuvant were used for immunization of the mouse 4T1 breast cancer model, which demonstrated a delay in tumor growth, decreased tumor size and longer survival post tumor challenge [216]. In their next studies, T cell responses were assessed upon immunization with the DNA-based mBORIS vaccine mixed with the molecular adjuvants interleukin 12 and 18, which demonstrated $CD4^+$ T cell proliferation, generation of $CD8^+$-cytotoxic T cells, inhibition of 4T1 tumor growth and prolonged survival of immunized mice upon tumor implantation [160, 217]. In addition, splenocytes isolated from mice immunized with the DNA-based vaccine were capable of killing 4T1 mouse mammary carcinoma, P815 mastocytoma and GL261 glioma cell lines, while this result was not obtained with a protein-based vaccine [160]. Taking the use of BORIS vaccine one step further, Mkrtichyan *et al.* investigated the effect of immunization with BORIS vaccine on existing tumors and metastasis [218]. This time, the vaccine consisted of antigen presenting dendritic cells loaded with recombinant mBORIS protein, which demonstrated similar T cell responses as reported previously [160, 217]. Furthermore, immunization post 4T1 tumor cell implantation resulted in the development of significantly smaller tumors and less metastasis [218].

THESIS RATIONALE AND OBJECTIVES

Melanoma is among the most aggressive cancers, due to its tendency to metastasize very early during tumor development, leading to life-threatening conditions. While surgery may be curative if the primary melanoma is discovered early, metastatic melanoma does not respond well to current therapies on the long term and the 5-year survival rate is only 15-20% [219]. CTAs are considered promising targets for immunotherapy due to their ability to elicit immune responses specifically directed at cancer cells [99]. BORIS is considered a CTA and is involved in processes underlying cancer progression [142, 149]. Immunization of mice bearing tumors and metastasis using dendritic cells loaded with a mutant BORIS protein demonstrated T cell responses, resulting in smaller tumors with less metastasis [218]. This highlights the potential for the use of BORIS-based vaccines. BORIS is expressed in melanoma and, while BORIS is proposed to play a role in melanoma progression [138], its biological functions in melanoma remain unknown.

The overall goal of this thesis was to obtain a better understanding of the role of BORIS in melanoma by identifying processes and pathways through which BORIS could promote melanoma progression. To achieve this, we set two objectives.

The first objective was based on previous work in the laboratory, which demonstrated that BORIS could alter expression of X-linked genes in melanoma cells [220]. Importantly, females with melanoma tend to have a better prognosis than males. One underlying reason for this sex-related effect could be differences in gene expression from the X-chromosome [221]. Our laboratory recently demonstrated that the Xi is often disrupted in melanoma [222]. Furthermore, multiple studies indicate an important role for CTCF in maintenance of the Xi [207, 211, 223]. While we have demonstrated that BORIS alters expression from the X-chromosome, it remains unknown if this is due to differential expression from the Xa or Xi. We hypothesize that BORIS specifically alters expression from the Xi by disrupting the Xi chromatin state (**Fig. 6**). To address the hypothesis, the objective was to modulate BORIS expression in a XaXi female melanoma cell line and assess allele-specific expression by exploiting single nucleotide polymorphisms (SNPs).

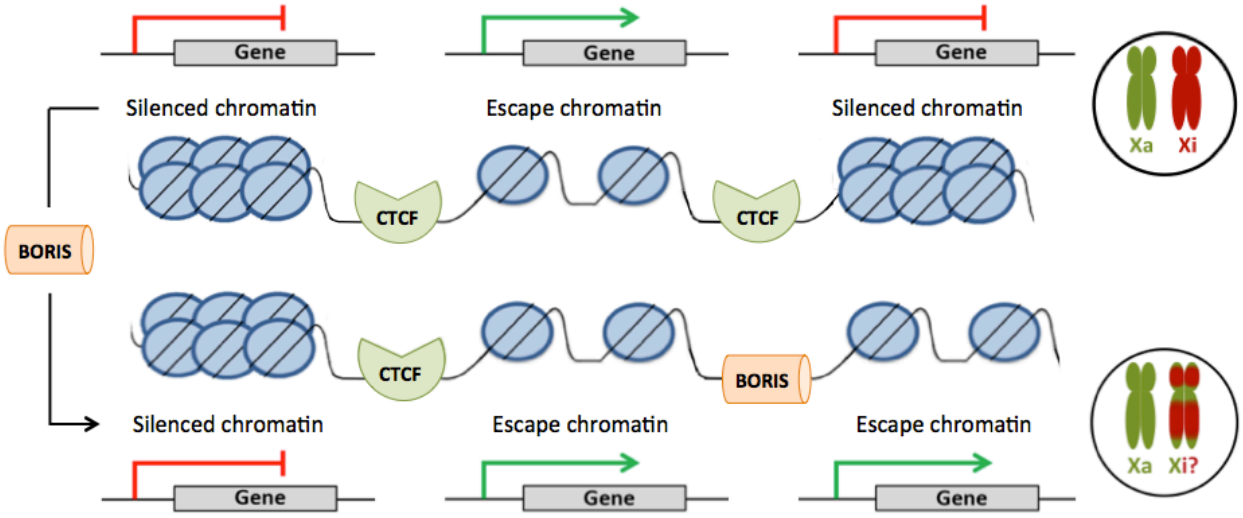


Figure 6. Schematic representation of the hypothesis for objective 1: Investigation of the effect of BORIS expression on allele-specific expression from the inactive X-chromosome. While BORIS is absent in normal female cells (top), female melanoma cells often express BORIS (bottom). BORIS can bind CTSes and, possibly via recruitment of specific interacting proteins, interfere with chromatin boundaries between silenced and escape chromatin, leading to aberrant expression of X-linked genes and disruption of the Xi. Figure adapted from Chaligné *et al.* and Berletch *et al.* [203, 223].

The second objective aimed to identify biological processes and pathways in melanoma that BORIS is involved in. Previous studies demonstrated that in cancer cells BORIS plays a role in various cellular processes [142, 149]. BORIS achieves this by altering gene expression [144, 195, 196] and/or by interacting with different protein partners [144, 167, 181-184]. All the different ways through which BORIS can alter transcription rely on BORIS' ability to bind the DNA at specific binding motifs [144, 153, 186]. For example, BORIS can directly change gene expression via its ability to bind in a gene promoter region and act as a transcriptional activator [151, 189-191] or recruit a transcriptional activator [182]. BORIS-DNA binding was mainly observed at regions of open and active chromatin that are associated with active histone modifications [144, 188]. To identify biological processes that BORIS is involved in and address how BORIS regulates these processes we aimed to use a multi-omics approach, including 1) the identification of BORIS-interacting proteins by mass spectrometry (MS), 2) the analysis of BORIS-induced transcriptional changes using RNA-seq, 3) the assessment of BORIS binding in gene promoter regions by ChIP followed by qPCR, and 4) the analysis of BORIS-induced changes on chromatin accessibility using the assay for transposase accessible chromatin followed by sequencing (ATAC-seq) (**Fig. 7**).

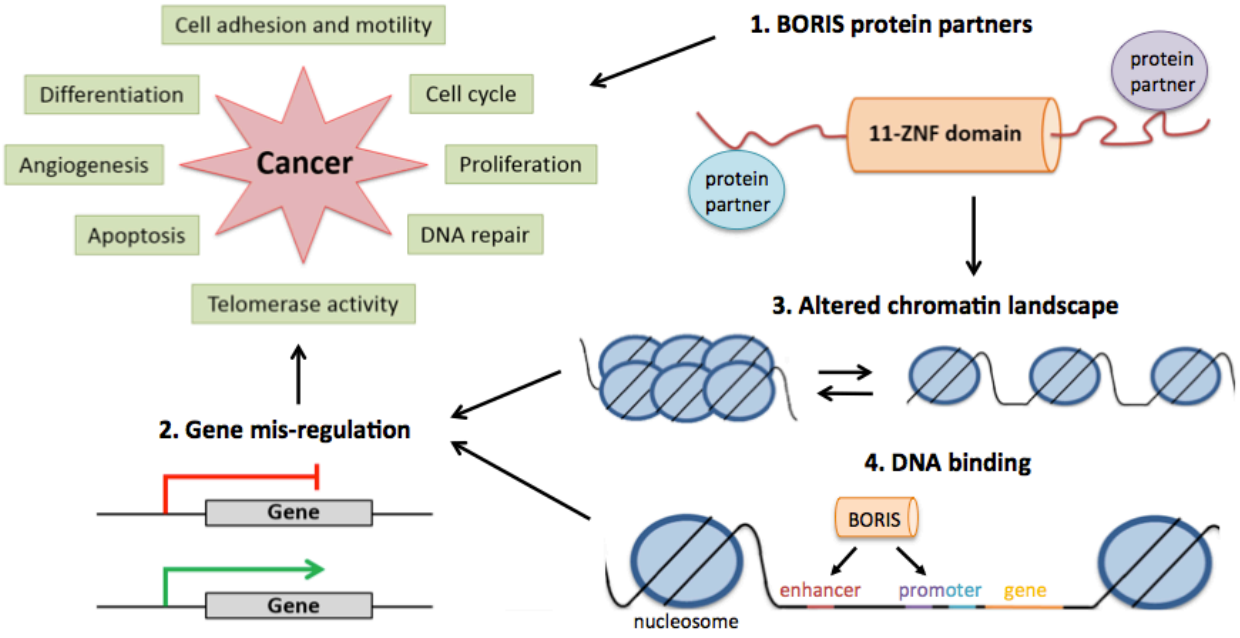


Figure 7. Schematic representation of objective 2: Identification of biological processes that BORIS is involved in and address how BORIS regulates these processes. To identify biological processes that BORIS is involved in and promote melanoma progression our objective is to examine BORIS protein partners by MS (1) and assess the effect of altered BORIS expression on transcriptional changes using RNA-seq (2). In addition, we aim to unravel the regulatory mechanisms behind BORIS-induced transcriptional changes by investigating BORIS-DNA binding at gene promoters using ChIP-qPCR (3) and determine if altered BORIS expression can induce changes in chromatin accessibility using ATAC-seq (4). We believe that potential alterations in chromatin accessibility could be due to BORIS-mediated recruitment of specific proteins involved in chromatin remodeling.

CHAPTER 2

-

MATERIALS AND METHODS

2.1 GENERATION OF EXPRESSION VECTORS

2.1.1 pLenti-His-Katushka

To construct a lentiviral vector expressing 6xHis-tag fused to Katushka (Red-shifted derivative of red fluorescent protein from Entacmaea quadricolor), we used the pTurboFP635-N vector as template and amplified a 755bp PCR fragment with the primers HK-F and HK-R (All primer sequences are listed in **Appendix I**). The resulting fragment was digested and cloned into pcDNA3-neo at HindIII-BamHI restriction sites. The final pLHK lentiviral vector was obtained by subcloning a fragment from the aforementioned plasmid (1048bp between SnaBI and BamHI) into pGIPZ-lentiviral non-silencing vector after removing the Turbo green fluorescent protein (GFP) - internal ribosome entry site (IRIS)-PuroR expressing sequence. This vector is referred to as HK-EV (empty vector). Full-length hBORIS cDNA was amplified by PCR on RNA extracted from HEK293T cells using iProof (Bio-Rad, 172-5301) with primers containing BamHI and MluI restriction sites, BORIS-2F-BamHI and BORIS-663R-Stop_MluI. The pLHK vector was cut using the fastdigest restriction enzymes BamHI (Fermentas) and MluI (Fermentas). The PCR product was ligated into the pLHK vector using T4 DNA ligase (Fermentas, EL0014), resulting in pLHK-BORIS, which is referred to as HK-BORIS.

2.1.2 pLenti-GFP-IRES-Flag

The lentiviral pLenti-TurboGFP-IRES-FLAG®-tag vector was obtained by subcloning a143bp sequence coding for one FLAG®-tag and a multiple cloning site into pGIPZ-lentiviral non-silencing vector after removing the 1048bp corresponding to the PuroR and shRNA expressing sequence. This resulted in the pLGIF vector, which is referred to as GIF-EV. Full-lenght cDNA of hBORIS and hCTCF were amplified by PCR on RNA exptracted from HEK293T cells using iProof (Bio-Rad, 172-5301) and inserted into pLGIF in the same way as specified above for pLHK, and referred to as GIF-BORIS and GIF-CTCF.

2.1.3 pInducer20

The plasmids pCMV6-mBORIS-Myc-DDK (Origene, MR222222) and pLHK-BORIS were amplified using iProof (Bio-Rad, 172-5301) with the primers mBORIS-F-NcoI, mBORIS-636R-AgeI, hBORIS-F-NcoI and hBORIS-663R-AgeI to obtain mBORIS and hBORIS, respectively.

The PCR products were digested with NcoI (Fermentas FastDigest) and AgeI (Fermentas FastDigest) and inserted into pDONR221 (Invitrogen cat # 12536017) by a BP reaction (Gateway™ BP Clonase™ II Enzyme mix, Invitrogen Cat # 11789020) to generate pENTR-hBORIS and pENTR-mBORIS with six C-terminal His-tags followed by a stop codon. The entry clones were recombined into the lentiviral destination vector pI20 (EV-6xH, a gift from Stephen Elledge [224], Addgene plasmid #44012) using Gateway LR Clonase™ II (Fisher, 11791020) to generate pI20-mBORIS-6xH and pI20-hBORIS-6xH. We will refer to these vectors as EV-6xH, mBORIS-6xH and hBORIS-6xH.

We used Gateway Cloning Technology to generate inducible lentiviral vectors that we called pIN-hBORIS-3xFlag and pIN-mBORIS-3xFlag. Briefly, a cDNA fragment of full-length hBORIS and mBORIS tagged at its C-terminal with three FLAG®-tags was amplified by PCR using iProof (Bio-Rad, 172-5301) and the following primers to introduce *attB* sites to the ends of hBORIS-3xFlag template: attB-BORIS-1F, and attB-mBORIS-MF_R-Stop, and attB-3Falg_R-Stop. The resulting PCR products were purified by QIAEXII Gel extraction kit (Qiagen Cat # 20021) followed by a BP reaction (Gateway™ BP Clonase™ II Enzyme mix, Invitrogen Cat # 11789020) with pDONR221 (Invitrogen cat # 12536017) to generate the pENTR-hBORIS-3xF and pENTR-mBORIS-3xF vectors. We then performed LR reactions (Gateway™ LR Clonase™ Enzyme mix, Invitrogen Cat # 11791019) using the pENTR-hBORIS-3xF, pENTR-mBORIS-3xF vectors and the pI20 to generate the final pIN-hBORIS-3xF and pIN-mBORIS-3xF. The pIN-Fluc-3xF (expressing Firefly luciferase tagged at its C-terminal with three FLAG®-tag) vector was obtained in a similar way and used as control. These vectors will be referred to as EV-3xF, mBORIS-3xF and hBORIS-3xF

2.1.4 pHaloTag®

The HaloTag® (HT) Control Vector (Cat # G6591) and pHTC HaloTag® CMV-neo Vector (Cat # G7711) were obtained from Promega. The vectors pHTC-BORIS and pHTC-CTCF were constructed by inserting the full-length coding sequence of either BORIS or CTCF between NheI and EcoRI restriction sites of pHTC HaloTag® CMV-neo Vector. PCR on RNA extracted from HEK293T cells was used to amplify cDNA sequences for BORIS and CTCF using the following

oligos: BORIS-1F_NheI, BORIS-663R_EcoRI, CTCF_1F-NheI and CTCF_727R-EcoRI. We will refer to these vectors as HT-EV, HT-BORIS and HT-CTCF

We used Sanger sequencing to verify the sequence of all vectors.

2.2 CELL CULTURE AND DRUG TREATMENTS

Early passage melanoma cell lines (MM102, MM117, MM073, MM079, MM071, MM057, MM111, MM050 and MM074) and Giant Congenital Nevi cell lines (GCN004, GCN008 and GCN011) were kindly provided by Dr. Ghanem Ghanem (Institut Jules Bordet, Brussels, Belgium) and cultured in Ham's F10 medium (Wisent, 318-050-CL). The following cell lines were cultured in Roswell Park Memorial Institute 1640 medium (Wisent, 350-000-CL): SK-MEL-2, SK-MEL-5 and SK-MEL-28 cell lines purchased from ATCC, the Wistar Melanoma (WM) cell lines WM164, WM983A, WM983B, WM852 and WM793 obtained from the Dr. Meenhard Herlyn (Wistar Institute, Philadelphia, PA, USA), A375P, A375M and M24Met received from Dr. Dirk Ruiter (Radboud University Medical Center, Nijmegen, The Netherlands), and K562 cells gifted by Dr. Sonia del Ricón (Lady Davis Institute for Medical Research, Montréal, QC, Canada). HEK293T cells, purchased from ATCC (CRL-3216), and mouse embryonic fibroblasts (MEFs), generated in our laboratory from E13.5 day post coitum embryos, were cultured in Dulbecco Modified Eagle Medium (Wisent, 319-005-CL). For all cell lines the medium was supplemented with 8% heat-inactivated fetal bovine serum (Wisent, 080-450), 1% penicillin-streptomycin (Wisent, 450-201-EL) or 0.1% gentamycin (Wisent, 450-135-XL) and 1% GlutaPlus (Wisent, 609-066-EL). Inducible cell lines were selected by the addition of G418/Neomycin (Wisent, 450-130-QL); MM057 150µg/ml; MM074 200µg/ml) for a minimum of 2 weeks or until non-infected control cells were killed. Upon selection, MM057 and MM074 cell lines were maintained in the presence of 50µg/ml G418. Doxycycline (dox) (Clontech, 631311) was added to the medium for the indicated dose and time, and refreshed every 48 hours. For the inhibition of poly ADP ribose polymerase (PARP), cells were treated with 1µM or 10µM Olaparib (a kind gift from Dr. Amber Yasmeen) for 72 hours. To induce demethylation of DNA CpG sites, cells were treated with 5µM 5-aza-dC (Abcam, ab120842) for 96 hours. Bimonthly tests for mycoplasma demonstrated the absence of contamination. All cultures were maintained at 37°C in a 5% CO_2 humidified atmosphere.

2.3 TRANSFECTION, VIRUS PRODUCTION AND INFECTION

For virus production, lentiviral expression vectors (10µg DNA) were transfected into HEK293T cells along with the viral packaging plasmids pSPAX2 (7.5µg DNA) and pMD2.G (5µg DNA) using polyethylenimine (PEI; Polysciences, 23966, DNA:PEI ration between 1:2 and 1:3) and Opti-MEM transfection medium (Invitrogen, 31985-070). Cells were seeded in a 100mm dish at 30% confluency 24 hours before transfection. The medium was replaced by fresh medium 3-4 hours after transfection. At 48-72 hours post transfection virus-containing supernatant was harvested, centrifuged, passed through a 0.45µm syringe filter (VWR, CA28145-505), and either used directly for infection or stored at -80^{0}C. Melanoma cells and MEFs were infected twice by incubation with a mixture of viral supernatant, medium (1:1 ratio) and polybrene (5-8ug/ml, Sigma-Aldrich, H9268) for 12 hours. The viral mixture was replaced by fresh medium for 24 hours before selection or fluorescent activated cell sorting. Expression of the constructs was verified using a combination of IF, immunoblot (IB) and qPCR.

2.4 DNA/RNA FLUORESCENCE *IN SITU* HYBRIDIZATION (FISH)

DNA FISH was performed for visualization of the X-chromosome and RNA FISH for *XIST*. Cells were allowed to attach to microscope slides (Fisherbrand Superfrost®Plus, 12-550-15), washed with PBS for 5 minutes and fixed using 3.7% formaldehyde for 10 minutes at room temperature. After rinsing with PBS cells were permeabilized in cold 70% ethanol for a minimum of 1 hour at 4°C and incubated in wash buffer (1x saline sodium citrate (SSC) and 10% formamide in ddH_2O) for 5 minutes at 37°C. A mixture of 125nM probe (Xist probe, Stellaris; cepX, Abbott) and hybridization buffer (25mg/ml dextran sulfate, 2x SSC, and 10% formamide in ddH_2O) was added and covered with coverslips. Hybridization was performed in a hybridizer at 37°C overnight. The next day coverslips were removed and slides were incubated in wash buffer for 30 minutes at 37°C. Upon incubation of the slides in 2x SSC for 10 minutes at room temperature, they were allowed to air dry in the dark and coverslips were mounted using mounting medium containing 4',6-diamidino-2-phenylindole (DAPI; DAPI II Counterstain, Abbott, 06J50). Images were captured with a Zeiss AX10 microscope equipped with a CV-M4+ monochrome megapixel camera (Jai Inc). The number of nuclear signals was counted to determine the number of X chromosomes (cepX) and the number of inactivated X chromosomes (*XIST*) in at least 100 cell nuclei.

2.5 MOUSE EMBRYONIC FIBROBLAST CELL LINES (MEFs)

2.5.1 Generation of MEFs

E13.5 day post coitum embryos from a cross between a male C57BL/6J and female FVB/NJ mice were kindly provided by the laboratory of Dr. Koromilas. Head, tail and blood organs were removed and the remaining embryo was finely minced using a blade and incubated in Trypsin/EDTA (Wisent, 325-043-CL) for 30min at 37°C. Culture medium was added and dissociated tissue was transferred to a cell culture dish and maintained at 37°C in a 5% CO_2 humidified atmosphere. To generated immortalized MEFs, cells were cultured for several passages until spontaneous immortalization occurred. Single cell clones were obtained by fluorescence activated cell sorting (FACS) of single cells into a 96 well plate.

2.5.2 Genotyping

Genomic DNA (gDNA) was extracted from the cells according to the following steps: cells were washed and lysed (10mM Tris-HCl pH7.5, 10mM EDTA, 10mM NaCL and 0.5% Sarkosyl in ddH_2O supplemented with proteinase K) overnight at 37°C. The next day, cell lysates were centrifuged at 2000rpm for 5 minutes, supernatant removed and DNA precipitated with ice cold NaCl/ethanol (75mM NaCl in 100% ethanol) for 2 hours at room temperature. Precipitated DNA was centrifuged for 15 minutes at 4000rpm at 4°C, supernatant removed and pellets were washed three times with 70% ethanol. The pellets were allowed to air dry for 3 hours and solubilized in Tris-HCl pH8.0 at 37°C overnight. Positive (male mouse) and negative (female mouse) control gDNA was kindly provided by Dr. Alicia Bolt (Lady Davis Institute). The gDNA was stored at -20°C. Before proceeding with the PCR, gDNA was thawed and mixed well. End-point PCR was carried out in a 25µl reaction using 0.5µl 5U/µl iTaq™ DNA polymerase (Bio-Rad, 170-8870), 1x iTaq™ PCR buffer, 0.5µl 10mM dNTPs, 0.75µl 50mM MgCl, 100nM primers and 20-60ng gDNA. Primer sequences for SRY and FABP2 are provided in **appendix I**. Amplification was 95°C for 3 minutes, then 40 cycles of 95°C for 30 seconds, 60°C for 30 seconds and 70°C for 30 seconds followed by plate read on a Bio-Rad C1000 Thermal Cycler. PCR products were visualized on an agarose (Bio-Rad, 1613101) gel.

2.6 SINGLE NUCLEOTIDE POLYMORPHISM (SNP) PROFILING

2.6.1 SNP detection in the MM057 melanoma cell line

Putative SNPs were identified from the SNP database within the UCSC genome browser (dsSNP build 141 Found in >= 1% of Samples) and only SNPs with a minor allele frequency of >15% were included. To validate the presence of these SNPs in the MM057 cell line, gDNA was extracted with the QIAamp DNA minikit (Qiagen, 51304) according to manufacturer's protocol. To assess allele-specific expression at the RNA level, total RNA was extracted using TRIsure™ (Bioline, BIO-38032) or the Nucleospin RNA kit (Macherey Nagel, 740955.5) and 400ng RNA was reverse transcribed using the iScript cDNA synthesis kit (Bio-Rad, 170-8890) according to manufacturer's specifications. Next, gDNA (400ng for initial SNP validation and 115ng for SNP validation in the cells with stable HK-EV and HK-BORIS expression) and cDNA (4µl of ¼ diluted) were amplified using iProof™ High-Fidelity DNA polymerase (Bio-Rad, 1725330) with the indicated primers (**Appendix I**) in a final volume of 25µl. PCR was performed for either 25 or 40 cycles on a Bio-Rad C1000 Thermal Cycler with the following cycling conditions: 98ºC for 3 minutes, 98ºC for 30 seconds, 60ºC for 30 seconds, 70ºC for 40 seconds and a final extension at 70ºC for 5 minutes.

2.6.2 SNP detection in the mouse embryonic fibroblast cell lines

Putative SNPs between the C57BL/6J and FVB/NJ mouse genetic backgrounds were identified from the Mouse Genome Informatics (MGI) database (dbSNP Build 137, http://www.informatics.jax.org/snp) [225] and can be found in **Appendix II**. For the SNP in *Xist*, both total RNA and gDNA were extracted using TRIsure™ (Bioline, BIO-38032) following manufacturer's instructions. An extra precipitation and cleanup of the gDNA was performed, using 70% Ethanol. Total RNA (400ng) was reverse transcribed using the iScript cDNA synthesis kit (Bio-Rad, 170-8890) according to manufacturer's specifications. Each 25µl reaction contained 150ng gDNA or 5µl cDNA (1/5 diluted) and was amplified using iProof™ High-Fidelity DNA polymerase (Bio-Rad, 1725330) with the primers indicated in **appendix I**. PCR was performed for 25 cycles on a Bio-Rad C1000 Thermal Cycler with the following cycling conditions: 98ºC for 3 minutes, 98ºC for 40 seconds, 60ºC for 40 seconds, 70ºC for 40 seconds,

98ºC for 20 seconds, 60ºC for 20 seconds, 70ºC for 20 seconds and a final extension at 70ºC for 5 minutes.

For all the above conditions, PCR products were visualized on an agarose (Bio-Rad, 1613101) gel, extracted with the QIAEX II Gel Extraction Kit (Qiagen, 20021) and submitted for Sanger sequencing at the Genomics platform of the Institute for Research in Immunology and Cancer (IRIC, Montréal, Canada). All fragments were sequenced from both strands.

2.7 RNA EXTRACTION, REVERSE TRANSCRIPTION AND QUANTITATIVE POLYMERASE CHAIN REACTION (QPCR)

Total RNA was extracted using the Nucleospin RNA kit (Macherey Nagel, 740955.5) according to manufacturer's instructions. DNase treatment excluded gDNA. RNA quantity was determined by NanoDrop or Qubit and quality was assessed by agarose (Bio-Rad, 1613101) gel upon incubation of 500ng RNA at 72 ºC for 2 minutes. For RNA-seq samples the quality and quantity were assessed on a Bioanalyzer (Agilent Technologies). RNA integrity numbers higher >8.9 were obtained (average for 12 samples is 9.6). Total RNA (400ng) was reverse transcribed using the iScript or qScript cDNA synthesis kit (Bio-Rad, 170-8890 and Quanta, 95047-100, respectively) according to manufacturer's specifications. qPCR was performed with SsoFast EvaGreen Supermix or PerfeCTa SYBR Green FastMix (Bio-Rad, 172-5201 and Quanta, 95072-012, respectively) on a CFX96 Touch™ (Bio-Rad, 1855196). Each 10µl reaction contained 4µl cDNA (1/5-1/10 diluted), 5µl PCR mix and 400nM forward and reverse primer. The thermal cycling program used was as follows: 30 seconds at 95°C followed by 40 cycles of 5 seconds at 95°C and 5 seconds (SsoFast protocol) or 20 seconds (PerfeCTa protocol) at 60°C, followed by 10 seconds at 95°C and a melt curve of 5 seconds at 65°C and 5 seconds at 95°C. Primers were designed to be exon spanning and when possible the forward or reverse primer was designed to span an exon-exon boundary. For each primer pair, the melting temperature, GC content and the secondary structure of the PCR product were assed using tools available from the University of California Santa Cruz Genome Browser (https://genome.ucsc.edu/cgi-bin/hgPcr) or Integrated DNA Technologies (https://www.idtdna.com/calc/analyzer and https://www.idtdna.com/UNA-Fold). Melting curve analyses were carried out to ensure product specificity. Primers were purchased from Integrated DNA Technologies and sequences are provided in **appendix I**.

Importantly, BORIS primers span exon 9 and 10, and were designed to amplify isoform family's sf1 and sf5 [146]. Samples were run in technical triplicate and biological duplicate or triplicate, and data were normalized to both *HPRT1* and *TBP*. All data was analyzed using the CFX Manager™ Software (Bio-Rad).

2.8 FLOW CYTOMETRY ANALYSIS

2.8.1 Cell cycle analysis

200,000 cells were seeded in a 6 well plate and treated with or without 50ng/ml dox the next day. Dox was refreshed every 36-48 hours. After five days, 500,000 cells were collected, washed with cold PBS, fixed with 70% ethanol and stored at -30°C. On the day of staining, cells were pelleted, ethanol was removed and the cells were washed with PBS. Next the cells were treated with 50µl of 100µg/ml DNase free RNase A (Sigma, R6148) and incubated with 200µl of 50µg/ml propidium iodide (PI, Invitrogen) for 30 min at 37°C in the dark. Samples were stored at room temperature in the dark until flow cytometry analysis. Samples were acquired within 6 hours after staining on a FACSCanto™ flow cytometer (BD Biosciences) and data was analyzed using ModFit LT™ software (Macintosh).

2.8.2 Apoptosis assay

200,000 cells were seeded in a 6 well plate and treated with or without 50ng/ml dox the next day. Dox was refreshed every 36-48 hours. After five days, 500,000 cells and their supernatant was collected and washed with cold PBS. Cells were stained for AnnexinV and PI by incubation in 100ul staining solution (1µl APC-conjugated AnnexinV (eBioscience, 88-8007-72), 0.1µg PI (Invitrogen), 1x binding buffer) for 15 minutes at room temperature in the dark. Next, cells were resuspended in 1x binding buffer and results were immediately acquired on a FACSCalibur™ flow cytometer (BD Biosciences) and analyzed using *FlowJo*® software (version 10.4.1).

2.9 PROLIFERATION ASSAY

For cell proliferation, cells were seeded in a 96 well plate as follows. Day 0: 15000, day 3: 5000, day 5: 4000, and day 7: 2000. The following day, cells were treated with 12.5, 25, 50, 100 or 200ng/ml dox (Clontech, 631311) for the indicated number of days and fixated using 4%

formaldehyde (Fisher, F79-500) for 20 minutes at room temperature. Next, cells were incubated with 0.05% crystal violet (Sigma-Aldrich, HT90132) at room temperature and washed 3 times with ddH_2O to remove excess crystal violet. Once the plates dried, crystal violet was solubilized using 150µl 0.2% Triton X-100 (Bioshop, TRX777.100) overnight and absorbance was measured using the FLUOstar OPTIMA (BMG Labtech) at 570nm.

2.10 3-(4,5-DIMETHYLTHIAZOL-2-YL)-2,5-DIPHENYLTETRAZOLIUM BROMIDE (MTT) ASSAY

Cell viability was assessed using the MTT assay. Cells were seeded in 96-well culture plates and treated as indicated. After treatment, culture medium was removed and 120µl of 5mg/ml MTT solution was added to each well and incubated for 3 hours at 37°C. The solution was discarded and formazan crystals were dissolved in 120µl solvent (4mM HCl and 0.1% NP-40 in isopropanol) by shaking for 15 minutes. Absorbance was measured using the FLUOstar OPTIMA (BMG Labtech) at 570nm.

2.11 MIGRATION AND INVASION ASSAY

In vitro cell migration and invasion were evaluated in transwell assays. Cells were treated with or without 100ng/ml dox for five days, counted and seeded at 5×10^4 (MM057) or 1×10^5 (MM074) cells in serum-free medium in the upper transwell insert chamber (PET membrane, 12-well, 8-µm pores; Corning Costar). For the invasion assays, the insert chamber was coated with Matrigel® (Corning) diluted 1 in 5 with cold, serum-free medium. Complete medium (900µl) with 8% FBS was added to the lower chamber and cells were incubated for either 16 hours (MM057) or 72 hours (MM074) at 37°C in a 5% CO_2 humidified atmosphere. The cells that did not migrate through the pores were removed from the top of the transwell membrane using a cotton swap. Migrated cells were fixed with 3.7% paraformaldehyde for 10 minutes at room temperature and stained with crystal violet (0.5%). Images were captured of six fields per insert chamber using the EVOS FL Cell Imaging System (ThermoFisher Scientific). Cells were counted with the Cell Counter in ImageJ. Representative images were captured using an inverted light microscope (Olympus BX43). Three independent experiments were performed, each consisting of two technical replicates.

2.12 PROTEIN ANALYSIS

2.12.1 Immunoblot (IB) analyses

Cells were grown to 70% confluency and harvested in ice cold PBS by scraping. The collected cell pellets were lysed on ice with lysisbuffer (50mM HEPES-KOH pH7.4, 150mM NaCl, 1% triton X-100, 0.2% sodium-deoxycholate, 5mM EGTA, 10% glycerol) supplemented with 1x protease inhibitor cocktail (Promega, G652A) and phosphatase inhibitor (Roche, 4693124001), and sonicated 3 times 15 seconds at 25%. Lysates were cleared by centrifugation and protein concentration was determined using the BCA kit (Fisher, 23225) according to manufacturer's instructions. Lysates were denatured in SDS sample buffer for 5 min at 95°C. For direct lysis, $2*10^5$ cells were lysed in 1x SDS sample buffer and denatured for 10min at 95°C. Protein extracts were separated on SDS-polyacrylamide gel and transferred onto polyvinylidene difluoride membranes (Bio-Rad, 162-0177). Membranes were blocked with 3% nonfat milk (Bio-Rad, 170-6404) or 3% bovine serum albumin (Bioshop, ALB003) in PBS containing 0.05% Tween 20 (TBST, Bioshop, TWN510.500) and probed with the appropriate primary antibody overnight at 4°C. Blots were washed with TBST and probed with the corresponding horseradish peroxidase-conjugated secondary antibody. Proteins were detected using enhanced chemiluminescence reagent (Bio-Rad, 170-5060), images were captured with a ChemiDoc™ Imaging System (Bio-Rad) and analyzed with Image Lab™ Software (Bio-Rad Laboratories, Version 5.0 build 18).

2.12.2 Co-immunoprecipitation (Co-IP)

For the co-IP experiments, MM057 cells stably infected with hBORIS-3xF were treated with or without 100ng/ml doxycycline for five days, harvested by scraping and washed with cold PBS. Cell pellets were resuspended and lysed in cell lysis buffer (50mM HEPES-KOH pH 7.4, 150mM NaCl, 1% Triton-X100, 0.2% sodium deoxycholate, 5mM EGTA, 10% Glycerol, 2mM NaF, 1mM sodium pyrophosphate, 1mM beta-glycerophosphate, 2mM Na3VO4) supplemented with 1x protease inhibitor cocktail (Promega #G6521). Cell lysates were incubated on ice for 20 minutes followed by mechanical disruption 10x with a 1cc 27G syringe. Supernatants were separated from the membrane debris by centrifugation at 13,000rpm for 15 min at 4°C. Before the Co-IP, supernatants were pre-cleared with 30µl of Dynabeads Protein G (Invitrogen

#10003D) for 1 hour on a rotator at 4°C. After pre-clearing, 30µl of each sample was kept as an input control and 0.5mg protein from each supernatant was incubated with 5µg mouse anti-FLAG M2 (Sigma, F1804, clone M2) and 30µl Dynabeads Protein G at 4°C on a rotator overnight. The immune-complexes were subsequently washed 5 times with 1ml lysis buffer. At the end, proteins were eluted from the beads in 1x SDS loading buffer (3M Tris pH 6.8, 20% SDS, 100% glycerol, bromophenol blue in ddH_2O) by incubation at 95°C for 5 minutes, loaded into acrylamide gel and blotted as described in the section "Immunoblot analyses".

2.12.3 Immunofluorescence (IF)

MM057 cells expressing HK-BORIS or hBORIS-6xH were allowed to attach to glass cover slips (VWR, 48366067) in a 6 well plate. Cells were washed with 1x PBS and fixed for 15 minutes at room temperature with 3.7% formaldehyde (Fisher, F79-500). After three washes with PBS^{++} (PBS supplemented with 1mM CaCl, 1mM MgCl) fixation was quenched in 0.1M Glycine in PBS for 30 minutes at room temperature to reduce fixative-induced auto-fluorescence. Next, cells were washed with PBS^{++}, permeabilized (PBS^{++} supplemented with 2% BSA (Bioshop, ALB003) and 0.5% Triton X-100 (Bioshop, TRX777.100) for 15 minutes at room temperature and incubated in blocking solution (2% BSA (Bioshop, ALB003) in PBS^{++}) for 1 hour at room temperature. Next, primary antibody diluted in PBS^{++} was added as specified and incubated overnight at 4°C. After three washes with PBS^{++} the cells were incubated with fluorescently labeled secondary antibody diluted in PBS^{++} for 1 hour at room temperature in the dark and washed again three times with PBS^{++}. All antibodies are listed in the method section "Antibodies". For visualization of the nucleus, cells were incubated with 0.1µg/ml DAPI (Sigma-Millipore, D9542) in PBS^{++} for 5 minutes at room temperature in the dark. Finally, cells were washed with PBS^{++} and mounted on microscopy slides (Fisherbrand Superfrost®Plus, 12-550-15) with Immunomount (Fisher, 2860042). Images were captured with a Nikon upright fluorescent microscope (Eclipse Ni-E) using NIS Elements.

2.12.4 Antibodies

For IB the membranes were probed with the following primary antibodies at a 1:1000 dilution in 0.3% BSA (Bioshop, ALB003) or 0.3% nonfat milk (Bio-Rad, 170-6404) in TBST: rabbit anti-human BORIS (Rockland, 600-401-907), rabbit anti-human CTCF (Millipore, 07-729), mouse

anti His-tag (Millipore 05-949), mouse anti-human UBF (Novusbio, H00007343-M02, clone 6C6), rabbit anti-human PARP (Cell Signaling, 9542), rabbit anti-human cleaved PARP (Cell signaling, 5625), rabbit anti-human Chk2 (Abcam, ab109413), rabbit anti-human pChk2-Thr68 (Cell signaling, 2197), mouse anti-human pH2AX-Ser139 (Millipore, 05-636I, clone JBW301), mouse anti-human SSRP1 (Biolegend, 609702, clone 10D1), mouse anti-human SPT16 (Biolegend, 607002, clone 8D2). The mouse anti-FLAG® M2 (Sigma, F1804, clone M2) antibody was used at 1:2000 and rat anti-human tubulin (Abcam, ab6160) at 1:7000. Secondary antibodies used for westernblot at a dilution of 1:15000 in TBST were goat anti-mouse (Jackson, 115-035-003), goat anti-rabbit (Jackson, 115-035-144), and anti-rat IgG (Cell Signaling, 7077) coupled to horseradish peroxidase.

The primary antibodies used for IF were mouse anti-human UBF at 1:500 (Novusbio, H00007343-M02, clone 6C6), mouse anti-human pH2AX-Ser139 (Millipore, 05-636I, clone JBW301) at 1:1000 and rabbit anti-human β-actin at 1:000 (Abcam, ab8227). Secondary antibodies used at a 1:2500 dilution were Alexa Fluor® 488 donkey anti mouse (Invitrogen, A21202), Alexa Fluor® 488 donkey anti rabbit (A21206) and Alexa Fluor 594 donkey anti mouse (Invitrogen, A21203).

2.13 CHROMATIN IMMUNOPRECIPITATION (CHIP)

MM057 cells stably infected with hBORIS-3xF were treated with or without 100ng/ml doxycycline for five days. The growth medium was removed and cells were crosslinked with 1% methanol-free formaldehyde (in PBS) for 7 minutes at room temperature with gentle agitation on a rocking platform. The formaldehyde was removed and the reaction was quenched using 0.125M glycine/PBS for 10 min at room temperature with gentle agitation on a rocking platform. Crosslinked cells were subsequently washed twice with cold PBS and harvested by scrapping in ice-cold PBS supplemented with 1x protease inhibitor cocktail (Promega #G6521) followed by centrifugation. The cell pellets were resuspended in cold hypotonic buffer (10mM Tris-HCl pH 7.4, 10mM NaCl, 3mM $MgCl_2$, 0.5% IGEPAL CA-630) and incubated on ice for 15 minutes. Nuclear pellets were obtained by centrifugation at 3500rpm for 5 minutes at 4°C. The nuclei were washed once with cold hypotonic buffer and were stored at -80°C till further use. Next, the nuclear fractions were resuspended in shearing buffer (50mM Tris-HCl pH 8.0, 1mM EDTA, 1%

Triton X-100, 0.3% SDS, 0.3% Na deoxycholate) supplemented with 1x proteases inhibitor cocktail (Promega #G6521). Chromatin was sheared using the Covaris AFA with the following settings; Intensity Peak: 75 watts, Duty cycle: 20%, Cycles per burst: 600 and Processing time: 900. Samples were transferred to a pre-chilled microcentrifuge tube containing dilution buffer (5mM Tris-HCl pH 8.0, 1mM EDTA, 210mM NaCl, 1% Triton X-100) and ChIP buffer (20 mM Tris-HCl pH 8.0, 1mM EDTA, 150mM NaCl, 1% Triton X-100, 0.1% SDS, 0.1% sodium deoxycholate) supplemented with 1x protease inhibitor cocktail (Promega #G6521), and centrifuged at 13000rpm for 5 minutes at 4°C. From each sample, 1% (5µl) was kept as input and 0.5 ml was combined with Dynabeads Protein G conjugated with either mouse anti-FLAG M2 antibody (Sigma, F1804, clone M2) or mouse preimmune IgG. Mixtures were incubated on a rotator overnight at 4°C. The ChIP'ed immune complexes were successively washed four times with Low Salt Buffer (20mM Tris-HCl pH 8.0, 2mM EDTA, 150mM NaCl, 1% Triton X-100, 0.1% SDS), High Salt Buffer (20mM Tris-HCl pH 8.0, 2mM EDTA, 500mM NaCl, 1% Triton X-100, 0.1% SDS) and LiCl Buffer (20mM Tris-HCl pH 8.0, 1mM EDTA, 0.25M LiCl, 0.5% IGEPAL CA-630, 0.5% Na deoxycholate). After the last wash, the final volumes of all samples, including the input controls, were adjusted to 20µl with ChIP-buffer. To all samples, 70µl H_2O, 3µl 1M Tris-HCl pH 8.0 and 2µl proteinase K (20 mg/ml) were added and samples were de-crosslinked overnight at 65°C. The next day, tubes were placed on a magnetic stand for 5 minutes and supernatants were transferred into new tubes. Samples were treated with RNase A (6µl NaCl and 2µl RNase A 10 mg/ml) while shaking at 37°C for 30 minutes and allowed to cool down to room temperature. For double-sided size selection, AMPure XP beads (0.6X; Beckman Coulter, A63881) were added to each sample (including the inputs), mixed by pipetting and vortexing and incubated at room temperature for 15 min. Following a quick centrifugation, tubes were placed on a magnetic stand to separate the beads from the supernatant and left until the solution became clear (≈ 3 minutes). Supernatants were transferred into a new tube and beads containing undesired large DNA fragments were discarded. Subsequently, AMPure XP beads (1.2X; Beckman Coulter, A63881) were added to the samples and the above steps were repeated to remove small DNA fragments. Upon incubation with the beads, this time the supernatants were discarded without disturbing the beads, which contained the desired ChIP'ed DNA. The beads were washed twice with 80% freshly prepared ethanol and allowed to air dry for 5 minutes. To elute the ChIP'ed DNA from the beads, 36µl ddH_2O was added, mixed with the

beads by pipetting and incubated at room temperature for 15 minutes. Tubes were placed on the magnetic stand until the solution became clear and the eluted DNA was transferred to a new tube. The ChIP'ed DNA was subjected to qRT-PCR analysis as described in the section 2.7 "RNA extraction reverse transcription and qRT-PCR" with the following modifications: 2.5µl of DNA was used as template and 200 nM primer. Specific primers were designed for putative BORIS binding sites and 3-4 kbp upstream of the putative BORIS binding sites (negative controls) and purchased from IDT (**Appendix I**). Putative BORIS binding sites were derived from the Encyclopedia of DNA Elements (ENCODE) BORIS/CTCFL ChIP-seq track for the K562 cell line. Amplification was 95°C for 5 minutes, then 40 cycles of 95°C for 15 seconds, 60.5°C for 20 seconds followed by 10 seconds at 95°C and a melt curve of 5 seconds at 65°C and 5 seconds at 95°C on a CFX96 Touch™ (Bio-Rad, 1855196). The delta-delta Ct method [226] was used to determine the initial amount of ChIP'ed DNA template. Then, all ChIP enrichments were calculated as fold enrichment of DNA ChIP'ed at the potential binding sites or control sites relative to the corresponding input samples.

2.14 BISULFITE SEQUENCING

Bisulfite sequencing is based on the conversion of unmethylated cytosine to uracil upon treatment with sodium bisulfite, while leaving 5-methylcytosine unchanged. This approach results in converted DNA in which each unmethylated cytosine (C) will be present as a thymine (T). First, the putative BORIS binding site within the PAX6 promoter region was identified by visualizing the BORIS ChIP-seq track in K562 cells from the ENCODE database in the UCSC table browser. The DNA sequence around the putative BORIS binding site was downloaded from the UCSC table browser and several potential 5-methylcytosines (CpG sites) were identified. PCR primers specific for bisulfite converted DNA were designed to assess 24 CpG sites in this DNA region. gDNA was isolated from the MM102 melanoma cell line with the QIAamp DNA minikit (Qiagen, 51304) according to manufacturer's protocol. Bisulfite conversion was performed on 200ng gDNA using the EZ DNA Methylation Kit (Zymo Research, D5001), according to manufacturer's specifications. Converted and non-converted gDNA was PCR amplified using JumpStart™ Taq DNA Polymerase (Sigma-Aldrich, D9307) in a final volume of 25µl with the primers Pax6-M_F and Pax6-M_R (**Appendix I**). End-point PCR was performed for 40 cycles on a Bio-Rad C1000 Thermal Cycler with the following cycling

conditions: 94ºC for 2 minutes, 94ºC for 30 seconds, 57ºC for 30 seconds, 72ºC for 1 minute, and a final extension at 72ºC for 7 minutes. The PCR products were visualized on an agarose (Bio-Rad, 1613101) gel, extracted with the QIAEX II Gel Extraction Kit (Qiagen, 20021) and ligated at 16ºC overnight into the TA cloning vector pGEM®-T (Promega, A3600). The ligation mixture was transformed into JM109 competent cells, which were spread onto Luria-Bertani (LB) agar plates containing 100µg/ml ampicillin, 50ul IPTG and 50ul Xgal and incubated at 37ºC overnight. Three white colonies were picked, inoculated in 5ml LB medium containing 100µg/ml ampicillin at 37ºC overnight and plasmid DNA was isolated using the QIAprep Spin Miniprep Kit (Qiagen, 27104) according to manufacturer's specifications. A control restriction enzyme digestion was performed using SphI (Fermentas FastDigest) and SalI (Fermentas FastDigest) to verify insertion of the converted gDNA PCR product into the pGEM®-T vector. In addition, the clones containing the pGEM®-T vector with the converted gDNA insert were digested with the restriction enzyme HpaII (Fermentas FastDigest) to confirm that the bisulfite conversion reaction worked. HpaII cuts the DNA at CCGG and can therefore not digest converted gDNA in which C's are converted into T's. The clones were submitted for Sanger sequencing at the Genomics platform of the Institute for Research in Immunology and Cancer (IRIC, Montréal, Canada). All fragments were sequenced from both strands using the M13 sequencing primers and the sequencing results were used to assess the methylation status of the 25 CpG sites in and around the putative BORIS binding site.

2.15 PROTEOMICS

2.15.1 Stable isotope labeling of amino acids in cell culture (SILAC)

HEK293T cells were cultured in Dulbecco Modified Eagle Medium depleted in L-Lysine and L-Argenine (Life Technologies, 89985-500) and supplemented with 10% dialised FBS (Invitrogen, 26400-036), 1% penicillin-streptomycin (Wisent, 450-201-EL) and 1% GlutaPlus (Wisent, 609-066-EL). To label proteins, the medium was supplemented with either a final concentration of 42mg/L L-arginine:HCl (Arg0, Sigma, A8094) and 146mg/L L-lysine:HCl (Lys0, Sigma, L8662) for light labeling or 53.18mg/L L-arginine:HCl 13C6 (Cambridge Isotope, CLM-2265-H-0.1) and 181.49mg/L L-lysine:2HCl 4,4,5,5,D4 (Cambridge Isotope, DLM-2640-0.25) (Arg6, Lys4) for heavy labeled cells. To achieve full incorporation of the labeled amino acids, cells

were cultured for 5 cell doublings using cell dissociation buffer (Life Technologies, 13151-014) for detachment, since trypsin-EDTA may contain non-isotopically labeled amino acids. Next, cells with light labeling and heavy labeling were used to perform the HaloTag® pull-down as specified in the method section "HaloTag® pull-down" for EV and BORIS, respectively and samples were submitted for MS.

2.15.2 Halotag® pull-down (HT-PD)

For HT-PD experiments, we used the HaloTag® Mammalian Pull-Down system (Promega, G6504), as this system allows for efficient capture of protein complexes via a highly specific and covalent bond with HaloLink™ Resin (Promega, PRG1913). The HaloTag® protein and HaloLink™ Resin have little nonspecific protein binding, resulting in a very low background signal. HEK293T cells were transfected with HT-EV, HT-BORIS or HT-CTCF using PEI (Polysciences, 23966, 1:2 ratio of DNA:PEI). Cells were harvested by scraping 48 hours post transfection, washed with 1x PBS and pellets were stored at -80°C for a minimum of 1 hour. Pellets were lysed in mammalian lysis buffer (50mM Tris-HCl pH 7.5, 150mM NaCl, 1% v/v Triton X100 (Bioshop, TRX777.100), 0.2% sodium deoxycholate), sonicated 4 times for 10 seconds at 20% and protein concentration was determined using the Pierce BCA Protein Assay kit according to the manufacturer's protocol (Fisher, 23225). Where indicated, lysate was incubated with 350 units/ml Benzonase® nuclease (Sigma, 71205-3) to exclude nucleic acid dependent interactions. For the pull-down, equal concentrations of protein were incubated with equilibrated HaloLink™ Resin on a tube rotator for 45-60 minutes at room temperature. Resin was washed three times with wash buffer (100mM Tris-HCl pH7.5, 150mM NaCl, 0.05% Igepal CA630 (Bioshop, NON999.500)), one time with final wash buffer (mammalian lysis buffer, 0.05% Igepal CA630 (Bioshop, NON999.500)) and bound proteins were eluted from the beads by incubation at 95°C for 10 minutes in sample buffer (2x SDS (Bio-rad, 1610301), 15% Beta-mercaptoethanol (Sigma, M3148). For MS, protein complexes linked to HaloLink™ Resin were submitted to the University of Victoria-Genome BC Proteomics Centre.

2.15.3 Mass spectrometry (MS)

Manual in-solution trypsin digestion: The protein sample was resolubilized with 250µL PBS prior to reduction with 12.5µl, 200mM dithiothreitol and incubated 30 min at 37°C. Cysteine

sulfhydryl groups were alkylated with 25µl, 200mM 2-iodoacetamide (30 min at 37°C in the dark). Alkylation was quenched with 25µl, 200mM DTT. The samples were diluted with 250µL water and 10µg of trypsin (Promega) was added for digestion at 37°C for 18h. Excessive salt was removed by Oasis HLB solid phase extraction followed by speed vac drying to reduce the volume. Half of the sample volume was desalted by C18 StageTip and speed vac concentrated prior to LC-MS analysis.

LC-MS/MS analysis: The peptide mixture (5µl of 20µl) was separated by on-line reverse phase chromatography using a Thermo Scientific EASY-nLC 1000 system with a reversed-phase pre-column Magic C18-AQ (100µm inner diameter), 2 cm length, 5µm, 100Å, and an in-house prepared reversed-phase nano-analytical column Magic C-18AQ (75µm I.D., 15 cm length, 5µm, 100Å, Michrom BioResources Inc, Auburn, CA), at a flow rate of 300nl/min. The chromatography system was coupled on-line with an Orbitrap Fusion Tribrid mass spectrometer (Thermo Fisher Scientific, San Jose, CA) equipped with a Nanospray Flex NG source (Thermo Fisher Scientific). Solvents were A: 2% Acetonitrile, 0.1% Formic acid; B: 90% Acetonitrile, 0.1% Formic acid. After a 300 bar (~12µl) pre-column equilibration and 300 bar (~3µl) nanocolumn equilibration, samples were separated by a 120-minute gradient (0 minutes: 5% B; 100 minutes: 40% B; 5 minutes: 80% B; 2 minutes: 100% B; hold 13 minutes: 100% B). The Orbitrap Fusion instrument parameters were as follows: Nano-electrospray ion source with spray voltage 2.5kV, capillary temperature 275°C. Survey MS1 scan m/z range 400-1700 profile mode, resolution 240,000 FWHM@200 m/z one microscan with maximum injection time 50 milliseconds. The Siloxane 445.120024 m/z was used for internal lock mass calibration. Data-dependent acquisition Orbitrap survey spectra were scheduled at least every 3 seconds, with the software determining "Top-speed" number of MS/MS acquisitions during this period. The automatic gain control target value for FTMS was 400,000. The most intense ions charge state 2-7 exceeding 20,000 counts was selected for normalized collision energy ion trap MS/MS fragmentation with detection in centroid mode. Dynamic exclusion settings were: repeat count: 1; exclusion duration: 30 seconds with a 10 parts per million mass window. The data dependent (ddMS2 IT normalized collision energy) scan used a quadrupole isolation window of 1.6 Dalton; IonTrap rapid scan rate centroid detection, 1 microscan, 250 milliseconds maximum injection time, automatic gain control 100 target value and stepped collision energy 30% ±5.

2.15.4 MS analysis

Data Analysis Parameters: All data was acquired with Thermo Orbitrap Fusion Tribrid 2.0.1258.14 software with raw files created by XCalibur 3.0.63 (Thermo Scientific) software and analyzed with Proteome Discoverer 1.4.0.228 software suite (Thermo Scientific). Parameters for the Spectrum Selection to generate peak lists of the CID spectra are: activation type: CID; s/n cut-off: 1.5; total intensity threshold: 0; minimum peak count: 1; precursor mass: 350-5000 Da. The peak lists were submitted to an in-house Mascot 2.4.1 server against the UP_human_Canon (69,085 sequences; 23,290,638 residues) database search as follows: precursor tolerance 8 parts per million; MS/MS tolerance 0.8 Da; Trypsin enzyme 2 missed cleavages; FT-ICR instrument type; fixed modification: carbamidomethylation (C); Acetyl (N-term), Oxidation (M), Label:13C(6) (R), and Label:2H(4) (K). Percolator settings: Max delta Cn 0.05; Target False Discovery Rate (FDR) strict 0.01, Target FDR relaxed 0.05 with validation based on *q* value.

Criteria for protein identification: Peptide identifications was performed with the Scaffold software package (version 4.8.3, Proteome Software, Inc, Portland, OR) and were accepted if they could be established at greater than 99.0% probability by the Scaffold Local FDR algorithm. Protein identifications were accepted if they could be established at greater than 99.0% probability and contained at least 2 identified peptides. Protein probabilities were assigned by the Protein Prophet algorithm [227]. Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony. Proteins sharing significant peptide evidence were grouped into clusters. This resulted in the inclusion of 6474 spectra covering 999 proteins for further analysis.

Quantitative data analysis: Statistical analyses of the Proteome Discoverer result files were performed with the Scaffold Q+S software package (version 4.8.3, Proteome Software, Inc, Portland, OR). Normalization was performed iteratively (across samples and spectra) by subtracting the average ratios in log-space (ratio-based analysis). Means were used for averaging. Spectra data were log-transformed, pruned of those matched to multiple proteins, and weighted by an adaptive intensity-weighing algorithm. Of 6474 spectra in the experiment at the given thresholds, 3310 (51%) were included in quantitation. The statistical test utilized in Scaffold Q+ was the permutation test with the following parameters: calculation type: mean; reference type:

individual spectrum reference; normalization between samples: ON; minimum dynamic range: 1%. Proteins were considered significant with $\log_2$ fold-change ($\log_2$FC)≥0.6 and p<0.05 after correction with the Benjamini-Hochberg method [228]. The ranking score was calculated as follows: $\log_2$FC*-$\log_2$(p value). The final list included 83 proteins (excluding BORIS), which were used for downstream analysis.

2.15.5 Molecular interaction search tool (MIST)

To identify protein complexes among the BORIS interacting proteins we used MIST [229]. The selected search type was "Protein List" to find interactors within input. All 83 BORIS interacting proteins were entered and a network search was performed for protein-protein interactions, filtering out low and moderate results.

2.16 NEXT GENERATION SEQUENCING

2.16.1 RNA-sequencing (RNA-seq)

Total RNA was extracted using the Nucleospin RNA kit (Macherey Nagel, 740955.5) according to manufacturer's instructions. DNase treatment excluded gDNA. Total RNA was submitted to the Centre for Applied Genomics in Toronto (Canada) for RNA quantitation and quality assessment, library preparation and sequencing. RNA quality and quantity were assessed on a Bioanalyzer (Agilent Technologies). RNA integrity numbers higher >8.9 were obtained (average for 12 samples is 9.6). RNA-seq libraries were prepared using the NEBNext Poly(A) mRNA Magnetic Isolation Module (New England Biolabs, E7490) to enrich for poly(A) mRNA and the NEBNext Ultra Directional RNA Library Prep Kit for Illumina (New England Biolabs, E7420), according to manufacturers' specifications. Libraries were indexed using the NEBNext Multiplex Oligos 1-8 for Illumina (Index Primers Set 1; New England Biolabs, E7335), according to manufacturers' specifications. The average library size for all 12 samples was 450bp (423-472). Samples were sequenced 8 per 1 lane on a HiSeq2500, using a high throughput V4 flowcell to generate paired-end reads of 126-bases.

2.16.2 RNA-seq data analysis

RNA sequencing data, as raw reads FastQ files were examined with FastQC [230], which indicated high-quality sequencing data. Each FastQ file was aligned to the human reference genome GRCh38.p12 using the STAR RNA-seq aligner [231] with default parameters. The Binary Alignment Map was used to store alignment results and each BAM file was indexed and name sorted with SAMtools [232]. Further downstream analyses, such as RNA transcripts quantification and DEG analysis, was conducted as in [233]. The HTseq framwork was used to count the number of reads that mapped to each gene [234], with the following command; *python -m HTSeq.scripts.count -f bam -r name -s no -i gene_name gencode.v28.annotation.gtf*. A total of 56,832 RNA transcripts were identified by htseq-count.

The mean total read count for all libraries was 27,268,767 (range: 24,488,222-29,136,496). Therefore, 1 Count Per Million (CPM) corresponds to an average of 27 reads (range: 24-29). The CPM cutoff was set at 0.5 CPM (13.5 reads on average, Range: 12-14.5 reads) in all 3 biological replicates of each treatment condition, meaning genes with expression lower than 0.5 CPM were considered as not expressed. The CPM cutoff was established based on the recommended minimum gene count in a library of 5-10 as described in the Biocunductor edgeR user guide. This analysis resulted in 16,308 genes for DEG analysis.

In order to identify DEGs, the Bioconductor edgeR [235] gene-wise exact-test was used. First, we identified DEGs between EV without dox and EV with dox, which indicates a dox-induced effect on gene expression. Every DEG identified was removed from further analyses (with a relaxed statistical cutoff, using a Benjamini and Hochberg [228] FDR cutoff of 0.5). A total of 14 genes were identified with this approach and removed. Next, using the same methodology, we compared BORIS without dox and BORIS with dox. Using a stringent cutoff of FDR<0.05 and a $|\log_2$ fold-change ($\log_2$FC)|>0.5 a total of 2291 genes were identified (801 downregulated and 1490 upregulated; excluding BORIS).

2.16.3 Assay for transposase accessible chromatin followed by sequencing (ATAC-seq)

For ATAC-seq library preparation cells were detached from the culture dish with citric saline (270mM KCl, 30mM sodium citrate), passed through a 40um filter and counted using a

hemocytometer to collect 50.000 cells for all four experimental condition. To isolate nuclei, cells were first lysed for 20 min on ice in lysis buffer #1 (0.1% Sodium citrate tribasic dehydrate (Sigma, S1804) and 0.1% Triton X-100 (Bioshop, TRX777.100)) followed by 20 min lysis on ice in lysis buffer #2 (10mM Tris-HCl pH7.4, 10mM NaCl, 3mM MgCl, 0.1% Igepal CA630 (Bioshop, NON999.500) and 0.1% Tween-20 (Bioshop, TWN510.500). The nuclei were pelleted by centrifugation for 10min at 600g (4°C) and supernatant was carefully removed to reduce mitochondrial DNA (mtDNA) content. For the transposase reaction, the nuclei pellet was re-suspended in 25µl transposase reaction mix containing of 4µl Tn5 transposase in 1x transposase reaction buffer (Nextera DNA library kit, Illumina, FC-121–1030) and 0.1% Tween-20 (Bioshop, TWN510.500), and incubated for 30min at 37°C with gentle shaking (400rpm). Immediately following transposition, DNA was purified using the MinElute PCR purification Kit (Qiagen, 28004), according to manufacturer's instructions. Library fragments were amplified in two steps. First, 5 cycles of PCR were performed using NEBNext® High-Fidelity PCR master mix (New England BioLabs, M0541) with 1.25µM barcoded Nextera PCR primers (**Appendix I**) in a final volume of 50µl on a BioRad C1000 Thermal Cycler. The cycling program used was as follows: 5 min at 72°C, 30 sec at 98°C, 5 times cycling for 10sec at 98°C, 30sec at 63°C and 1 min at 72°C. To reduce GC and size bias during library amplification, 5µl from the PCR mix was used for a 15µl qPCR side-reaction containing the same reagents with the addition of SYBRgreen (9X, Life Technologies, S7563) to determine the number of additional PCR cycles and stop amplification prior to saturation. The CFX96 Touch™ (Bio-Rad, 1855196) qPCR machine was used with the following cycling conditions: 30 sec at 98°C, 20 times cycling for 10sec at 98°C, 30sec at 63°C and 1 min at 72°C followed by melt curve analysis. Using the CFX Manager™ Software (Bio-Rad), the fluorescent intensity was plotted versus cycle number and the additional number of cycles was established as the cycle number that corresponds to ¼th of the maximum fluorescent intensity. Next, the initial PCR reaction mix was amplified for 3 additional PCR cycles. Library preparation for each experimental condition was performed in triplicate and samples were combined post PCR amplification for quality controls and sequencing. One third of the library was cleaned up using KAPA Pure Beads (1.8X; Roche, 7983298001) and analyzed on a 2.5% agarose (Bio-Rad, 1613101) gel to verify nucleosomal patterning. The remainder of the library was size selected using KAPA Pure Beads (0.55X-1.8X; Roche, 7983298001) to enrich for 150-750 bp fragments. Part of the DNA from the size selected

library was diluted 1 in 5 and used for qPCR to confirm enrichment for open chromatin and estimate mtDNA content. The qPCR reaction contained 2.5µl DNA template, 5ul PerfeCTa SYBR Green FastMix (Quanta, 95072-012) and 300nM forward and reverse primer. The thermal cycling program used was as follows: 3 minutes at 95°C, 40 times cycling for 15 seconds at 95°C and 30 seconds at 60°C followed by melt curve analysis. To assess enrichment of open chromatin primers were designed in genomic regions that are often marked by open chromatin (GAPDH and KAT6B) or closed chromatin (SLC22A3 and RHO). A minimum difference in quantitation cycle (Cq) value of 3, representing an 8-fold enrichment, was considered as enrichment of open chromatin in the libraries. The sequences of these primers as well as the primers directed at mtDNA are listed in **appendix I**. The remaining DNA of the libraries was submitted to the McGill University and Genome Québec Innovation Centre and fragment distribution was assessed by TapeStation (Agilent). Due to primer contamination, libraries were re-purified using AMPure XP beads (1.8X; Beckman Coulter, A63881) and re-submitted to the McGill University and Genome Québec Innovation Centre. Again, fragment distribution was assessed by TapeStation (Agilent), libraries were quantified using Quant-iT™ PicoGreen® dsDNA Reagent® (ThermoFisher Scientific), pooled and sequenced paired-end (100bp read length) on one lane of an Illumina HiSeq4000.

2.16.4 ATAC-seq data analysis

ATAC-Seq data, as raw reads FastQ files, were first examined with FastQC [230]. Warnings by FastQC in some fields were resolved as follows; tiles with poor quality reads were removed with the *filterbytile.sh* script from the BBMap tool [236], using default parameters. Adapter contamination of the Nextera Transposase Sequence was removed, using the *trim_adapters* script by the atactk [237] toolkit with default parameters. In addition, the first 17 base-pairs (bp) of all reads were trimmed with cutadapt [238] due to unequal per base sequence content distribution. The final trimmed reads had a sequence length of 83 bp and an average of 164,561,210 (Range; 164,424,288-187,302,698) paired-end total sequences, for all four experimental conditions. Next, trimmed reads were aligned the to the human reference genome GRCh38.p12 with BWA-MEM [239], using the following command; *bwa mem -M -a -t 30 hg38.fa. read1.fastq read2.fastq*, where *read1.fastq* and *read2.fastq* were modified to indicate the forward and reverse reads of each library. The resulting alignment files in the Binary Alignment

Map were further indexed and name-sorted with SAMtools [232]. Duplicate alignments were marked and removed with the MarkDuplicates tool from Picard Tools [240], using the default parameters with the following changes: *REMOVE_DUPLICATES* was set to *true* and *VALIDATION_STRINGENCY* was set to *LENIENT*. The average read pair duplication rate (including PCR and optical duplicates), as determined by MarkDuplicates was 31.88% (Range; 30.2%-35.1%). After duplication removal, an additional filtering step was implemented with SAMtools, as recommended by the Parker laboratory (see below). First, we defined autosomal references with the command*; export CHROMOSOMES=$(samtools view -H file_name.bam | grep '^@SQ' | cut -f 2 | grep -v -e _ -e chrM -e chrY -e 'VN:' | sed 's/SN://' | xargs echo)*; once the variable CHROMOSOMES was defined the following command was ran; *samtools view -b -h -f 3 -F 4 -F 8 -F 256 -F 1024 -F 2048 -q 20 file_name.bam $CHROMOSOMES.* The result of this sifting step was the exclusion of reads mapped to Mitochondrial DNA and Chromosome Y, in addition to reads that were not mapped/paired properly or mapped to multiple places in the genome. Lastly, reads with Q-score<20 were removed. The average final number of properly aligned PE reads saved in Binary Alignment Map format for downstream analyses were 85 million (Range: 77,670,407-95,952,118). The filtered Binary Alignment Maps served as input for MACS2 [241, 242]. The MACS2 call-peak function was used for peak calling with the following command; *macs2 callpeak -f BAMPE -t filel.bam -g hs -q 0.01 --nomodel --llocal 20000 --shift -37 --extsize 73 --call-summit –B*. The results of MACS2 were saved as BedGraph files, in addition to excel files that contain peak location, length, enrichment, and statistic-index values. From the BedGraph files we generated UCSC Genome Browser compatible BigWig files with the following commands; *(1) env LC_COLLATE=C sort -k1,1 -k2,2n file_name.bdg > file_name.sorted.bdg (2) bedGraphToBigWig file_name.sorted.bdg hg38.chrom.sizes.txt file_name.bw*. The pipeline described here is based on recommendations from the Parker laboratory: https://github.com/ParkerLab/bioinf525.

2.17 OVERREPRESENTATION ANALYSIS

2.17.1 Protein analysis through evolutionary relationship (PANTHER)

The list of BORIS-interacting proteins as identified through SILAC-based MS was used as input for overrepresentation analysis in PANTHER (version 13.1) [243] for GO-Slim Biological

Processes. The test type was Fisher's Exact with FDR multiple test correction. REduce and Visualize Gene Ontology (REViGO) [244] was used to remove highly redundant processes based on their GO terms. The *p* values from PANTHER were provided and the allowed similarity stringency was set to medium (0.7). In REViGO the semantic similarity measure was performed based on Jiang and Conrath and GO terms were set to only include Homo Sapien. The interactive graph function was used for visualization and the output was manually annotated to obtain a comprehensive, clustered overview of the main overrepresented biological processes. Bubble color indicates the user-provided *p* value; bubble size indicates the frequency of the GO term in the underlying GOA database (bubbles of more general terms are larger). Highly similar GO terms are linked by edges in the graph, where the line width indicates the degree of similarity.

2.17.2 Consensuspath database (CPDB)

Reactome pathway analysis in CPDB (release 33) [245, 246] was used for both BORIS-interacting proteins and differentially expressed genes. The list of BORIS-interacting proteins as identified through SILAC-based mass spectrometry, BORIS downregulated DEGs (FDR<0.05) and BORIS upregulated DEGs (FDR<0.05) were provided separately and the overrepresentation test was performed for Reactome pathways with a minimum overlap of 4 proteins with the input list and a $p<0.01$. Identified pathways with a $q<0.01$ (BORIS-interacting proteins) or <0.05 (BORIS DEGs) were visualized and clustered using the concept overlap graph function within CPDB allowing a relative overlap of 0.5.

2.18 GENE SET ENRICHMENT ANALYSIS (GSEA)

GSEA (version 3.0 build 0160) [247] was used on the RNA-seq data for enrichment analysis. The RNA-seq recommended GSEA-PreRanked algorithm was used with 10^4 permutations, ranking genes by $\log_2$FC values (as obtained from the edgeR analysis of BORpos vs BORneg samples). The entire ranked gene list was used in accordance with GSEA requirements. The enrichment gene sets in GSEA that reached a nominal $p<0.01$ and FDR<0.01 were considered statistically significant. The GO gene set for biological processes and hallmark EMT gene set from the Molecular Signatures Database [248] were used. In addition, published melanoma gene sets were used, two proliferation and two invasiveness gene sets; Verfaillie [59] ('proliferative':

770 genes, 'invasive': 642 genes) and Hoek [51] ('proliferative': 54 genes, 'invasive': 51 genes). Furthermore, the published set of upregulated genes upon NFATC2 knockdown (1758 genes) was utilized [61].

For visualization of enriched biological processes the Cytoscape (version 3.6.1. Java version: 1.8.0_171) [249] plugin EnrichmentMap (version 3.1.0) [250] was used with the q value cutoff set at 0.05 and edge cutoff at 0.6. The plugin AutoAnnotate (version 1.2) [251] and WordCloud (version 3.1.2) [252] were used for initial annotation with the following setting: cluster algorithm: MCL Cluster, edge weight column: similarity coefficient, label algorithm: WordCloud adjacent words. The layout network was chosen to prevent cluster overlap and the minimum number of GO's per cluster was 4.

2.19 PREDICTION OF TRANSCRIPTION REGULATORS

Transcriptional regulator prediction was performed using the iRegulon plugin (version 1.3) [253] in Cytoscape (version 3.6.1. Java version: 1.8.0_171) [249]. The list of genes, either BORIS upregulated DEGs or BORIS upregulated invasive genes (FDR<0.05), was submitted to iRegulon and analyzed using the following options: 1) motif collection on 20-kb region centered around the TSS for 7 species, which included 9,713 position weight matrices, and 2) track collection on 20-kb region centered around the TSS for 1120 ChIP-seq tracks of ENCODE raw signals. The top 5 motifs and tracks are displayed.

CHAPTER 3

–

RESULTS

CHAPTER 3.1

–

BORIS Expression Alters X-linked Gene Expression in a XaXi Melanoma Cell Line

Sanne M. Janssen, Mounib Elchebly, Anke F. van Rijk, Fanny Beltzung, Mohammad Amraei, Sophia Borchers, Léon C.L. van Kempen and Alan Spatz

Background and objectives

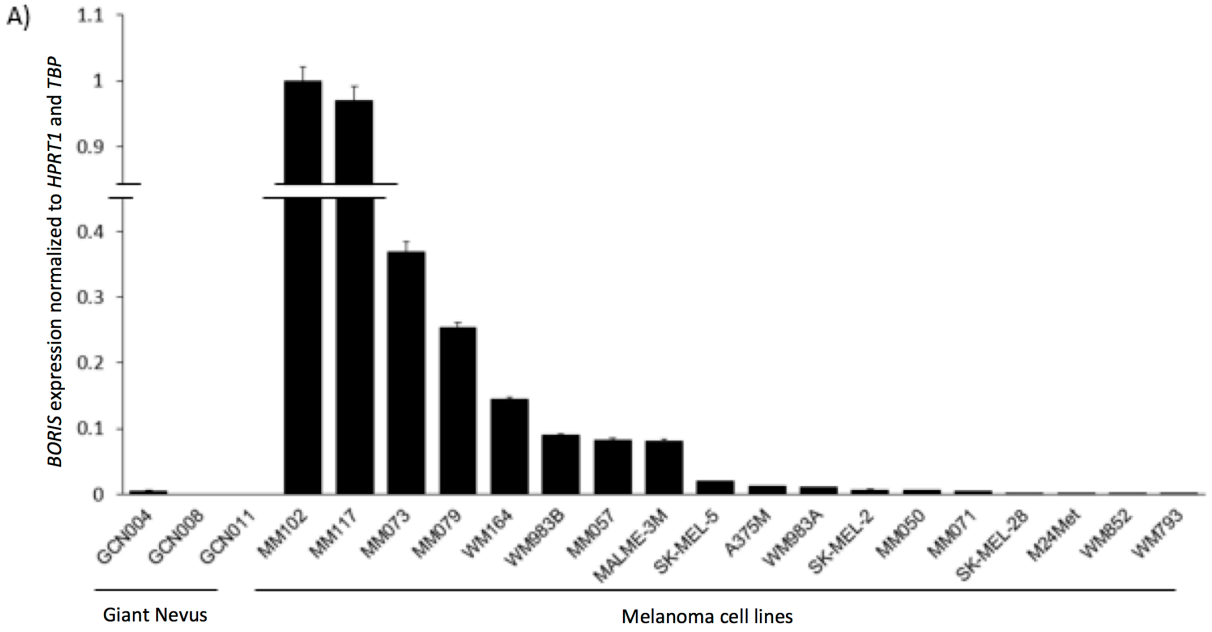
BORIS and CTCF have a highly similar DNA binding domain and share a subset of DNA binding sites across the genome [144, 188]. At shared binding sites, BORIS and CTCF were shown to compete for DNA binding, which resulted in a different outcome on gene expression [144]. One important function of CTCF is as an insulator protein. Interestingly, while CTCF binds both alleles of the X-chromosome, allele-specific CTCF binding profiles revealed distinct binding on the Xi at escape genes, indicating an important role for CTCF as insulator on the Xi in females [211]. The X-chromosome is of particular interest in melanoma, since females have a better outcome compared to males [63, 65]. Our laboratory previously demonstrated loss of the Xi in ~23% of melanoma metastasis from female patients and showed a correlation with poor distant metastasis-free survival [222], stressing the importance of X-linked gene expression regulation in melanoma. Furthermore, NanoString gene expression profiling of melanoma cell lines with genetically altered BORIS expression revealed large-scale changes in expression levels of X-linked genes [220]. Currently, no studies have investigated if BORIS expression can interfere with CTCF's function as insulator, specifically on the X-chromosome. Here, we set out to address if BORIS expression leads to mis-regulation of X-linked gene expression and if this is due to altered expression from the Xa or Xi.

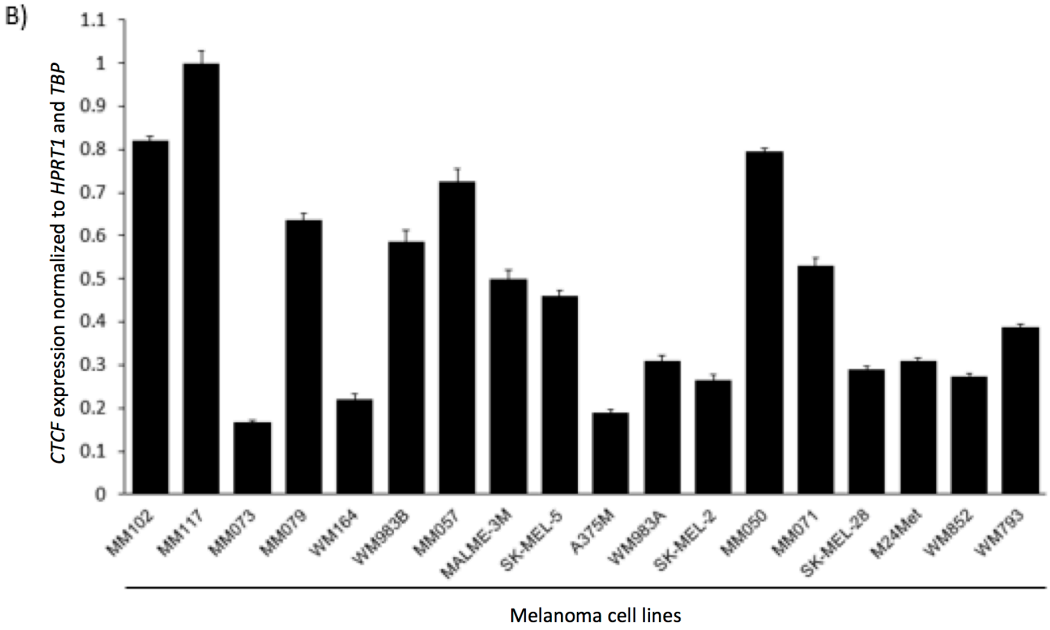
Results

3.1.1 BORIS and CTCF expression in melanoma

Before addressing the effect of BORIS expression on X-linked gene expression in melanoma, we aimed to confirm that BORIS is expressed in melanoma cell lines as reported previously [138]. In addition, we wanted to determine if a correlation exists between BORIS and CTCF expression level in melanoma. Expression of *BORIS* and *CTCF* mRNA was measured using qPCR in a panel of melanoma cell lines (**Fig. 8A and B**). *BORIS* mRNA was detected in ~60% (11 out of 18) of the melanoma cell lines. We did not detect *BORIS* expression in non-malignant giant nevi cells. Furthermore, no correlation was observed between *BORIS* and *CTCF* expression. Next, BORIS and CTCF protein expression were determined by IB. The chronic myelogenous leukemia cell line K562 was used as a positive control for BORIS protein expression [144, 195]. In accordance with previous findings in melanoma tissue [138], BORIS expression was detected in ~33% (4 out of 12) of the metastatic melanoma cell lines. CTCF protein was observed in all the cell lines **(Fig. 8C**). Again, no correlation was observed between BORIS and CTCF protein expression. Furthermore, *BORIS* mRNA did not correspond to BORIS protein level, as we observed very low to no detectable *BORIS* mRNA in 3 out of 4 melanoma cell lines with BORIS protein expression. It is important to note that it has proven difficult to reproduce BORIS IB data due to the lack of a clean and sensitive commercially available BORIS antibody following the inability to reproduce the commonly used Abcam antibody (ab18337). Together, these data indicate that BORIS is expressed in melanoma cells at the mRNA and protein level, and BORIS expression is not correlated to CTCF mRNA or protein level.

Figure 8


A)
BORIS expression normalized to HPRT1 and TBP
1.1
1
0.9
0.4
0.3
0.2
0.1
0
GCN004
GCN008
GCN011
MM102
MM117
MM073
MM079
WM164
WM983B
MM057
MALME-3M
SK-MEL-5
A375M
WM983A
SK-MEL-2
MM050
MM071
SK-MEL-28
M24Met
WM852
WM793
Giant Nevus
Melanoma cell lines


B)
CTCF expression normalized to HPRT1 and TBP
1.1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
MM102
MM117
MM073
MM079
WM164
WM983B
MM057
MALME-3M
SK-MEL-5
A375M
WM983A
SK-MEL-2
MM050
MM071
SK-MEL-28
M24Met
WM852
WM793
Melanoma cell lines

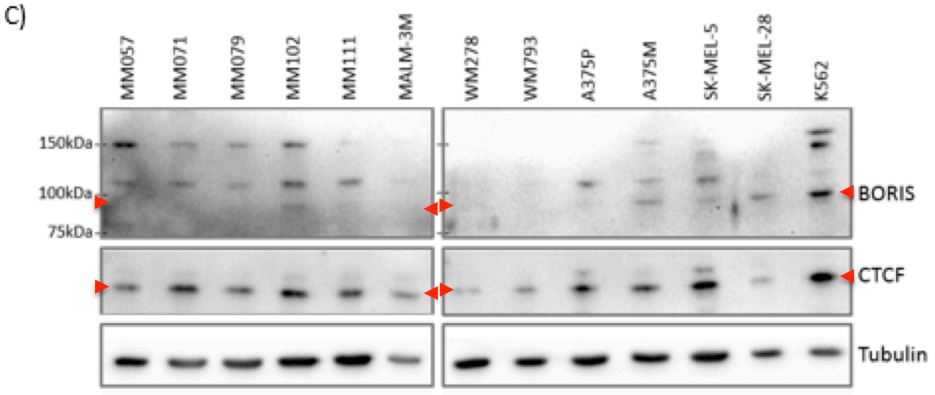

C)
MM057
MM071
MM079
MM102
MM111
MALM-3M
WM278
WM793
A375P
A375M
SK-MEL-5
SK-MEL-28
K562
150kDa
100kDa
75kDa
BORIS
CTCF
Tubulin

Figure 8. BORIS and CTCF expression in a panel of melanoma and non-malignant giant nevi cell lines.

A) *BORIS* and B) *CTCF* mRNA expression was measured by qPCR and normalized to *HPRT1* and *TBP*. Data are represented as mean normalized expression scaled to the highest expressing cell line (designated as 1.0). For each sample, measurements were performed in triplicate. Data was analyzed with the Bio-Rad CFX Manager™ Software. Error bars represent the standard error of mean. C) BORIS and CTCF protein expression were measured by IB using anti-BORIS and anti-CTCF antibody. Red arrows indicate the anticipated molecular weight of BORIS and CTCF protein. Anti-Tubulin was used as a loading control.

3.1.2 Modulation of BORIS expression in a XaXi melanoma cell line

Next, we wanted to determine if expression of BORIS has an effect on X-linked gene expression in melanoma cells. As we wanted to address expression from the Xa and Xi, we first tested ten melanoma cell lines for the number of X-chromosomes using DNA- FISH. Out of the tested cell lines most displayed >2 X-chromosomes, except for the three cell lines, MM079, MM073 and MM057, for which we counted 2 X-chromosomes in 98%, 91% and 95% of the cells, respectively (**Fig. 9**). Among these cell lines we found that MM057 expressed *Xist* from one of the two X-chromosomes as detected by RNA-FISH (**Fig. 10A**), indicating that this cell line contains one Xa and one Xi. Based on the characterization of BORIS expression in our melanoma cell lines (section 3.0.2), we knew that the MM057 cell line has low *BORIS* mRNA expression and no detectable BORIS protein (**Fig 8A and C**). Lack of detectable endogenous BORIS protein ruled out the use of a knockdown system in this cell line. Therefore, BORIS level was modulated by infection with an expression vector encoding a His-tag fused to the fluorescent protein Katushka (TurboFP635) fused to either full length hBORIS (HK-BORIS) or empty vector (HK-EV) control. This allowed us to visualize ectopic BORIS expression and assess infection efficiency with a fluorescent microscope. To validate HK-BORIS expression, cells were sorted by FACS 72h post infection and HK-BORIS protein and mRNA expression were validated by IB (**Fig. 10B**) and qPCR (**Fig. 10C**). In sum, we identified a XaXi melanoma cell line in which we can modulate BORIS expression.

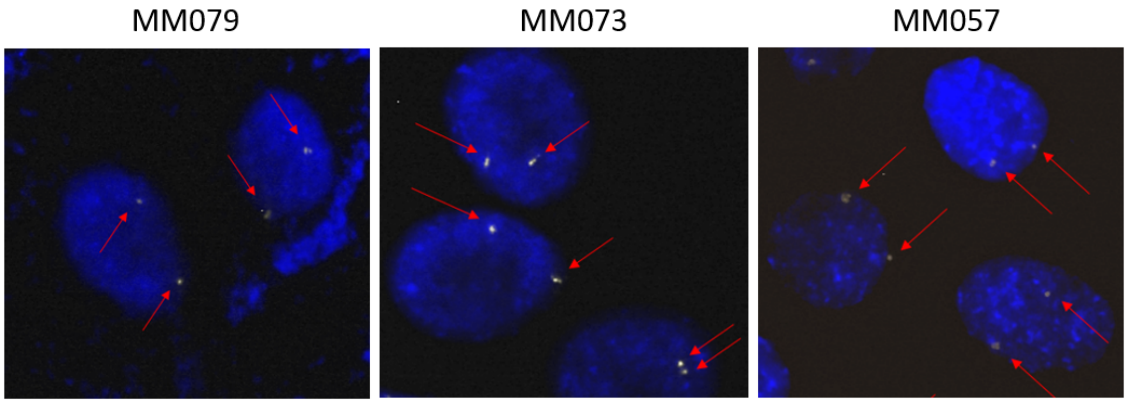


Figure 9. Three melanoma cell lines with 2 X-chromosomes.

Visualization of the X-chromosomes using DNA-FISH towards the X-chromosome centromere (CEP-X; white). Nuclei were visualized with DAPI (blue) and images were captured with a Zeiss AX10 microscope. Red arrows point to the X-chromosomes.

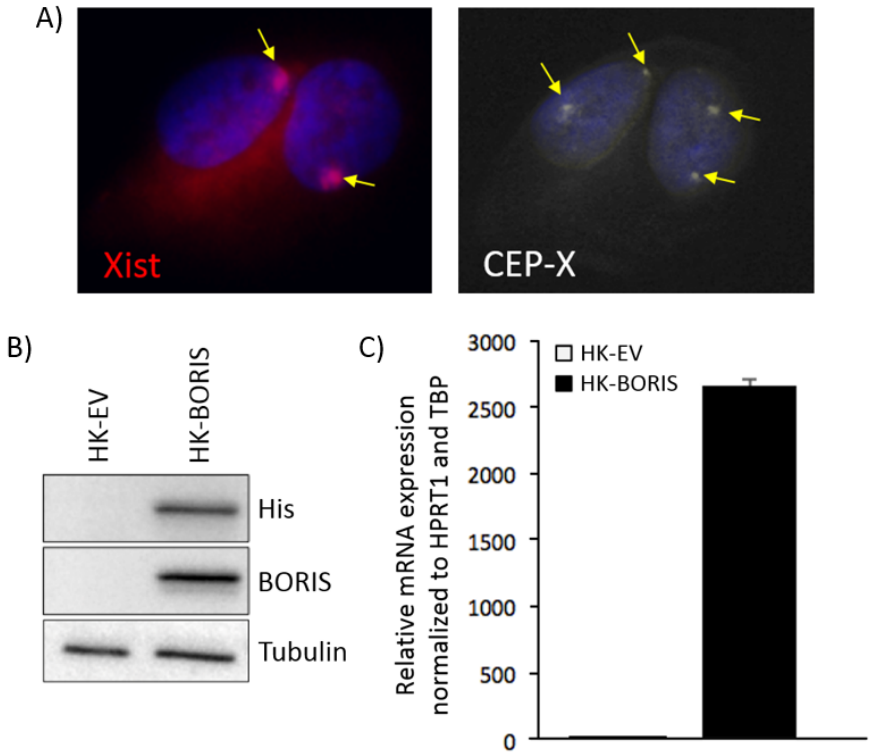


Figure 10. Ectopic BORIS expression in a female XaXi melanoma cell line.

A) Visualization of the inactive X-chromosome (left) followed by the X-chromosomes (right) using RNA-FISH targeting *XIST* (red) and DNA-FISH towards the X-chromosome centromere (CEP-X; white), respectively. Nuclei were visualized with DAPI (blue) and images were captured with a Zeiss AX10 microscope. B and C) MM057 cells were genetically engineered to stably express HK-BORIS or HK-EV control. Katushka positive cells were sorted using FACS. B) Expression of His-tag and BORIS protein was validated by IB using anti-His and anti-BORIS antibodies. Anti-Tubulin was used as a loading control. C) *BORIS* mRNA expression was assessed by qPCR and normalized to *HPRT1* and *TBP*. For each sample, measurements were performed in triplicate. Data was analyzed with the Bio-Rad CFX Manager™ Software. Error bars represent the standard error of the mean.

3.1.3 BORIS expression results in differential expression of X-linked genes

To address if BORIS expression can alter X-linked gene expression, we carefully selected a subset of 12 X-linked genes (**Table II**) based on the following criteria: 1) the X-inactivation status to include both inactive and escaping genes [254], 2) involvement in cancer according to a literature search, and 3) presence of a putative BORIS/CTCF binding site within the promoter region to include genes expected to be regulated by BORIS. ENCODE ChIP-seq tracks for the K562 cell line were used to identify putative BORIS (ENCSR000BNK) and CTCF (ENCSR000DMA) binding sites. HK-BORIS positive cells were collected by FACS and BORIS expression was confirmed by qPCR (**Fig 11A**). Next, the effect of HK-BORIS on the expression of the X-linked genes was assessed by qPCR. For the inactive genes, we observed significant downregulation of *BEX2* and upregulation of *RAP2C*, *ZBTB33 (KAISO)*, *MST4* (STK26) and *PIM2* compared to the EV control (**Fig 11B**). *FGF13* mRNA expression was not detectable. *BEX2* and *PIM2* do not have a BORIS binding site in their promoter region. It is possible that these genes are regulated via higher order chromatin conformation changes, such as chromatin looping. Indeed, visualization of Chromatin Interaction Analysis by Paired-End Tags (ChIA-PET) data track for CTCF in the UCSC genome browser (ENCODE project, GSM970216, K562 cell line) revealed that both *BEX2* and *PIM2* are located within BORIS/CTCF-mediated chromatin loops, which could be altered by BORIS-CTCF competition for the DNA binding sites. The two genes without a known inactivation status, *BMX* and *SAGE1*, were both significantly upregulated upon BORIS expression (**Fig. 11C**). Similarly to the inactivated genes, BORIS mediated regulation of these genes is likely via direct binding in the promoter region and/or altered chromatin looping events. In contrast to the genes that are inactivated or have an unknown status, none of the escaping genes were upregulated in the presence of HK-BORIS. We did observe a significant downregulation of the escaping genes *RBBP7* and *PPP2R3B* compared to the EV control (**Fig. 11D**). This was surprising as both *RBBP7* and *PPP2R3B* do not contain a BORIS binding site in their promoter region, while *ZFX* does. Looking into CTCF ChIA-PET data revealed that *RBBP7* is present within a BORIS/CTCF-mediated chromatin loop, though downregulation of *PPP2R3B* cannot be attributed to chromatin looping events. Regarding *ZFX*, a close look at the BORIS biding site demonstrates that this binding site is within the promoter region of both *ZFX* as well as its antisense non-coding RNA, which could inhibit *ZFX* expression. It is important to note that while we use publicly available data to gain inside into

possible involvement of BORIS in the regulation of these X-linked genes, validation of the regulatory mechanisms of these genes is outside the scope of this work. Together, these results indicate that modulating BORIS levels can alter the expression of X-linked genes.

Table II: List of X-linked genes for differential gene expression

Gene ID	Inactivation status	Involvement in cancer	BORIS binding site
BEX2	Inactive	[255-257]	No
RAP2C	Inactive	[258]	Yes
ZBTB33	Inactive	[259-262]	Yes
MST4 (STK26)	Inactive	[221, 263, 264]	Yes
FGF13	Inactive	[221, 265, 266]	No
PIM2	Inactive	[221, 267, 268]	No
BMX	Unknown	[221, 269, 270]	No
SAGE1	Unknown	[271, 272]	Yes
ZFX	Escaping	[273-275]	Yes
KDM5C	Escaping	[276, 277]	No
RBBP7	Escaping	[221, 278, 279]	No
PPP2R3B	Escaping	[222]	No

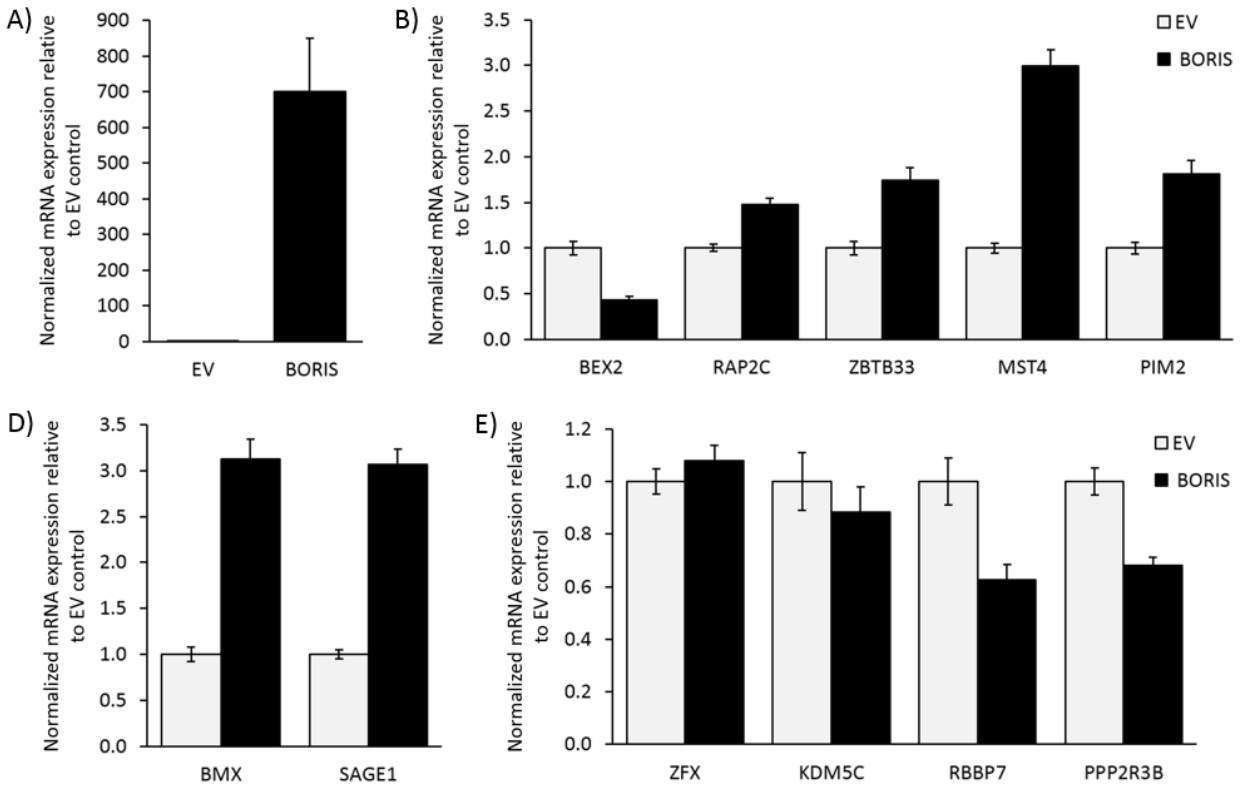


Figure 11. Ectopic BORIS expression alters the expression of X-linked genes.

qPCR analysis for A) *BORIS*, B) inactivated X-linked genes, C) X-linked genes with an unknown inactivation status, and D) escaping X-linked genes in the MM057 cell line with ectopic HK-BORIS (BORIS) or HK-EV expression (EV). Expression was normalized to *HPRT1* and *TBP*. For each sample, a technical triplicate was performed for a minimum of two biological replicates. Data was analyzed with the Bio-Rad CFX Manager™ Software. Error bars represent the standard error of the mean.

To determine if the upregulation of inactivated X-linked genes is a result from increased expression from the Xa or Xi, it is necessary to discriminate between these two X-chromosome alleles. This can be done with the use of SNPs within the 5' UTR, 3' UTR or coding region of a gene. SNPs are changes in the DNA sequence with a frequency >1% that occur on one of the two chromosomal alleles. Putative SNPs were identified for differentially expressed genes using the SNP database within the UCSC genome browser (dsSNP build 141 Found in >= 1% of Samples). Only SNPs with a minor allele frequency, the frequency at which the second most common allele occurs, of >15% were included (**Fig. 12A**). No SNPs were found for *RAP2C*, *PIM2* and *BMX*. To validate the presence of the identified SNPs in EV and BORIS positive MM057 cells, the gDNA regions containing the SNPs were PCR amplified and subjected to Sanger sequencing. We were able to validate one SNP for the inactivated gene *ZBTB33* (rs201958171), which results in a TGA insertion (**Fig. 12B**). In addition, several SNPs for the escaping gene *PPP2R3B* (rs1133530, rs14127, rs1133535 and rs17855192) were confirmed (**Fig. 12C**). Unfortunately, none of the other SNPs could be validated in our model system. Next, mRNA of the HK-EV and HK-BORIS positive cells was reverse transcribed into cDNA and the regions containing SNPs for *ZBTB33* and *PPP2R3B* were PCR amplified followed by Sanger sequencing. This allowed us to determine whether RNA expression occurs from one or both alleles. In the MM057 cells with HK-EV expression, we observed *ZBTB33* expression mainly from the allele containing the insertion, indicating that this is the active allele. On the contrary, we observed expression from both alleles with HK-BORIS expression (**Fig. 12B**), suggesting that expression from the inactive allele contributes to *ZBTB33* upregulation. *PPP2R3B* demonstrated RNA expression from both alleles with HK-EV expression, which was expected since *PPP2R3B* is a known XIC-escaping gene. Though *PPP2R3B* is downregulated with HK-BORIS expression, we did not observe reduced expression from a specific allele (**Fig. 12C**). These data suggest that BORIS expression can result in upregulation of inactivated genes via increased expression from the inactivated allele, while this effect was not observed for escaping genes. Unfortunately, due to the lack of SNPs in the MM057 cell line for the upregulated genes *RAP2C, MST4, PIM2, BMX* and *SAGE1*, we were only able to demonstrate increased expression from the inactive allele for *ZBTB33*.

Figure 12

A)

Gene Name	SNP #	AF 1	AF 2	AF 3
BEX2	rs7557	73.5% C	26.5% T	
RAP2C	no SNP			
ZBTB33	rs201958171	43.7% -	56.3% TGA	
ZBTB33	rs7882480	3.1% -	34.4% C	62.5% T
ZBTB33	rs11537755	46.6% A	53.4% G	
ZBTB33	rs1044413	0.6% -	51.1% A	48.3% G
MST4*	rs5933055	76.5% C	20.5% T	
MST4	rs995246	76.6% A	23.4% T	
MST4	rs201033420	21.9% -	78.1% A	
PIM2	no SNP			
BMX	no SNP			
SAGE1	rs5930805	20.1% A	79.2% G	
SAGE1**	rs12014884	19.4% C	80.6% T	
SAGE1	rs4829799	1.5% -	18.9% C	79.6% T
SAGE1	rs4829584	0.4% -	13% T	86.6% G
RBBP7***	rs67984110	76.9% C	23.1% T	
PPP2R3B	rs1133530	48.5% A	51.5% G	
PPP2R3B	rs14127	76.7% C	23.3% T	
PPP2R3B	rs1133535	49.2% C	50.8% T	
PPP2R3B	rs1133523	55% C	45% T	
PPP2R3B	rs1133520	75% C	25% T	
PPP2R3B	rs2738376	87% C	13% T	
PPP2R3B	rs3813594	60% C	40% T	
PPP2R3B	rs3813593	10% A	90% G	
PPP2R3B	rs17855192	30% A	70% G	

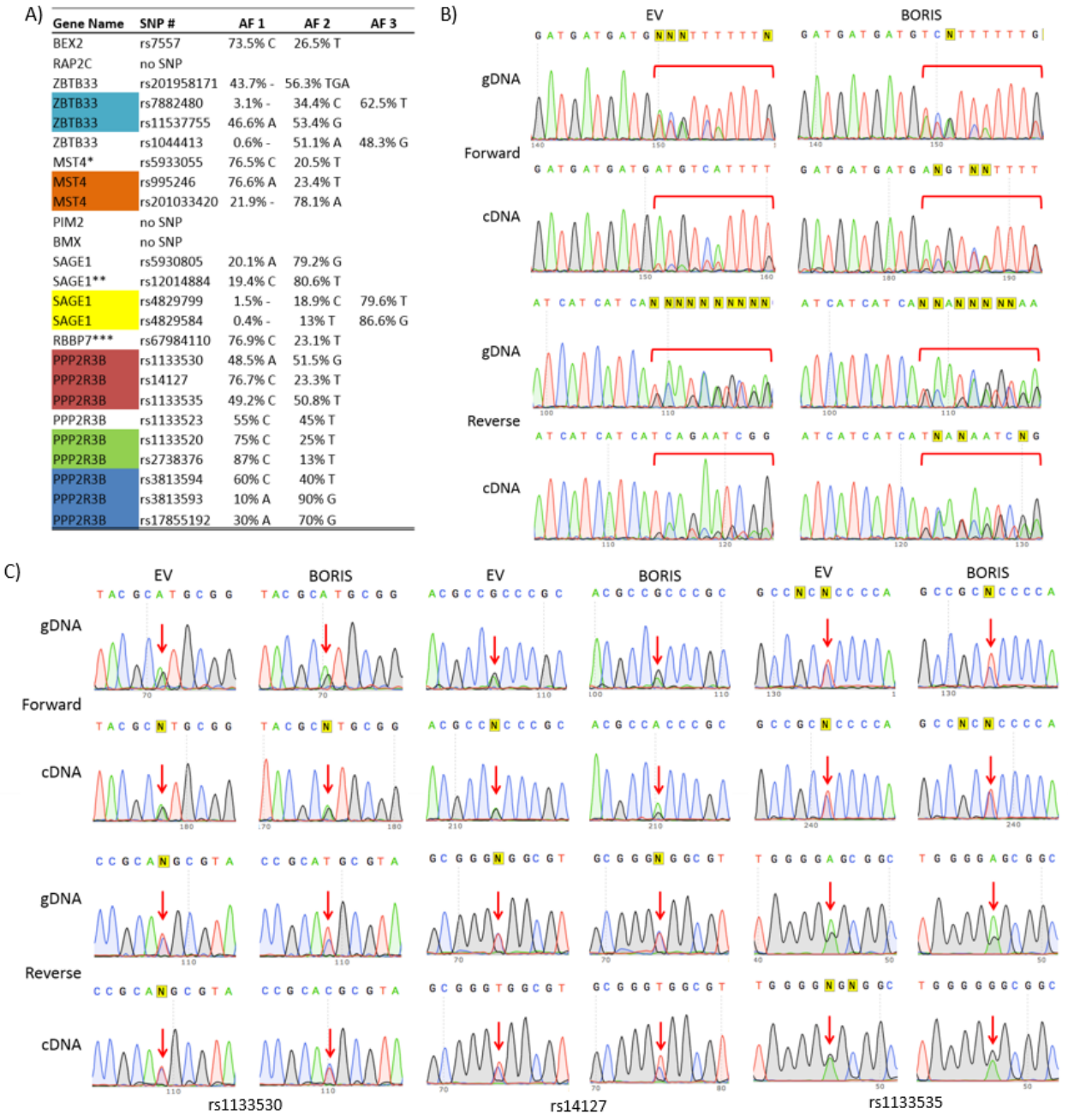


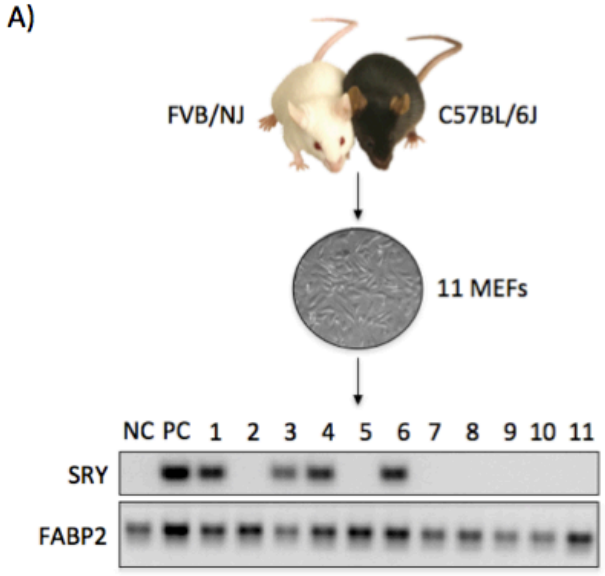
Figure 12. Ectopic BORIS expression results in mRNA expression from the Xi for the inactivated X-linked gene *ZBTB33*.

A) Overview of putative SNPs within differentially expressed X-linked genes as identified from the SNP database within the UCSC genome browser. Displayed are the gene name, the reference SNP (rs) cluster identification number and the allele frequency (AF) at which the allele occurs within the population. Genes for which multiple SNPs could be detected with one primer pair are color highlighted. B and C) Sanger sequencing results of the forward and reverse strand for B) SNP rs201958171 (TGA insertion) in ZBTB33 and C) SNPs rs1133530, rs14127 and rs1133535 in PPP2R3B at the gDNA and mRNA (cDNA) level of MM057 cells with ectopic HK-BORIS (BORIS) or HK-EV expression (EV). The region marked by a red bar/red arrow indicates the location of the SNP/allelic expression difference.

3.1.4 Establishing a XaXi cell model to study the effect of BORIS expression on X-linked allele-specific gene expression

With the absence of SNPs in the MM057 cell line, a better model was needed to determine the effect of BORIS expression on allele-specific gene expression. Human cell lines seem to have many limitations, including a wide variety in X-chromosome number, lack of SNPs and unknown parental origin of the Xi and Xa. Therefore, we decided to make use of the benefits of well-characterized mice genomes to further investigate the role of BORIS in allele-specific expression. Genomes of various mice strains have been well studied and information on SNPs between different backgrounds of mice strains is available in the Mouse Genome Informatics database from the Jackson Laboratory [225]. A male C57BL/6J and female FVB/NJ mice were crossed to generate MEF cell lines with known SNPs. According to the MGI database (dbSNP Build 137) offspring from this cross will have SNPs in 204 X-linked genes (**Appendix II**). In total, 11 MEF cell lines were established. Genotyping by PCR on the gDNA of the MEFs revealed that 4 out of 11 cell lines expressed the male SRY gene, indicating that the remaining 7 cell lines are of female origin (**Fig. 13A**). Interestingly, the non-coding RNA *Xist* was among the list of X-linked genes with four putative SNPs (**Fig. 13B**). Importantly, a SNP in *Xist* would allow us to determine wether the maternal or paternal X-chromosome is inactivated. This means that, in addition to identifying if expression changes originate from the Xi or Xa, we can determine if the expression changes originate from the paternal or maternal X-chromosome. Next, single cell clones from two female MEF cell lines (#2 and #9) were expanded and gDNA and RNA isolated to validate the presence of one of the *Xist* SNPs, rs31391216 (**Fig. 13B**). First, gDNA was amplified by PCR with primer specific for the DNA region containing the SNP followed by Sanger sequencing. The *Xist* SNP was detected in MEF #9 clone B12, MEF #9 clone C12 and MEF #2 clone E5 (**Fig. 14**). Next, primers specific to detect the SNP on the RNA level were used to PCR amplify the cDNA region containing the SNP, which was subjected to Sanger sequencing. We observed expression at the RNA level from one allele for clone B12 and C12 and the other allele for clone E5 (**Fig. 14**). The allele expressed by clone B12 and C12 corresponds to the maternal X-chromosome and the allele expressed by clone E5 corresponds to the paternal X-chromosome. Since *Xist* is only expressed from the Xi, this demonstrates that the maternal X-chromosome is inactivated in clones B12 and C12 and the paternal X-chromosome is inactivated in clone E5.

A)



B)

SNP ID	Map Position	C57BL/6J	FVB/NJ
rs31372548	ChrX:103474157	A	C
rs31391216	ChrX:103474838	C	T
rs31308397	ChrX:103474914	A	C
rs31295127	ChrX:103480851	G	C

Figure 13. Generation of mouse embryonic fibroblast cell lines containing single nucleotide polymorphisms in *Xist*.

A) Top: A female FVB/NJ mouse was crossed with a male C57BL/6J mouse and eleven embryonic fibroblast cell lines were isolated on day E13.5. Bottom: gDNA was extracted from the cells and PCR performed for the male *Sry* gene and control *Fabp2*. NC: negative control (gDNA from a female mouse), PC: positive control (gDNA from a male mouse). B) Table containing information regarding the SNPs in *Xist* according to MGI database (dbSNP Build 137). The red box highlights the *Xist* SNP used for characterization of the X-chromosome inactivation status of selected MEF clones.

Figure 14

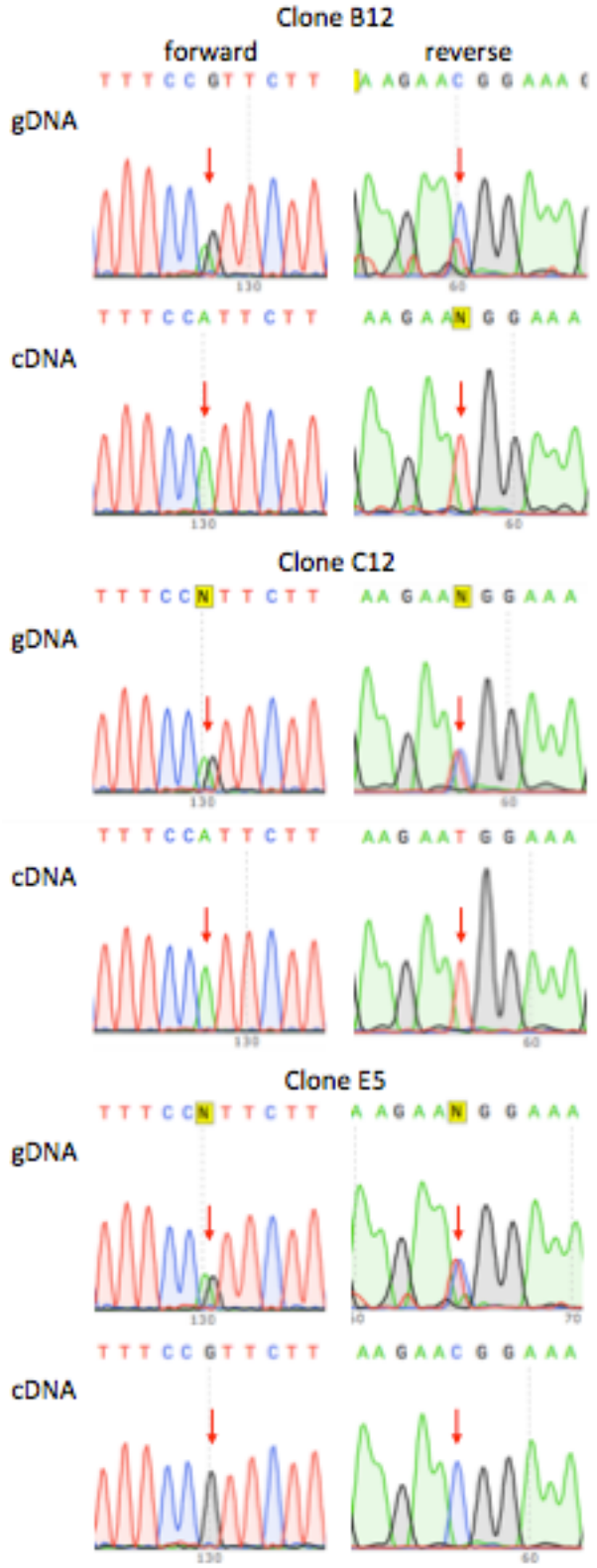

Clone B12
forward
reverse
gDNA
cDNA
Clone C12
gDNA
cDNA
Clone E5
gDNA
cDNA

Figure 14. Identification of paternal or maternal X-chromosome inactivation for three female MEF clones.

Sanger sequencing results of the forward and reverse strand for SNP rs31391216 in *Xist* at the gDNA and mRNA (cDNA) level of MEF clones B12, C12 and E5. The red arrow indicates the location of the SNP.

So far, we have shown that BORIS can change X-linked gene expression and have established a mouse cell line model with known parental origin of the Xi and presence of SNPs in X-linked genes. While the mouse and human full length Boris and BORIS protein, respectively is only 56% identical, the ZF domains are 82% identical (BLAST alignment, **Fig. 4**). Despite the high similarity between the ZFs, we specifically wanted to modulate Boris expression in the new cell lines to ensure proper functioning of the Boris protein with respect to N- and C-terminal interacting protein partners expressed in murine cells. During previous experiments with HK-BORIS expression in the MM057 cell line, we observed a reduction in Katushka positive cells by FACS over time. Due to this effect, it was challenging to obtain enough cells for IB and qPCR. Therefore, we decided to use a dox-inducible system, as this allows us to turn expression "on" and "off" and control for the level and duration of expression. MEF clones were stably infected to express either EV control or Boris fused to a 6x His-tag (EV-6xH and mBORIS-6xH, respectively). Cells with inducible EV-6xH expression were generated to control for the addition of dox to the media as well as the effect of transducing cells with an expression vector. To determine if dox has a toxic effect on the cells, the B12 clone was treated with an increasing concentration of dox for 96h. Compared to untreated cells, none of the dox concentrations tested displayed a toxic effect on the B12 cells (**Fig. 15A**). Next, we assessed expression of His-tagged Boris in the E5 clone by IB upon induction with dox, using an antibody directed against the 6x His-tag. Based on the expression of endogenous and non-tagged BORIS at ~90kDa, we anticipated the Boris protein slightly lower at ~85 kDa. As expected, we observed a band at ~80-85kDa for his-tagged Boris in the presence of dox for both MEF clones (**Fig. 15B**). However, the His antibody resulted in a high background signal, which made clear detection of Boris difficult. Since our human BORIS antibodies do not react with mouse Boris and no reliable mouse Boris antibody is commercially available [280], we were unable to confirm Boris expression using a human BORIS antibody. To circumvent the possibility that the detected band for his-tagged Boris is in fact a background band, we generated an inducible expression plasmid encoding either EV or Boris fused to a triple Flag (3xDYKDDDDK)-tag (EV-3xF and mBORIS-3xF, respectively), as we know that the Flag antibody does not produce background signal. Currently, we have validated expression of Flag-tagged Boris upon induction with dox in HEK293T cells (**Fig. 15C**). Once inducible Boris expression is established in the MEF clones B12 and E5, allele-specific expression of X-linked genes will be assessed by RNA-seq.

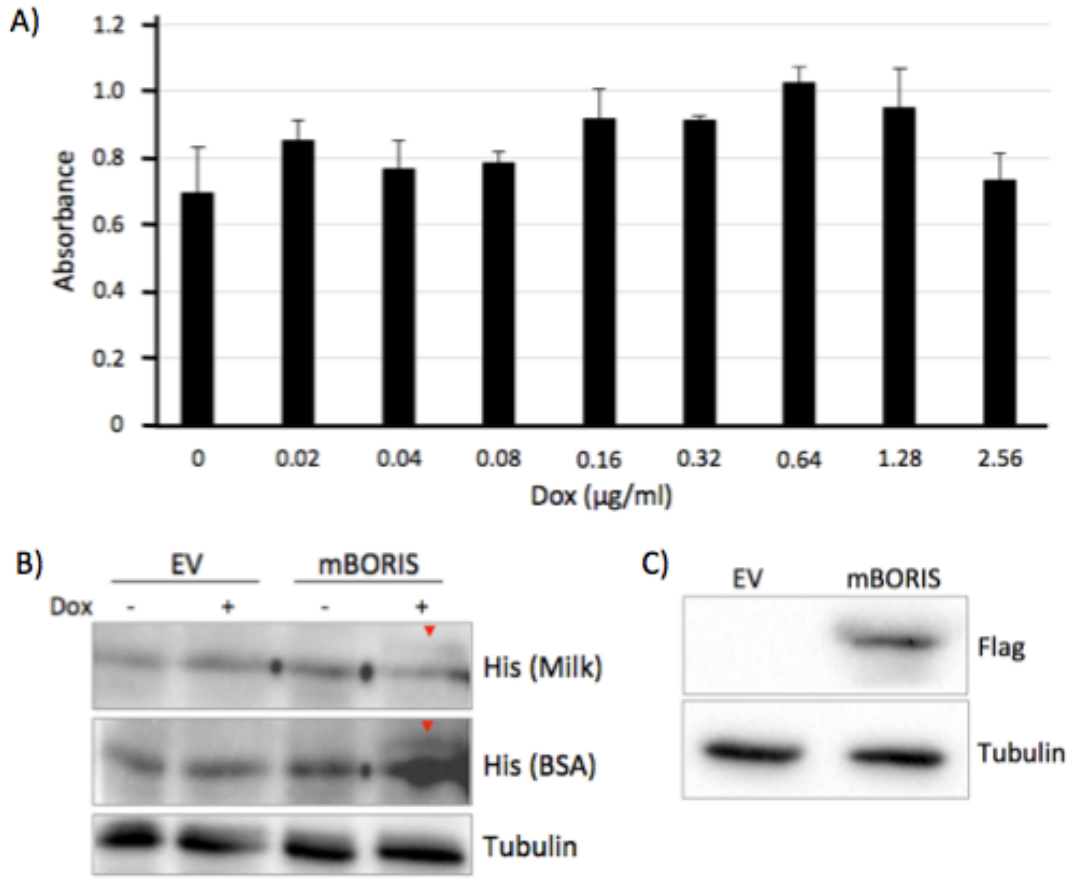


Figure 15. Establishing inducible Boris expression in the MEF clones.

A) MEF clone B12 was exposed to an increasing concentration of dox for 96 hours. The number of metabolically active cells in the absence and presence of dox was assessed by MTT assay and results are displayed as a bar graph showing the absorbance measured at 570nm. B) MEF clone E5 was genetically engineered to stably express EV-6xH or mBORIS-6xH upon induction with dox. Expression was induced with 3µg/ml dox for 48 hours and assessed by IB using anti-His antibody. The red arrows indicate His-tagged Boris protein. C) HEK293T cells were genetically engineered to express EV-3xF or mBORIS-3xF upon induction with dox. Cells were exposed to 1µg/ml dox for 48 hours and expression of Flag-tagged Boris was confirmed by IB using an anti-Flag antibody. For IBs anti-Tubulin was used as a loading control.

CHAPTER 3.2

–

BORIS Expression Alters the Transcriptome of Melanoma Cells

Sanne M. Janssen, Mounib Elchebly, Roy Moscona, Andreas Papdakis, Léon C.L. van Kempen, Eitan Rubin and Alan Spatz

Results in this chapter are included in the following manuscript currently under review:
Janssen SM, Moscona R., Elchebly M., Papadakis AI, van Kempen LCL, Rubin E, Spatz A. BORIS/CTCFL promotes a switch from a proliferative towards an invasive phenotype in melanoma cells. (BioRxiv doi: 10.1101/560474)

Background and objectives

In the previous chapter we demonstrated that BORIS can alter expression of genes from the X-chromosome in melanoma cells. However, BORIS expression has been linked to various cellular processes [142, 149], suggesting that BORIS can likely alter gene expression at a larger scale. Investigating genome-wide (differential) gene expression has become a major tool in molecular biology. Currently, RNA-seq is the standard method for gene expression profiling, as it provides higher specificity and sensitivity, and a wider dynamic range compared to microarrays. Despite the availability of improved techniques, only a few studies have assessed the effect of BORIS expression on the transcriptome of cancer cells [144, 195, 196]. Here we set out to address if BORIS expression can result in genome-wide differential gene expression in melanoma cells. Furthermore, we anticipate that the identification of these DEGs will allow us to identify cellular processes in which BORIS plays a role in melanoma cells.

Results

3.2.1 Establishing a melanoma cell line model with inducible BORIS expression

Due to the loss of cells with stable HK-BORIS expression as described in chapter 3.1.4, we aimed to establish a new model system with inducible hBORIS expression. From the characterization of BORIS expression in melanoma cell lines we knew that BORIS expression in the MM057 cell line is low (chapter 3.1, **Fig. 8**), which made it a good model to assess the effect of ectopic BORIS expression on gene expression. A dox-inducible system was established in the MM057 cell line to ectopically express either full-length hBORIS or EV control fused to a 6x His-tag (hBORIS-6xH and EV-6xH, respectively). Upon induction of BORIS expression for various time points we observed a dose dependent reduction in cell number over time compared to untreated cells, which was not observed in control cells (**Fig. 16A**). Expression of BORIS at each time point was confirmed using IB (**Fig. 16B**). Induction of BORIS expression with a concentration of 50ng/ml dox for 5 days resulted in a significant reduction in cell number without any cytotoxic effect of dox on the control cells (**Fig. 16A**). Subsequent experiments were performed under these conditions. Besides a reduction in cell number, no clear BORIS-expression induced phenotypic differences were observed in the MM057 cell line (**Fig. 16C**). In sum, we established a melanoma model system with dox-inducible BORIS expression and show that BORIS expression induced with a low, non-toxic dose of dox leads to a reduction in cell number.

Figure 16

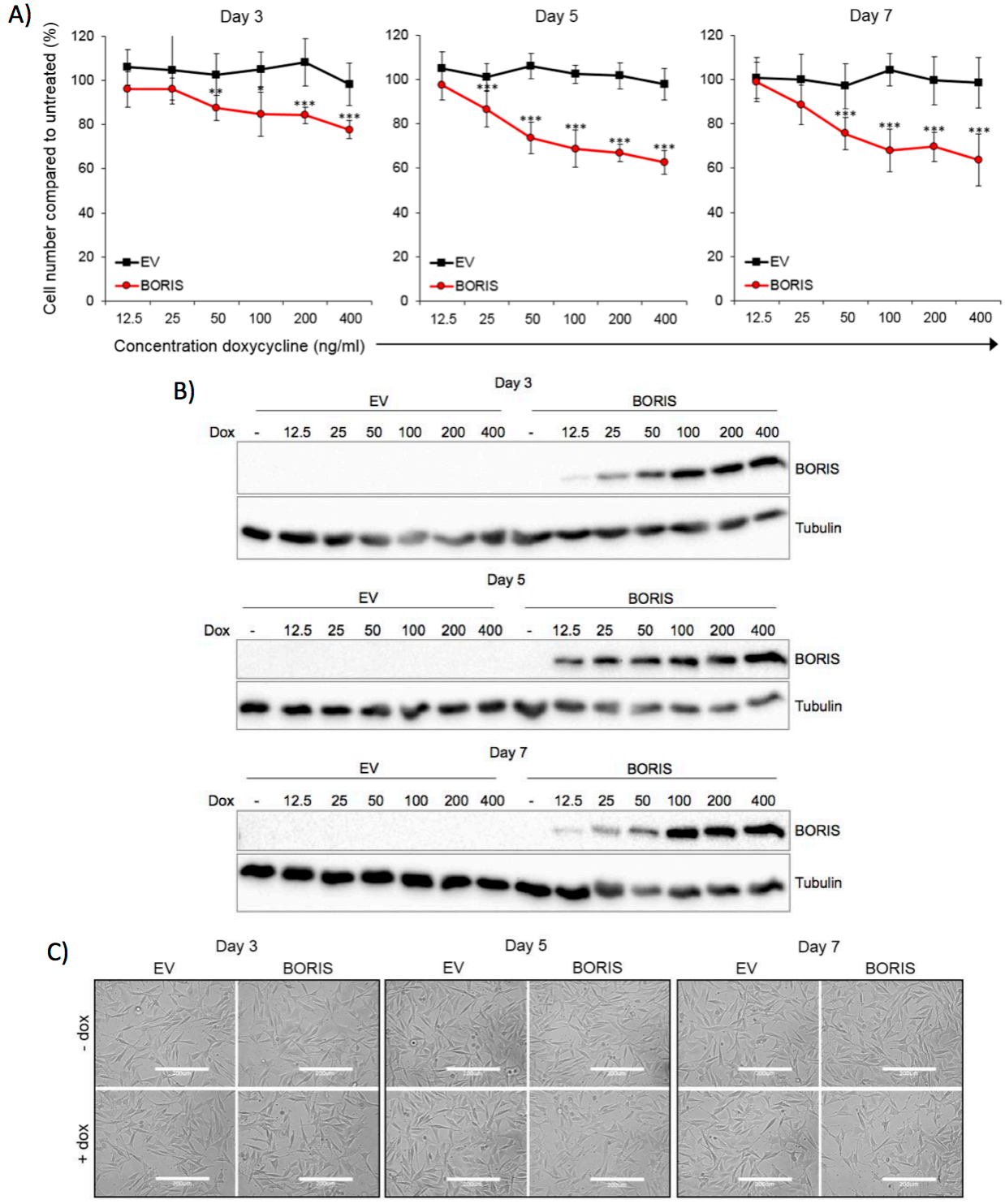

A)
Day 3
Day 5
Day 7
Cell number compared to untreated (%)
EV
BORIS
Concentration doxycycline (ng/ml)
B)
Day 3
Day 5
Day 7
EV
BORIS
Dox
BORIS
Tubulin
C)
- dox
+ dox

Figure. 16. A dose dependent reduction in cell number with ectopic BORIS expression.
A-C) MM057 cells were genetically engineered to stably express EV-6xH or hBORIS-6xHis upon induction with dox. Expression of BORIS protein was induced with an increasing dose of dox for 3, 5 or 7 days. A) Proliferation was assessed by measuring crystal violet absorbance at 570nm. B) Validation of BORIS expression by IB using anti-BORIS antibody. Anti-Tubulin was used as a loading control. C) Bright field images were made of BORIS and control expressing cells in the presence or absence of 50ng/ml dox. * indicates significantly different from untreated (*p<0.05. **p<0.01, ***p<0.001).

3.2.2 BORIS expression results in large-scale differential gene expression

Next, to address the effect of BORIS expression on the transcriptome of melanoma cells we performed RNA-seq on the MM057 cells with EV-6xH or hBORIS-6xH. We used the established concentration and exposure time of dox (50ng/ml for 5 days) to investigate DEGs between EV without dox (EVneg) and with dox (EVpos), and BORIS without dox (BORneg) and with dox (BORpos). Three independent experiments were performed for which BORIS expression was validated by qPCR and IB (**Fig. 17A**). A high correlation was observed between the replicates (Pearson correlation r≥0.97, **Fig. 17B**). Therefore, data of all replicates was used for downstream analysis. DEGs were defined using the following criteria: $|\log_2 FC| \geq 0.5$, FDR<0.05 (=$p \leq 0.01$) and CPM>0.5. To visualize expression differences between all samples we used multidimensional scaling based on the $\log_2 FC$. As expected, this analysis demonstrated clustering of the samples into two groups, one representing the controls and the other BORIS positive samples (**Fig. 18A**), suggesting substantial differences in gene expression between the controls and samples with BORIS expression. We observed a small batch effect between replicates, though this could be discarded based on the high correlation among replicates (**Fig. 17B**). The differential expression analysis revealed only 5 DEGs between EVneg and EVpos, which were likely differentially expressed as a result of dox and therefore removed from the list of DEGs between BORneg and BORpos. Between BORneg and BORpos, 2291 DEGs were detected, including 1490 upregulated and 801 downregulated genes (**Fig. 18B**). Analysis of the $\log_2 CPM$ for the top 100 up- and down-regulated genes demonstrated robust differential expression as well as clustering of all control samples versus BORIS positive samples (**Fig. 18C**). To validate our RNA-seq dataset, DEGs were ranked by FDR and differential expression of every 300^{th} upregulated gene (**Fig. 19A**) and 200^{th} downregulated gene (**Fig. 19B**), including the last gene of the list, was validated by qPCR. Together, these data demonstrate that ectopic BORIS expression results in large-scale differential gene expression and confirm that we have a robust set of DEGs.

Figure 17

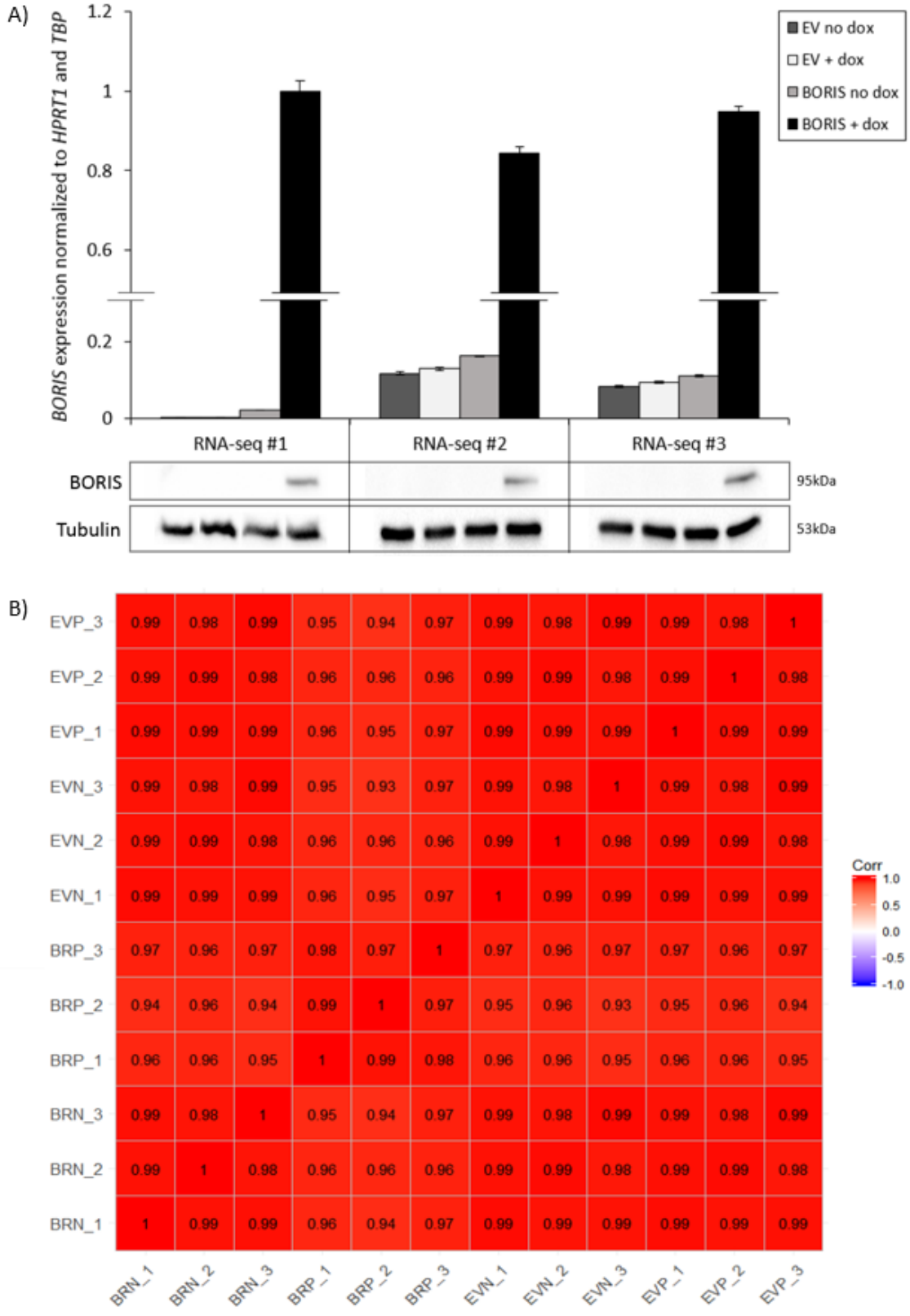

A)
BORIS expression normalized to HPRT1 and TBP
1.2
1
0.8
0.6
0.2
0
EV no dox
EV + dox
BORIS no dox
BORIS + dox
RNA-seq #1
RNA-seq #2
RNA-seq #3
BORIS
Tubulin
95kDa
53kDa
B)
EVP_3 0.99 0.98 0.99 0.95 0.94 0.97 0.99 0.98 0.99 0.99 0.98 1
EVP_2 0.99 0.99 0.98 0.96 0.96 0.96 0.99 0.99 0.98 0.99 1 0.98
EVP_1 0.99 0.99 0.99 0.96 0.95 0.97 0.99 0.99 0.99 1 0.99 0.99
EVN_3 0.99 0.98 0.99 0.95 0.93 0.97 0.99 0.98 1 0.99 0.98 0.99
EVN_2 0.99 0.99 0.98 0.96 0.96 0.96 0.99 1 0.98 0.99 0.99 0.98
EVN_1 0.99 0.99 0.99 0.96 0.95 0.97 1 0.99 0.99 0.99 0.99 0.99
BRP_3 0.97 0.96 0.97 0.98 0.97 1 0.97 0.96 0.97 0.97 0.96 0.97
BRP_2 0.94 0.96 0.94 0.99 1 0.97 0.95 0.96 0.93 0.95 0.96 0.94
BRP_1 0.96 0.96 0.95 1 0.99 0.98 0.96 0.96 0.95 0.96 0.96 0.95
BRN_3 0.99 0.98 1 0.95 0.94 0.97 0.99 0.98 0.99 0.99 0.98 0.99
BRN_2 0.99 1 0.98 0.96 0.96 0.96 0.99 0.99 0.98 0.99 0.99 0.98
BRN_1 1 0.99 0.99 0.96 0.94 0.97 0.99 0.99 0.99 0.99 0.99 0.99
BRN_1 BRN_2 BRN_3 BRP_1 BRP_2 BRP_3 EVN_1 EVN_2 EVN_3 EVP_1 EVP_2 EVP_3
Corr
1.0
0.5
0.0
-0.5
-1.0

Figure 17. RNA-sequencing of melanoma cells with inducible BORIS expression.

MM057 melanoma cells were genetically engineered to stably express EV-6xH or hBORIS-6xH upon induction with dox. BORIS expression was induced with 50ng/ml dox for 5 days and RNA-seq was performed in triplicate. A) Expression of BORIS mRNA and protein was confirmed by qPCR and IB, respectively. IB was performed using anti-BORIS antibody and anti-Tubulin was used as a loading control. B) Correlation matrix displaying Pearson's correlation between each sample. EVN: EVneg, EVP: EVpos, BRN: BORneg, BRP: BORpos.

Figure 18

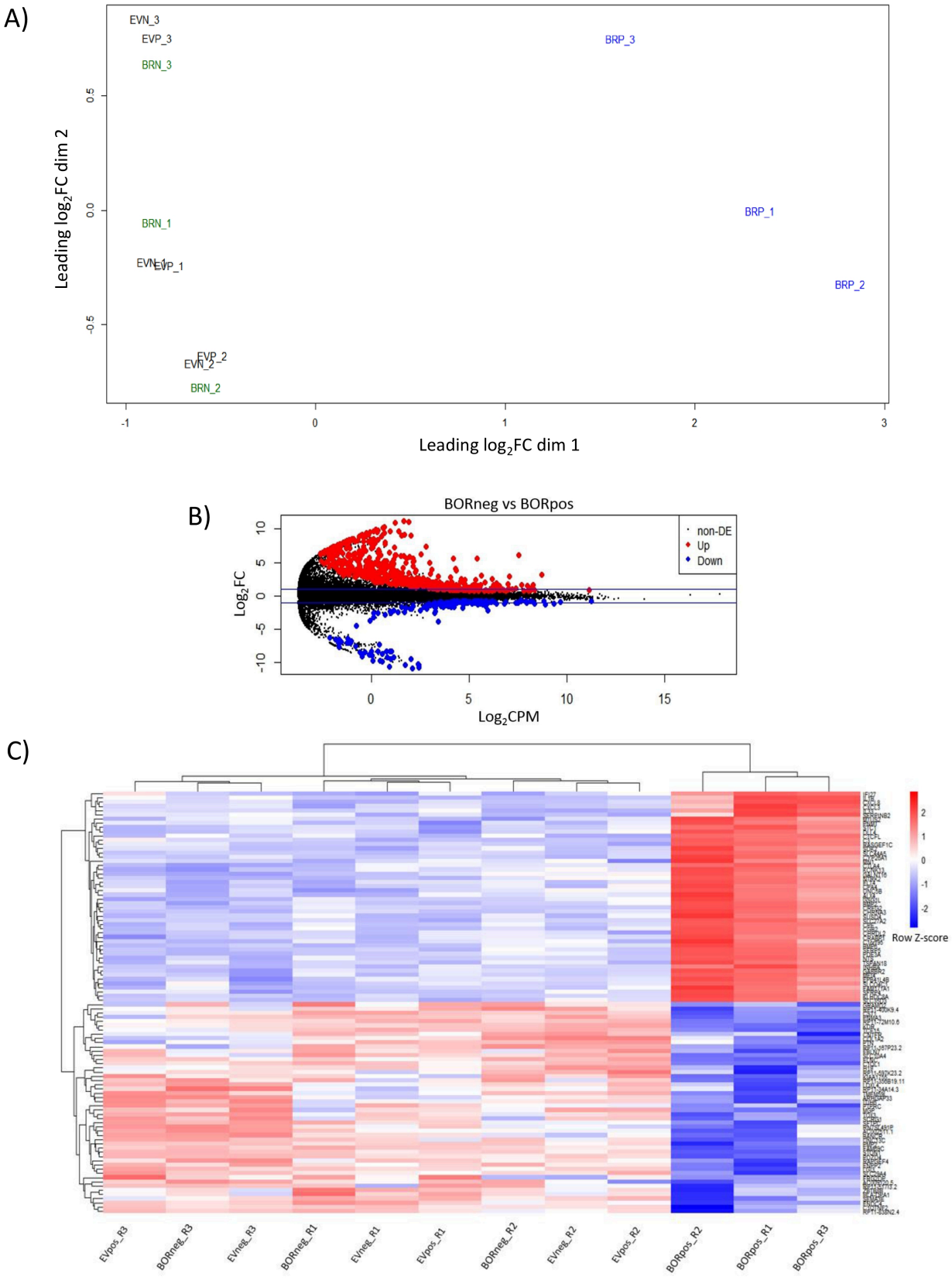

A)
EVN_3
EVP_3
BRN_3
BRP_3
BRP_1
BRN_1
EVN_1
EVP_1
BRP_2
EVP_2
EVN_2
BRN_2
Leading log2FC dim 2
Leading log2FC dim 1
B)
BORneg vs BORpos
non-DE
Up
Down
Log2FC
Log2CPM
C)
Row Z-score
EVpos_R3
BORneg_R3
EVneg_R3
BORneg_R1
EVneg_R1
EVpos_R1
BORneg_R2
EVneg_R2
EVpos_R2
BORpos_R2
BORpos_R1
BORpos_R3

Figure 18. Ectopic BORIS expression results in robust differential gene expression.

A) Expression differences between and clustering of all RNA-seq samples is displayed in a multidimensional scaling plot based on the $\log_2$ fold change ($\log_2$FC). B) MA plot of the differentially expressed genes between BORneg and BORpos. Red dots indicated significantly upregulated genes and blue dots significantly downregulated genes. C) Heat map based on the RNA-seq data of all samples visualizing the $\log_2$ counts per million ($\log_2$CPM) for the top 100 up- and down-regulated genes.

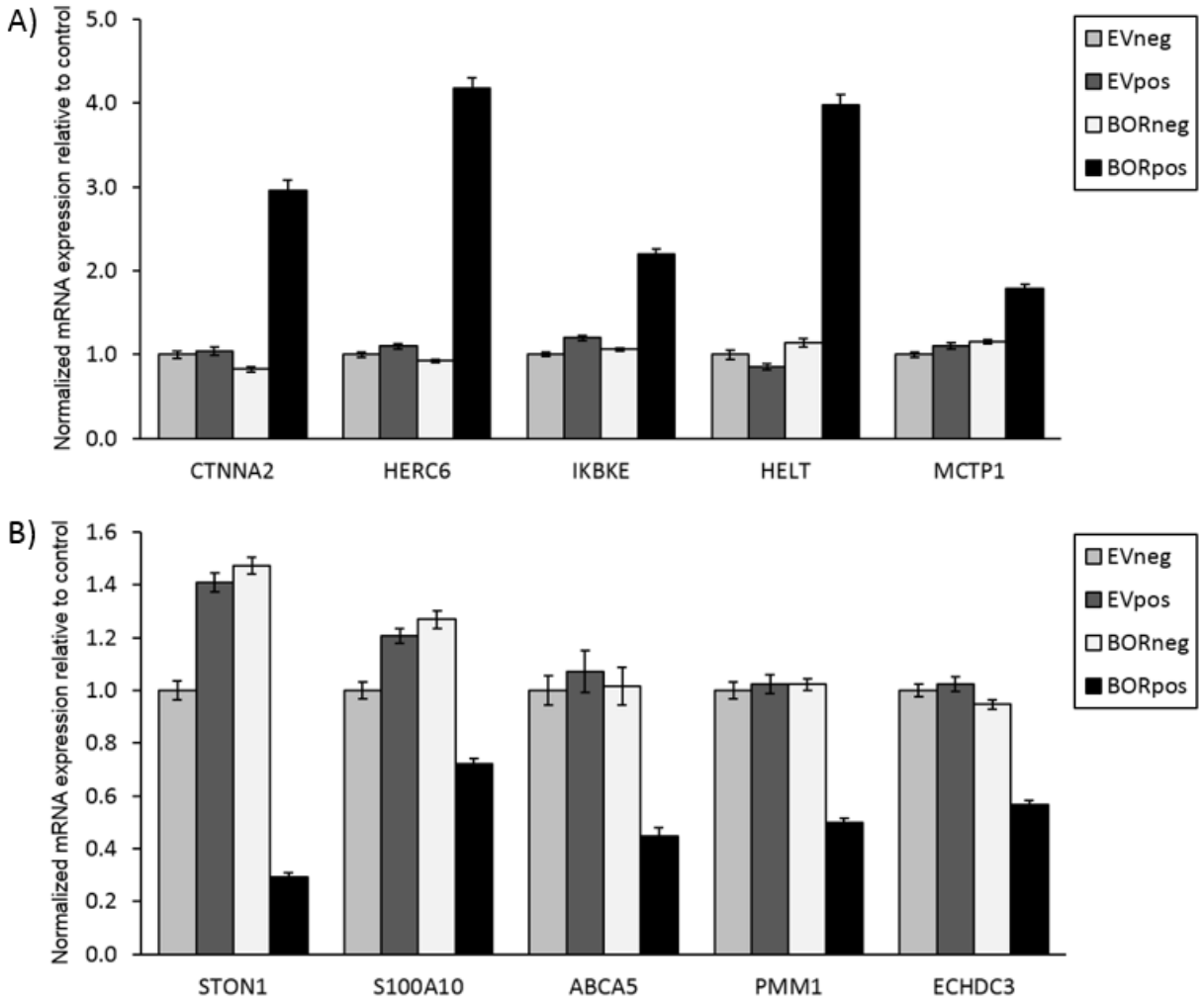


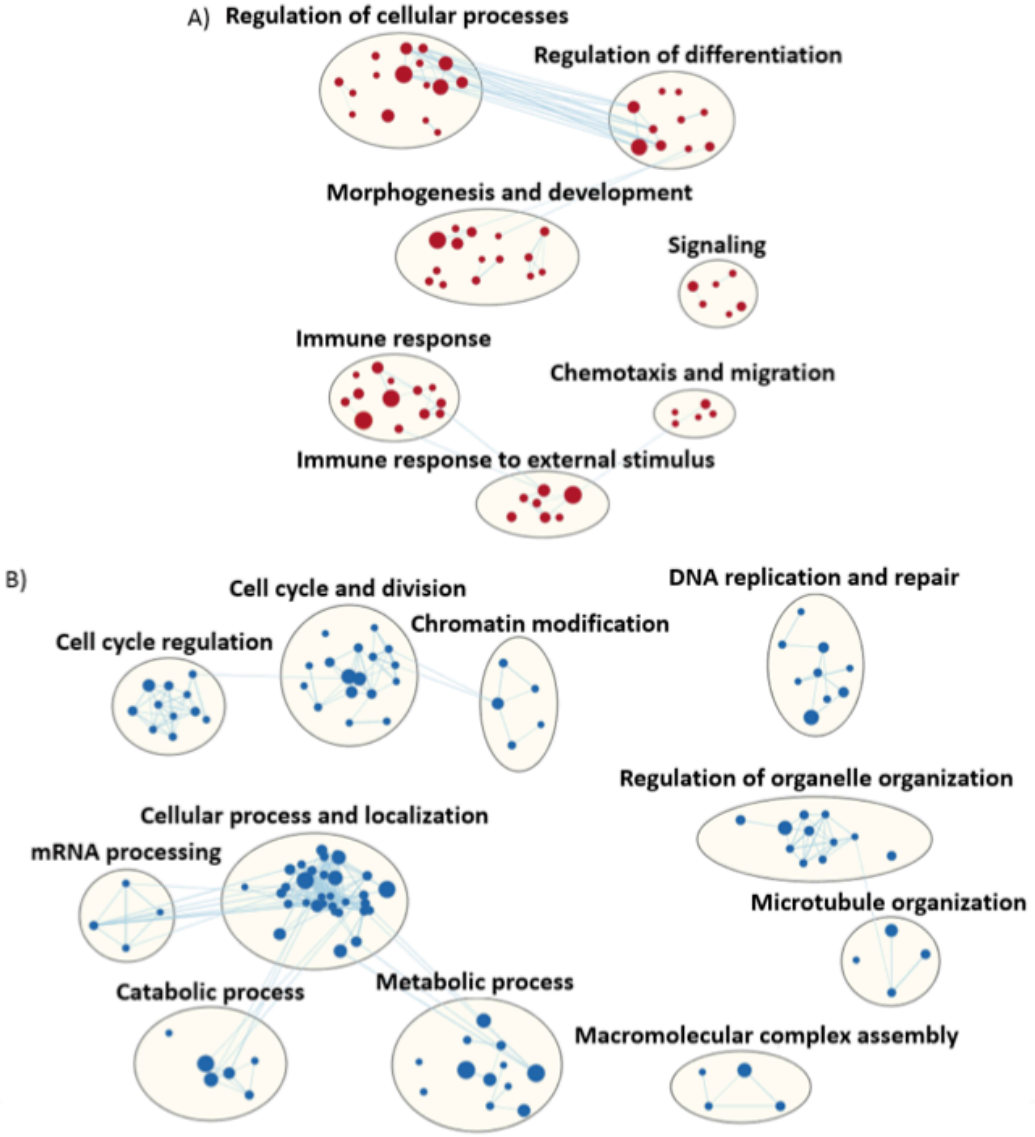
Figure 19. Validation of the RNA-sequencing data.

Up- and down-regulated genes were ranked by their FDR from most significant to least significant. A) qPCR for every 300th upregulated gene, including the last gene of the list. B) qPCR for every 200th downregulated gene, including the last gene of the list. For all qPCR experiments, a technical triplicate was performed for two biological replicates. Expression was normalized to *HPRT1* and *TBP*. Data was analyzed with the Bio-Rad CFX Manager™ Software. Error bars represent the standard error of the mean.

3.2.3 Identification of BORIS functions by gene set enrichment and overrepresentation analysis

Now that we confirmed to have a robust RNA-seq data set we wanted to use this data to identify cellular processes in which BORIS is involved. GSEA [247] was performed for biological processes on a pre-ranked gene list. The ranked gene list was established by including genes with a CPM>0.5, FDR<0.2 (=$p \leq 0.05$) and ranked according to their $\log_2$FC. This analysis revealed a significant (FDR<0.05) positive enrichment for 90 processes (**Appendix III**), meaning that these processes are enriched among upregulated genes. Furthermore, it revealed a significant (FDR<0.05) negative enrichment for 130 processes (**Appendix III**), meaning that these processes are enriched among downregulated genes. For better visualization, the enriched biological processes were clustered and annotation was manually edited for enhanced comprehension (**Fig. 20A and B**). Based on the clustering, the top 3 positively enriched biological processes include, cell-cell signaling, response to type I interferon and defense response and the top 3 negatively enriched biological processes include, cell cycle, RNA processing and DNA metabolic process (**Fig. 20C**). According to these data BORIS DEGs are negatively enriched for processes involved in cell cycle and division (**Fig. 20B and C**), which is in agreement with our observation that BORIS expression results in reduced cell number (**Fig. 16A**).

Figure 20



C)

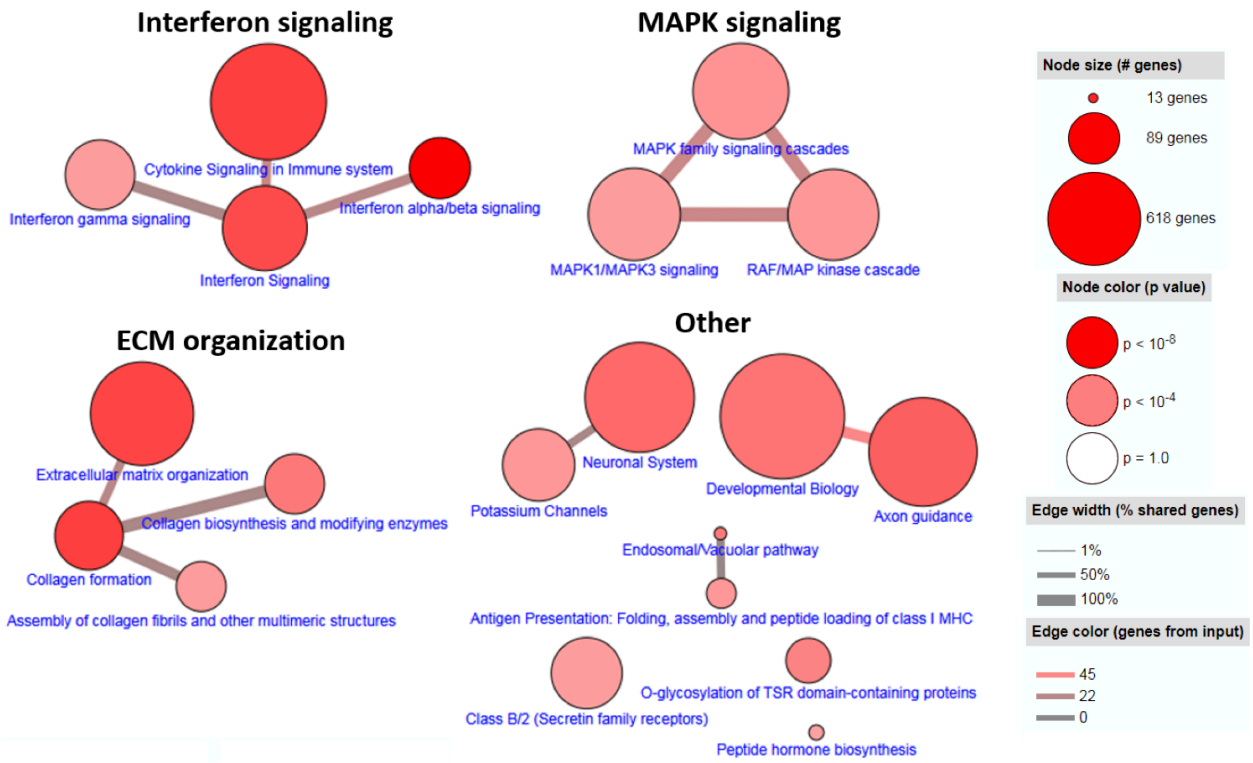
Biological process (positive enrichment)	Cluster	NES	p-value	FDR	ID
Cell-cell signaling	Signaling	3.35	<0.0001	<0.0001	GO:0007267
Response to type I interferon	Immune response	3.25	<0.0001	<0.0001	GO:0034340
Defense response	Immune response to external stimulus	3.06	<0.0001	2.10E-04	GO:0006952

Biological process (negative enrichment)	Cluster	NES	p-value	FDR	ID
Cell cycle	Cell cycle and division	-5.83	<0.0001	<0.0001	GO:0007049
RNA processing	mRNA processing	-4.62	<0.0001	<0.0001	GO:0006396
DNA metabolic process	DNA replication and repair	-3.89	<0.0001	<0.0001	GO:0006259

Figure 20. Gene set enrichment analysis for biological processes on differentially expressed genes.

GSEA was performed for biological processes on the significant DEGs between BORneg and BORpos ranked according to the log_2FC. A and B) The Cytoscape [249] plugin EnrichmentMap [250] was used to generate a network map of enriched gene ontologies. Red and blue nodes represent gene ontologies and edges represent mutual overlap, which are used to cluster highly redundant gene ontologies (clustered within ellipses). The node size represents the number of DEGs represented within the gene ontology, the more genes the larger the node size. The Cytoscape plugins AutoAnnotate [251] and WordCloud [252] were used for initial annotation, which was manually edited for enhanced comprehension. A) For positive enrichment, 12 gene ontologies did not cluster and were removed from the image. B) For negative enrichment, 14 gene ontologies did not cluster and were removed from the image. C) Overview of the top 3 positively and negatively enriched biological processes as identified by gene set enrichment analysis according to the network clustering.

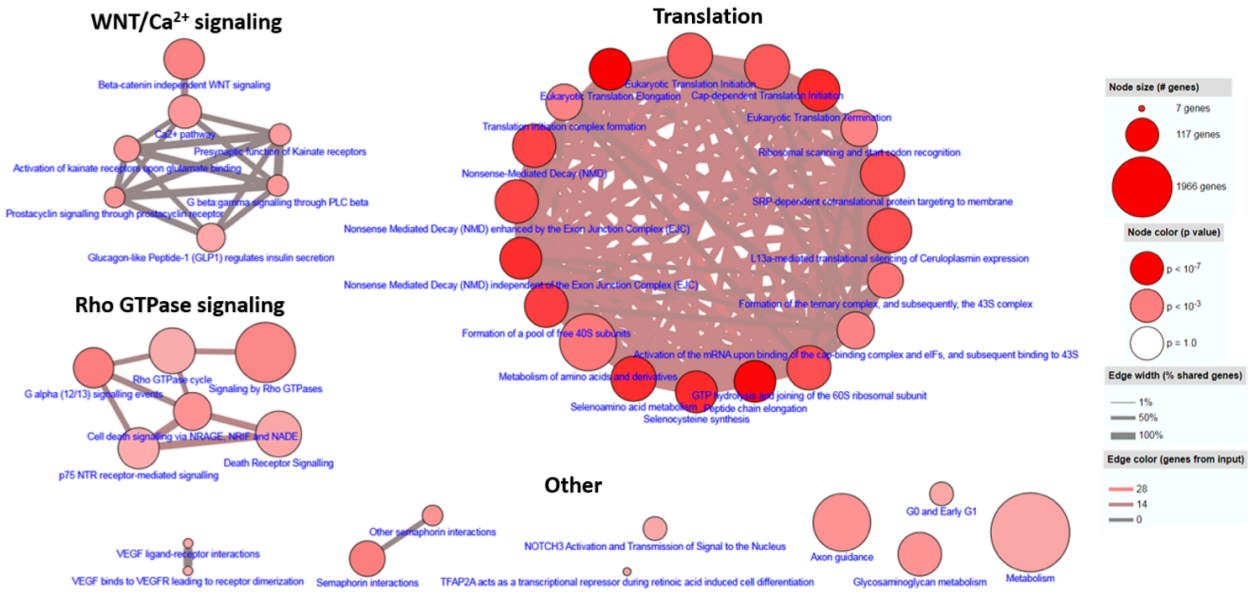
In addition to GSEA for biological processes, overrepresentation analysis was performed for Reactome pathways on the list of upregulated and downregulated DEGs. The analysis was performed in the CPDB [245] and revealed 20 significantly overrepresented pathways ($q < 0.05$) among the upregulated genes and 42 among the downregulated genes (**Appendix III**). The concept overlap graph function within CPDB was used to cluster, manually annotate and visualize the results. Based on the clustering, the top 3 overrepresented pathways for the upregulated DEGs included, interferon alpha/beta signaling, collagen formation, and MAPK family signaling cascades (**Fig. 21**). For the downregulated genes these included, eukaryotic translation elongation, G alpha (12/13) signaling events, and beta-catenin independent Wnt signaling (**Fig. 22**). Overall, these analyses point to a role for BORIS in; 1) various signaling pathways (MAPK, interferon, WNT/Ca^{2+}, and Rho GTPase), 2) migration and ECM organization, 3) immune response, 4) cell cycle and division, 5) mRNA processing, 6) DNA replication and repair, and 7) translation.



Pathway (upregulated genes)	Cluster	p-value	q-value	ID
Interferon alpha/beta signaling	Interferon signaling	7.95E-09	5.66E-06	R-HSA-909733
Collagen formation	ECM organization	8.91E-07	2.12E-04	R-HSA-1474290
MAPK family signaling cascades	MAPK signaling	4.63E-04	2.75E-02	R-HSA-5683057

Figure 21. Pathway overrepresentation analysis on upregulated genes.

Overrepresentation analysis was performed for Reactome pathways on the significantly upregulated genes between BORneg and BORpos. The overlap graph function within the CPDB [245] was used to generate a network map of overrepresented pathways. Clusters were manually annotated for enhanced comprehension. The table provides an overview of the top 3 overrepresented pathways among upregulated genes based on the network clustering.



Pathway (downregulated genes)	Cluster	p-value	q-value	ID
Eukaryotic Translation Elongation	Translation	1.29E-08	7.74E-06	R-HSA-156842
G alpha (12/13) signaling events	Rho GTPase signaling	1.05E-04	3.69E-03	R-HSA-416482
Beta-catenin independent WNT signaling	WNT/Ca^{2+} signaling	1.81E-04	5.15E-03	R-HSA-3858494

Figure 22. Pathway overrepresentation analysis on downregulated genes.

Overrepresentation analysis was performed for Reactome pathways on the significantly downregulated genes between BORneg and BORpos. The overlap graph function within the CPDB [245] was used to generate a network map of overrepresented pathways. Clusters were manually annotated for enhanced comprehension. The table provides an overview of the top 3 overrepresented pathways among downregulated genes based on the network clustering.

3.2.4 BORIS expression contributes to a switch from a proliferative to invasive state

Three lines of evidence suggest a role for BORIS in phenotype switching. First, previous studies have revealed that BORIS expression can modulate the transition between epithelial and mesenchymal cell states in breast cancer and neuroblastoma cell lines [214, 215]. Second, our analysis for enriched biological processes and overrepresented pathways revealed a potential role for BORIS in migration, extracellular matrix organization, and interferon, MAPK and non-canonical Wnt signaling. These processes and pathways are associated with phenotype switching [56]. Third, our observation of a reduced number of cells upon BORIS expression (**Fig. 16A**), suggests that BORIS expression leads to decreased proliferation. To our advantage, a recent study by Verfaillie *et al.* characterized a panel of 11 melanoma cell lines by RNA-seq and found that the MM057 cell line harbors a proliferative gene signature [59]. Based on these data, we set out to determine if BORIS expression can mediate phenotype switching in the MM057 melanoma cell line.

As mentioned, our initial finding that BORIS expression in the MM057 cells results in a reduced number of cells (**Fig. 16A**) supports the idea that BORIS expression leads to reduced proliferation. However, the reduced number of cells could also be the result of increased apoptosis or senescence. We first investigated if cell death could be responsible for the decrease in cell number. Compared to EV-6xH control cells, expression of hBORIS-6xH in the MM057 cell line resulted in increased expression of cleaved PARP, a marker for cell death (**Fig. 23A**). In addition, flow cytometry analysis showed an increase in the number of early apoptotic, apoptotic as well as necrotic cells with BORIS expression compared to controls (**Fig. 23B and C**), further indicating the cells were undergoing apoptosis. However, the percentage of death cells was only ~12%, which cannot solely explain the 30% reduction in cell number observed in the same timeframe of 5 days. We therefore also assessed senescence, but did not observe any increase in senescent cells (data not shown). These data show that expression of BORIS results in apoptosis and suggests that in addition other cellular processes, like proliferation, might play a role in the BORIS-associated reduction of cell number.

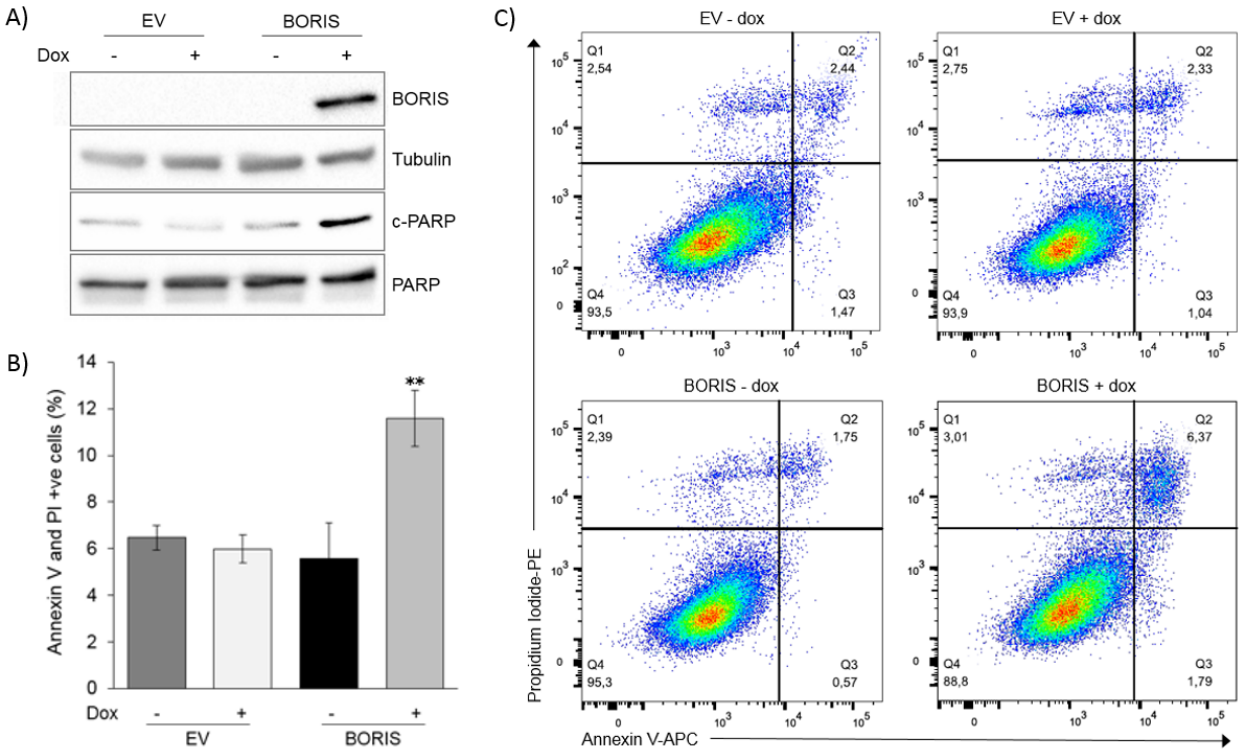


Figure 23. Increased cell death with ectopic BORIS expression.

Expression of EV-6xH and hBORIS-6xH was induced in MM057 cells with 50ng/ml dox for 5 days. (A) Whole cell lysate was used for IB with anti-BORIS, anti-cleaved PARP (c-PARP) and anti-PARP antibodies. Anti-Tubulin was used as a loading control (B) Bar graph representing the percentage Annexin V and PI positive cells as analyzed by FlowJo®. (C) Representative image of the percentage of apoptotic and necrotic cells as determined by flow cytometry analysis of Annexin V and PI staining. ** indicates significant difference compared to controls (p<0.01).

To investigate the effect of BORIS expression on phenotype switching, the pre-ranked gene list from the RNA-seq data was used to perform GSEA. First, we looked for enrichment among the hallmark EMT gene set [248] and observed a significant enrichment (normalized enrichmet score (NES)=3.12, FDR<0.0001, **Fig. 24A**), meaning that upregulated DEGs are enriched among the genes that belong to the hallmark EMT signature. Since melanoma cells do not undergo typical EMT, we next performed GSEA for the Hoek [51] and Verfaillie [59] gene signatures that correspond to either the proliferative or invasive state of melanoma cells. For these gene sets, we identified a strong positive correlation with the Hoek (NES=2.21, FDR=0.001) and Verfaillie (NES=5.99, FDR<0.0001) invasive gene signatures (**Fig. 24B**) and a negative correlation with the Hoek (NES=-1.98, FDR=0.006) and Verfaillie (NES=-3.84, FDR<0.0001) proliferative gene signatures (**Fig. 24C**), suggesting BORIS expression promotes a transcriptional switch from a proliferative to invasive state. We observed 17.9% (115 out of 642) of the invasive genes from the Verfaillie invasive gene signature among BORIS upregulated DEGs (**Fig. 24D**), highlighting the significance of the contribution of BORIS expression to the switch from a proliferative to invasive gene signature. In addition, qPCR analysis was performed for four "invasive" and four "proliferative" genes that were among the identified BORIS DEGs and present in both the Hoek and Verfaillie gene signatures. We observed significant upregulation of the "invasive" genes *DKK1, FBN1, INHBA* and *VEGFC* (**Fig. 25A**) and significant downregulation of the "proliferative" genes *GYG2, LZTS1, MLANA* and *TFAP2A* (**Fig. 25B**) upon hBORIS-6xH expression in the MM057 cells compared to controls. Together, these results indicate that BORIS promotes a switch from a proliferative to invasive gene signature.

Figure 24

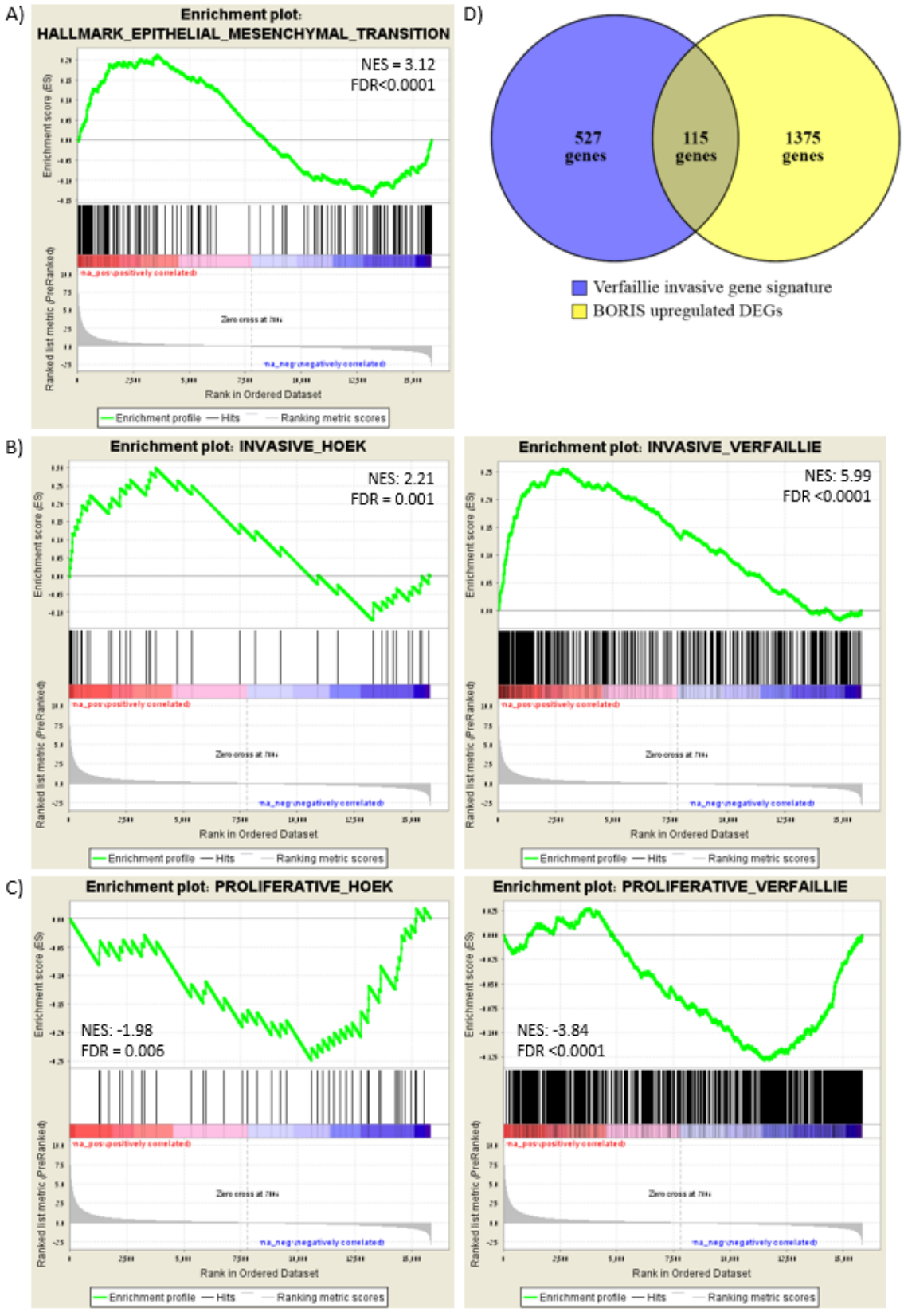

A)
Enrichment plot:
HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION
NES = 3.12
FDR<0.0001
Enrichment score (ES)
Ranked list metric (PreRanked)
Rank in Ordered Dataset
Enrichment profile
Hits
Ranking metric scores
D)
527 genes
115 genes
1375 genes
Verfaillie invasive gene signature
BORIS upregulated DEGs
B)
Enrichment plot: INVASIVE_HOEK
NES: 2.21
FDR = 0.001
Enrichment plot: INVASIVE_VERFAILLIE
NES: 5.99
FDR <0.0001
C)
Enrichment plot: PROLIFERATIVE_HOEK
NES: -1.98
FDR = 0.006
Enrichment plot: PROLIFERATIVE_VERFAILLIE
NES: -3.84
FDR <0.0001

Figure 24. BORIS promotes a switch from a proliferative to invasive gene signature.
A-C) GSEA on the RNA-seq data (ranked according to log_2FC) for A) the hallmark EMT gene set from the Molecular Signatures Database [248], and the 'Hoek' [51] and 'Verfaillie' [59] A) invasive and B) proliferative gene signatures. D) Venn diagram [281] of the overlap between invasive genes from the Verfaillie gene signature and BORIS upregulated DEGs. NES: normalized enrichment score, FDR: false discovery rate.

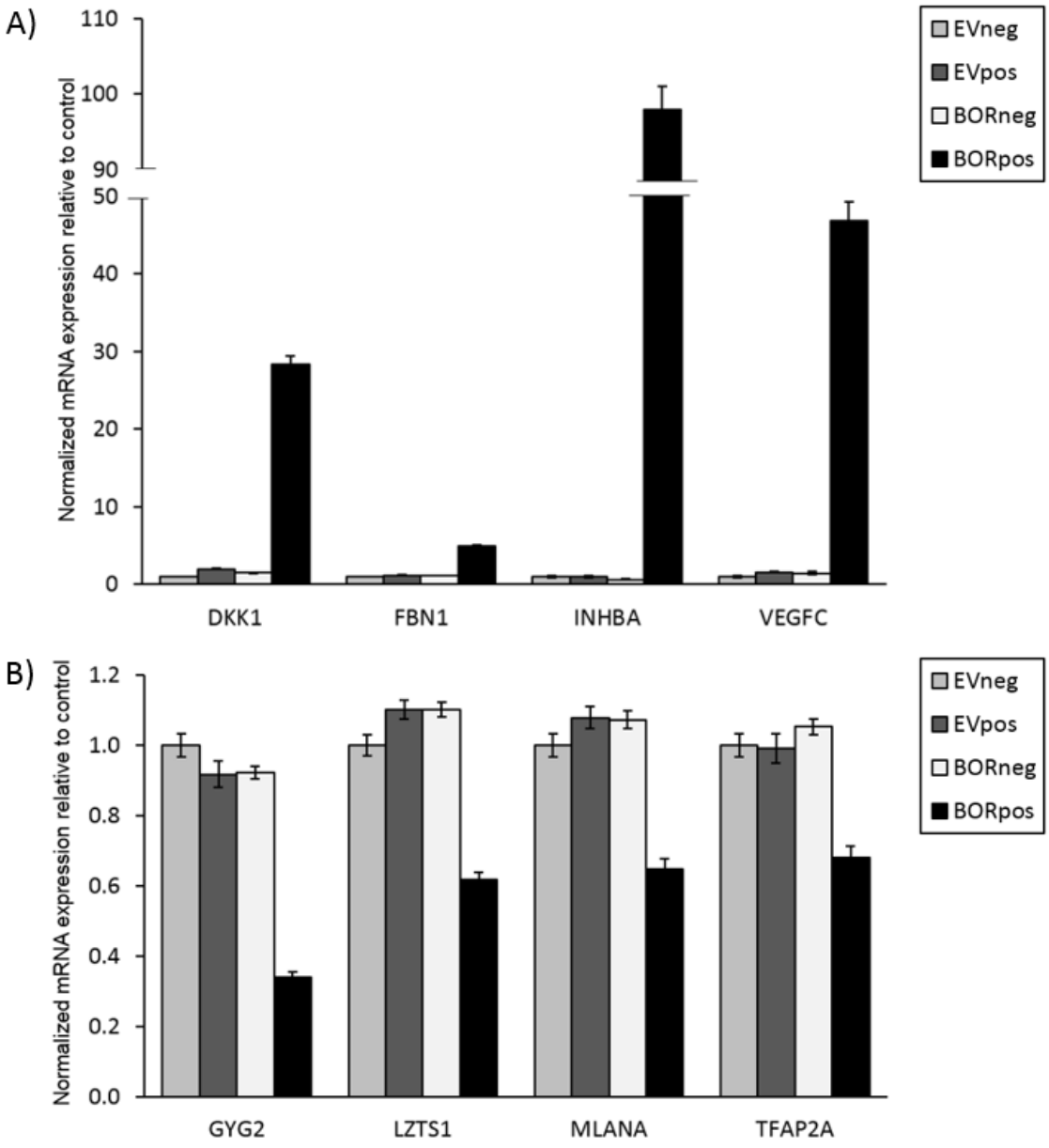


Figure 25. BORIS expression results in the upregulation of invasive genes and downregulation of proliferative genes.

EV-6xH and hBORIS-6xH expression was induced in MM057 cells with 50ng/ml dox for 5 days followed by qPCR for the indicated A) "invasive" and B) "proliferative" genes. For each sample, a technical triplicate was performed for two biological replicates. Expression was normalized to *HPRT1* and *TBP*. Data was analyzed with the Bio-Rad CFX Manager™ Software. Error bars represent the standard error of mean.

In melanoma, reduced expression of MITF has been reported in invasive cells compared to proliferative cells [32]. MITF is regulated by the switch in expression from ZEB2 in melanocytes to ZEB1 in melanoma cells, with ZEB1 functioning as a transcriptional repressor of E-cadherin (*CDH1*) [34, 35]. In the MM057 cells BORIS did not significantly reduce *MITF* expression (**Figure 26A**). Similarly, BORIS did not significantly alter *ZEB1*, *ZEB2* (**Figure 26A**) or *CDH1* expression (not expressed). In contrast, expression of the transcription factor and melanoma EMT inducer SNAIL (*SNAI1*) [282] was increased in BORIS-expressing MM057 cells (**Figure 26A**). Expression of SNAIL can be upregulated via multiple mechanisms, including elevated levels of the matricellular protein SPARC [283] and TGF-beta [282]. Indeed, we observed significantly higher expression of *SPARC* and *TGFB1* in BORIS-induced cells compared to non-induced cells (**Figure 26B**). Furthermore, the RNA-seq data reveals increased transcription of multiple genes belonging to the TGF-beta family (**Figure 26C**). These data suggest that the BORIS-induced transcriptional switch may be mediated via TGF-beta signaling and increased *SNAI1* expression.

Next, we aimed to address if, in addition to the transcriptional switch from a proliferative to invasive signature, the cells undergo a phenotypic switch as well. To this end, we induced hBORIS-6xH expression with dox in the MM057 cells and performed *in vitro* transwell cell migration and invasion assays. Compared to the control cells we observed an increased number of cells migrating (**Fig. 27A**) and invading (**Fig. 27C**) through the transwell membrane in the presence of BORIS. Similar results were obtained upon expression of hBORIS-6xH in the proliferative melanoma cell line MM074 for both migration (**Fig. 27B**) and invasion (**Fig. 27D**). These data indicate that BORIS expression promotes gene expression changes that lead to an invasive transcriptome and phenotype.

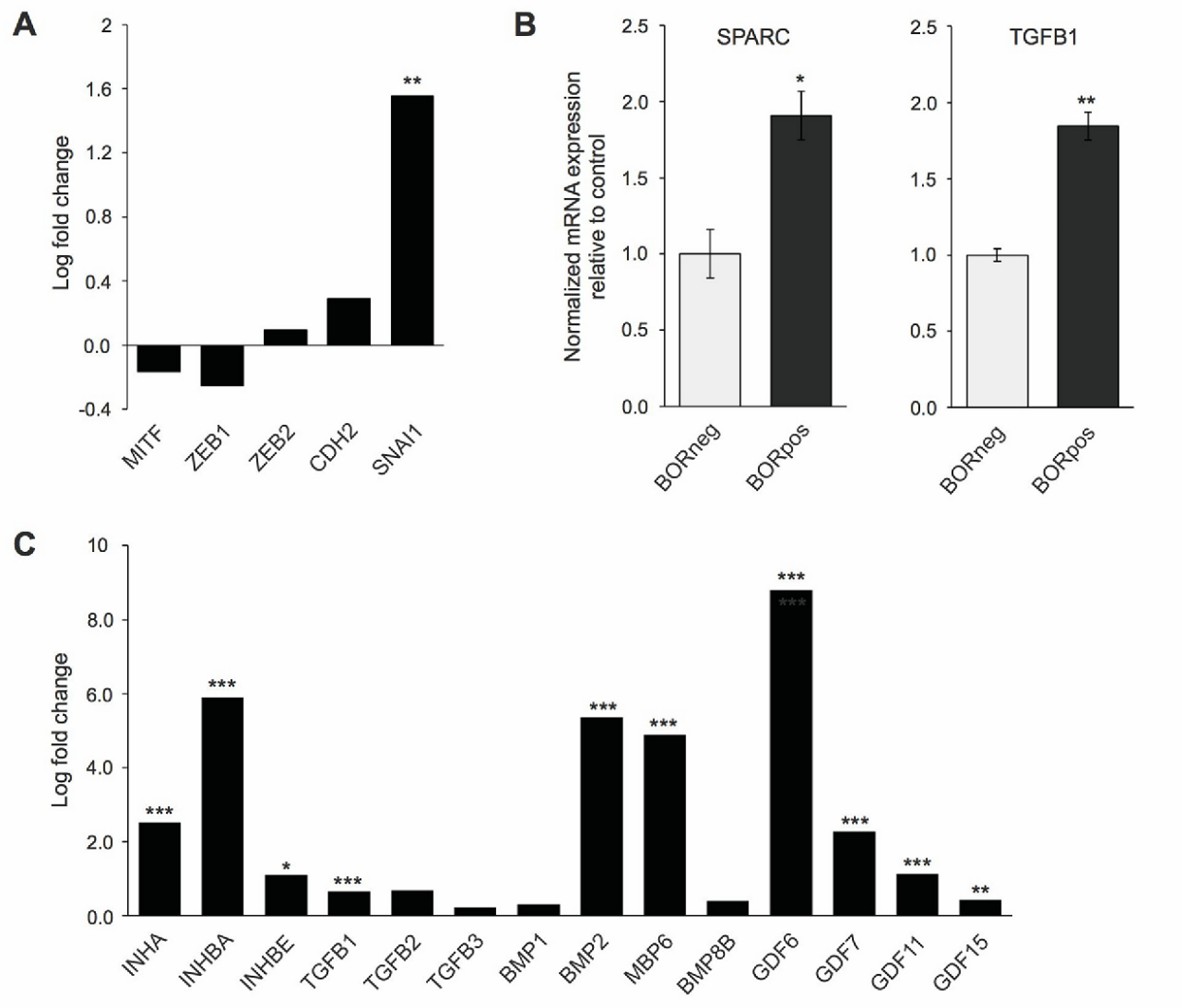


Figure 26. BORIS expression leads to increased expression of *SNAI1*, *SPARC* and members of the TGF-beta family.

Bar graph representing the $\log_2$ fold change between BORneg and BORpos RNA-seq samples for A) EMT-TFs and C) TGF-beta family members. B) EV-6xH and hBORIS-6xH expression was induced in MM057 cells with 50ng/ml dox for 5 days followed by qPCR for *SPARC* and *TGFB1*. For each sample, a technical triplicate was performed for two biological replicates. Expression was normalized to *HPRT1* and *TBP*. Error bars represent the standard error of mean. * indicates significantly different from BORneg (*p<0.05. **p<0.01, ***p<0.001).

Figure 27

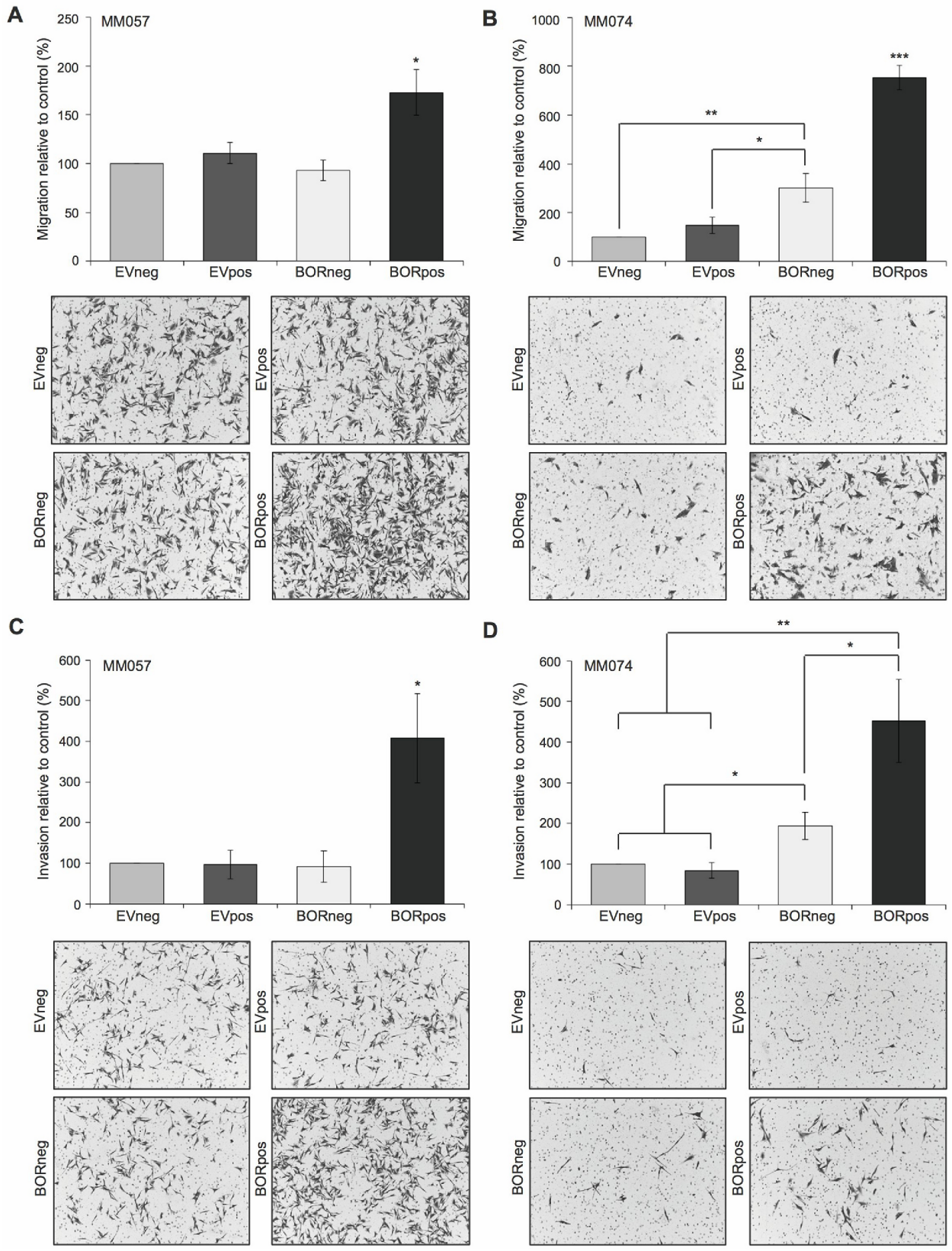

A
MM057
Migration relative to control (%)
250
200
150
100
50
0
*
EVneg
EVpos
BORneg
BORpos
B
MM074
Migration relative to control (%)
1000
800
600
400
200
0
**
*
***
EVneg
EVpos
BORneg
BORpos
C
MM057
Invasion relative to control (%)
600
500
400
300
200
100
0
*
EVneg
EVpos
BORneg
BORpos
D
MM074
Invasion relative to control (%)
600
500
400
300
200
100
0
**
*
*
EVneg
EVpos
BORneg
BORpos

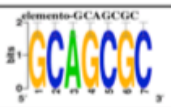
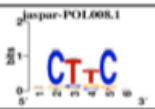
Figure 27. BORIS expression leads to increased migration and invasion.

Expression of EV-6xH and hBORIS-6xH was induced in A, C) MM057 cells or B, D) MM074 cells with 100ng/ml dox for 5 days followed by *in vitro* transwell cell migration and invasion assays. A, B) Bar graphs display the percentage of cells per field compared to the EVneg control that migrated through the insert membrane. C, D) Bar graphs display the percentage of cells per field compared to the EVneg control that invaded through matrigel and the insert membrane. For each assay, images of six fields were captured and counted using the Cell Counter option in ImageJ. Error bars represent the standard error of the mean. Three independent experiments were performed, each consisting of two technical replicates. Below each bar graph are representative images of migrated and invaded cells. * indicates significantly different from untreated (*p<0.05. **p<0.01. ***p<0.001).

3.2.5 Identification of transcriptional regulators of invasive genes upregulated upon BORIS expression

Next, we aimed to obtain a better understanding of how BORIS expression can lead to upregulation of invasive genes. BORIS was shown to be directly involvement in gene regulation by binding in the promoter region for the majority of the DEGs (55-70%) in the study by Pugacheva *et al.* [144], suggesting BORIS is a likely candidate for regulation of the invasive genes. We used iRegulon [253] to investigate transcriptional regulators of the BORIS-induced genes. First, we looked for enriched regulators among all BORIS upregulated genes. As expected we found the binding motif for CTCF/BORIS as most enriched (NES=3.61) among the upregulated DEGs (**Fig. 28A**). To confirm this prediction, we performed track discovery in iRegulon and found the ChIP-seq ENCODE track of CTCF from Hepg2 (NES=3.42) and H1 human embryonic stem cells (NES=3.23) as well as the track of BORIS from K562 cells (NES=3.16) amid the highest ranked tracks based on the normalized enrichment score (**Fig. 28B**). Combining predicted target genes of both the BORIS binding motif and ChIP-seq track indicates that 38% of the upregulated DEGs (574 out of 1490) contain a BORIS binding site within the promoter region (20kb centered around the TSS, **Fig. 28C**). Overall, this analysis suggests that BORIS is directly involved in the regulation of a subset of the identified upregulated DEGs.

A)

# Rank	Motif ID	AUC	NES	ClusterCode	Transcription factor	Motif
1	jaspar-PF0156.1	0.043	3.62	M1	CTCF/CTCFL	
2	elemento-GCAGCGC	0.041	3.26	M2		
3	jaspar-POL008.1	0.041	3.23	M3		
4	jaspar-PF0051.1	0.041	3.15	M4		
5	homer-M00115	0.040	3.07	M5	NANOG (Homeobox)	

B)

# Rank	Track ID	AUC	NES	ClusterCode	Transcription factor
1	wgEncodeSydhTfbsK562Stat2Ifna6hStdPk.narrowPeak.gz	0.044	3.50	T1	STAT2
2	wgEncodeHaibTfbsHepg2Ctcfsc5916V0416101PkRep1.broadPeak.gz	0.043	3.42	T2	CTCF
3	wgEncodeHaibTfbsH1hescCtcfsc5916V0416102PkRep2.broadPeak.gz	0.042	3.23	T2	CTCF
4	wgEncodeHaibTfbsK562Ctcflsc98982V0416101PkRep2.broadPeak.gz	0.041	3.17	T3	CTCFL
5	wgEncodeSydhTfbsNt2d1Suz12UcdPk.narrowPeak.gz	0.041	3.13	T4	SUZ12

C)

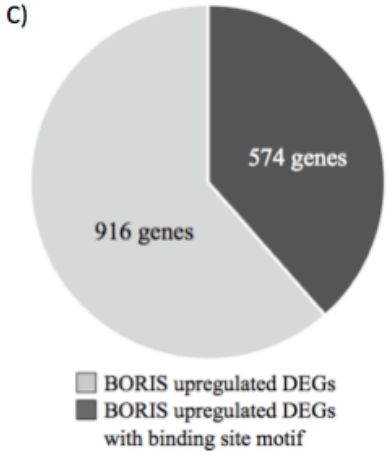


Figure 28. iRegulon analysis revealed BORIS as a potential direct regulator of BORIS-upregulated genes.

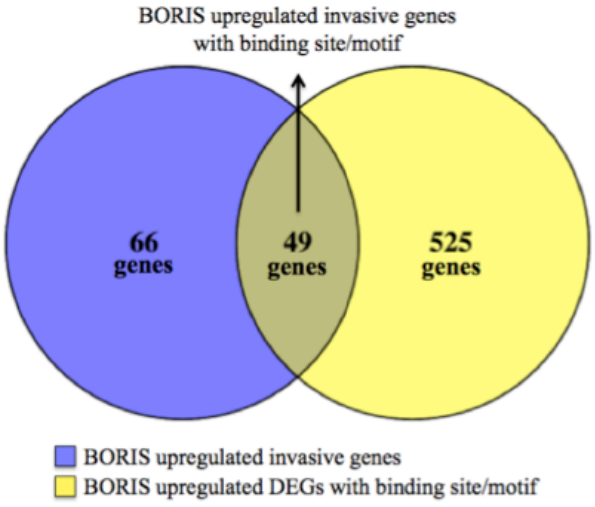
Upregulated DEGs as identified by RNA-seq were used as input for the Cytoscape plugin iRegulon [253] to assess enrichment of A) DNA binding motifs and B) transcription factor binding based on ENCODE ChIP-seq tracks 20 kilobases centered around the TSS of the DEGs. C) Pie chart demonstrating the overlap between upregulated DEGs with and without predicted BORIS TSS binding.

To address potential regulation of invasive genes by BORIS, we compared the predicted 574 BORIS target genes as identified above to the 115 invasive genes among BORIS DEGs (**Fig. 24D**) and found that 49 genes overlap (**Fig. 29A**). This indicates that BORIS binding within the gene promoter region could regulate some of the invasive genes, but is not solely responsible for the switch towards an invasive gene signature. In agreement, we did not identify enrichment of the BORIS motif (**Fig. 29B**) or ChIP-seq track (**Fig. 29C**) among master regulators of the upregulated invasive genes. The two most enriched binding motifs that we identified among the upregulated invasive genes were for the transcriptional regulators STAT/IRF and AP-1 (**Fig. 29B**). Members of these transcription factor families were also identified through track discovery (STAT1/2 and FOSL2), which in addition revealed NFIC, a member of the nuclear factor I (NF-I) family of transcription factors (**Fig. 29C**). These data are in line with previously identified regulators of the invasive state, which include AP1, NFIB, NFATC2 and TEADs [59, 61]. Interestingly, our RNA-seq data reveals that *FOSL2* (log_2FC=0.53, FDR=0.05), *FOSB* (log_2FC=1.82, FDR=0.01) and *JUNB* (log_2FC=0.67, FDR=3.3E-4), all members of the AP-1 transcription factor family, are upregulated with BORIS expression. Analysis of predicted AP-1 target genes within the invasive gene signature (112 genes) revealed an overlap of 28 genes with BORIS upregulated invasive genes (115 genes, **Fig. 30A**), suggesting BORIS indirectly contributes to the invasive gene signature via AP-1. Furthermore, we observed downregulation of *NFIB* (log_2FC=0.88, FDR=6.2E-6), which corresponds to its role as transcriptional repressor [61]. However, we observed very little overlap, 6 genes, between predicted invasive NFIB target genes (25 genes) and BORIS upregulated invasive genes (115 genes, **Fig. 30B**). Besides NFIB, Aibar *et al.* identified NFATC2 as another transcriptional repressor that regulates expression of invasive genes by repressing AP-1 and TEAD target genes [61]. While we did not identify this factor via motif or track discovery, we did observe significant downregulation of *NFATC2* (log_2FC=-1.36, FDR=5.1E-13) upon BORIS expression in the RNA-seq data. In their study, Aibar *et al.* established NFATC2 knockdown in the A375 melanoma cell line and performed RNA-seq, which revealed a significant enrichment of the invasive gene signature among upregulated genes [61]. When comparing the invasive genes upregulated upon NFATC2 knockdown (172 genes) with the invasive genes among BORIS upregulated DEGs (115 genes) we observed an overlap of 46 genes (**Fig. 30C**), suggesting that a subset of the invasive genes upregulated upon BORIS expression could be a result of NFATC2 downregulation. In addition,

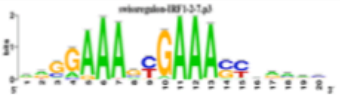
GSEA was performed on the list of upregulated genes upon NFATC2 knockdown, which demonstrated enrichment of the BORIS DEGs among this gene set (NES=3.07; FDR<0.0001; **Fig. 30D**). Another set of regulators that was identified in the study from Verfaillie *et al.* are the TEAD trasncription factors [59]. We did identify TEAD4 with ChIP-seq track discovery (NES=3.2) and found an overlap of 25 genes between the invasive genes within the TEAD target gene signature (104 genes) [59] and BORIS upregulated invasive genes (115 genes, **Fig. 30E**). However, no differential expression of any of the four TEAD family members was observed in our RNA-seq data, suggesting that aberrant TEAD expression is not involved in the regulation of BORIS upregulated invasive genes. Together, these data suggest that BORIS contributes to the switch from a proliferative to invasive gene signature both directly via binding within the promoter region and indirectly via downregulation of NFATC2 and upregulation of AP-1 family members.

Figure 29

A)



B)

# Rank	Motif ID	AUC	NES	ClusterCode	Transcription factor	Motif
1	swissregulon-IRF1-2-7.p3	0.093	4.14	M1	STAT/IRF	
2	yetfasco-1362	0.093	4.11	M2	AP-1	
3	encode-UW.Motif.0044	0.090	3.92	M3		
4	homer-M00095	0.089	3.90	M1	STAT/IRF	
5	yetfasco-1180	0.088	3.82	M4		

C)

# Rank	Track ID	AUC	NES	ClusterCode	Transcription factor
1	wgEncodeSydhTfbsK562Stat2Ifna6hStdPk.narrowPeak.gz	0.109	5.20	T1	STAT2
2	wgEncodeHaibTfbsSknshFosl2V0422111PkRep2.broadPeak.gz	0.100	4.70	T2	FOSL2
3	wgEncodeHaibTfbsSknshNficsc81335V0422111PkRep1.broadPeak.gz	0.094	4.33	T3	NFIC
4	wgEncodeSydhTfbsK562Stat1Ifna6hStdPk.narrowPeak.gz	0.093	4.31	T4	STAT1
5	wgEncodeHaibTfbsA549Gata3V0422111PkRep2.broadPeak.gz	0.091	4.18	T5	GATA3

Figure 29. iRegulon analysis identifies STAT and AP-1 transcription factors as regulators of BORIS-upregulated invasive genes.

A) Venn diagram [281] demonstrating the overlap between BORIS-upregulated invasive genes and BORIS-upregulated DEGs with predicted BORIS TSS binding as identified in Fig 7. B and C) BORIS-upregulated invasive genes revealed by RNA-seq were used as input for the Cytoscape plugin iRegulon [253] to assess enrichment of B) DNA binding motifs and C) transcription factor binding based on ENCODE ChIP-seq tracks 20 kilobases centered around the TSS of the BORIS-upregulated invasive genes.

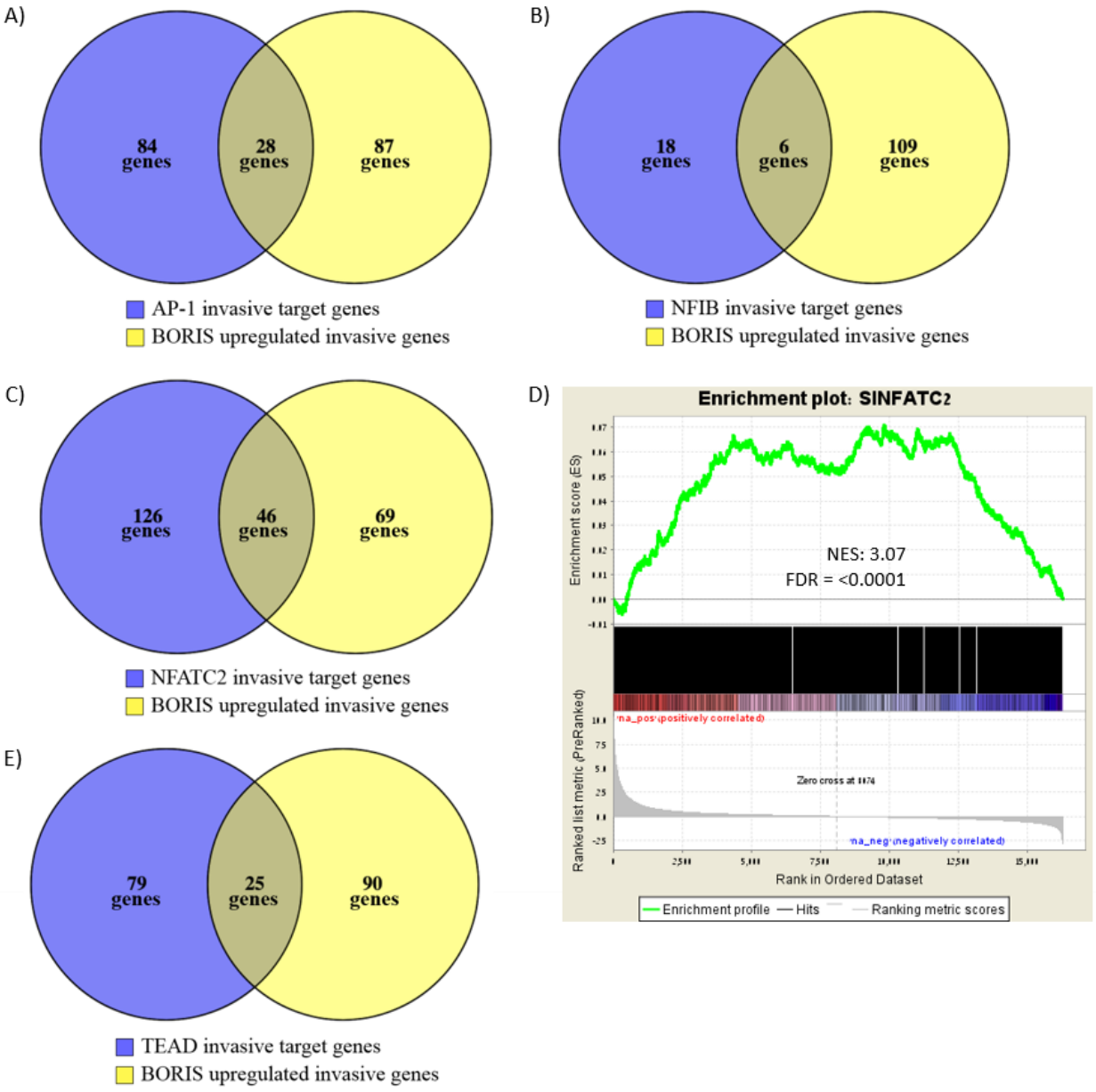


Figure 30. Target gene analysis points to AP-1 and NFATC2 as transcriptional regulators of BORIS-upregulated invasive genes.

Venn diagrams [281] demonstrating the overlap between BORIS-upregulated invasive genes and A) AP-1, B) NFIB, C) NFATC2, and E) TEAD target genes present in the Verfaillie invasive gene signature [59]. D) GSEA on the RNA-seq data (ranked according to log_2FC) for genes upregulated upon NFATC2 knockdown in A375 melanoma cells [61].

CHAPTER 3.3

–

Identification of BORIS Protein Partners

Sanne M. Janssen, Mounib Elchebly, Léon C.L. van Kempen and Alan Spatz

Background and objectives

While previous chapters have focused on the role of BORIS in melanoma cells by looking at BORIS-associated changes in gene expression, proteins that interact with BORIS influence BORIS' functions as well [167, 182, 184]. As mentioned previously, BORIS and CTCF share a similar DNA binding domain. However, their N- and C-terminal protein domains are highly divergent. This allows the interaction with different proteins that can lead to distinct functional outcomes for BORIS and CTCF. Currently, 25 proteins have been reported to interact with BORIS [144, 167, 181-184], including UBF as the single common interacting partner between CTCF and BORIS [183]. Considering the multi-functionality of BORIS, we anticipate that a number of BORIS-interacting proteins remain to be identified. Here we set out to further investigate BORIS protein partners, using more recent techniques. Our aims are to 1) reveal new BORIS interacting proteins, 2) compare BORIS and CTCF protein partners and 3) to identify cellular processes in which BORIS is involved.

Results

3.3.1 SILAC-based mass spectrometry reveals 79 new BORIS-only interacting proteins

To identify BORIS protein partners, we used SILAC-based mass spectrometry (MS). In short, HEK293T cells were cultured in the presence of either "light" or "heavy" amino acids for 5 passages to allow incorporation into synthesized proteins (**Fig. 31A**). To isolate BORIS protein partners, cells labeled with light amino acids were transfected with an expression plasmid containing HaloTag®-Empty Vector (HT-EV) control and cells with heavy labeled amino acids with HaloTag®-BORIS (HT-BORIS). HEK293T cells were used, since it is difficult to achieve high transfection efficiency in melanoma cell lines. Next, transfected cells were harvested and lysed, lysates were pooled in a 1:1 ratio based on protein concentration, and the HaloTag®-Pull Down (HT-PD) was performed to obtain proteins bound to the HT fusion proteins. MS was performed on the PD sample to identify proteins co-purifying with BORIS based on the ratio of light to heavy isotopes in the purified proteins (**Fig. 31B**). Expression of the HT-BORIS fusion protein was confirmed by IB on the total cell extract (TCE) and HT-PD sample (**Fig. 31C**). Interestingly, while we were unable to detect endogenous BORIS protein in the TCE, we observed several bands in the HT-PD sample. This suggests that BORIS is present at a low endogenous level that cannot be detected with the BORIS antibody and that BORIS can interact with itself. IB for CTCF demonstrated that CTCF is pulled down with HT-BORIS, which is in accordance with the known interaction between BORIS and CTCF ([144, 181] and **Fig. 31C**). Together this demonstrates that the HT-BORIS is expressed and can pull down CTCF and potential BORIS-interacting proteins.

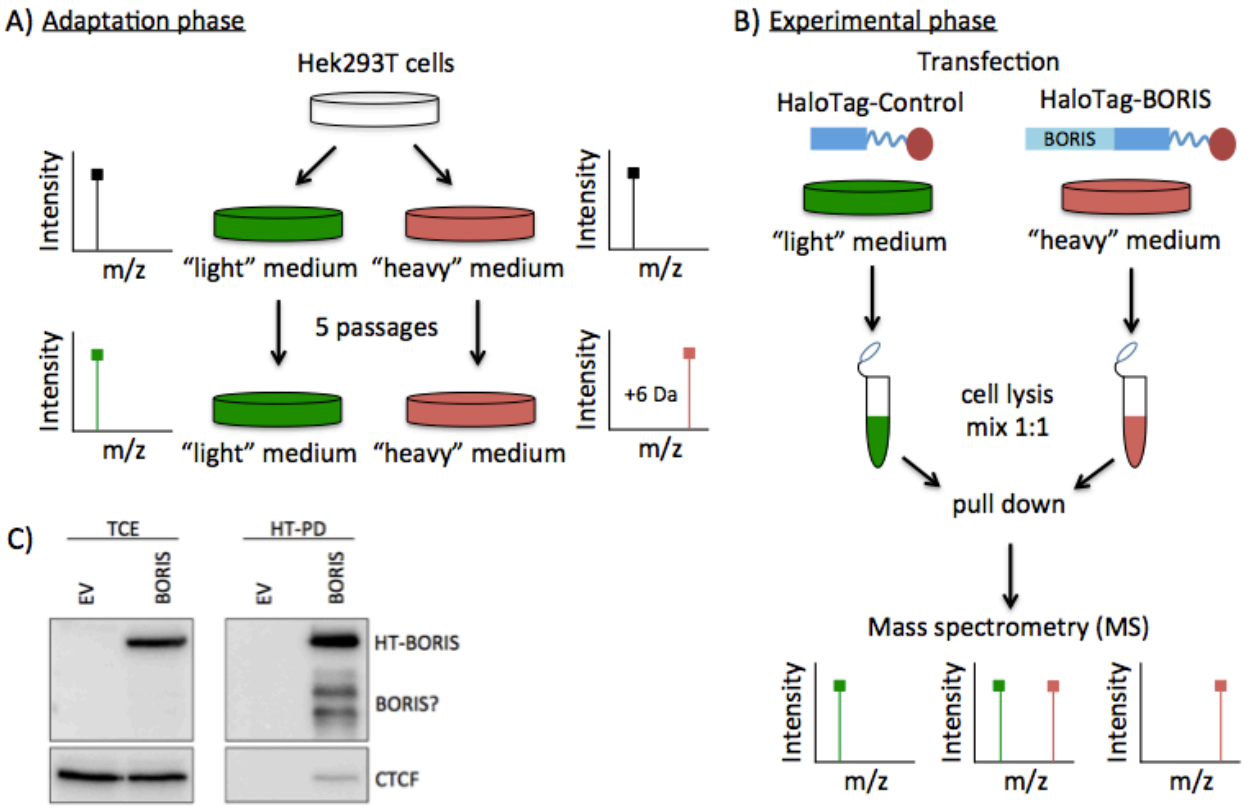


Figure 31. Identification of BORIS binding partners using Stable Isotope Labeling of Amino Acids in Cell Culture-based mass spectrometry followed by HaloTag® pull down. Schematic overview of the A) SILAC approach and B) HT-PD followed by MS. "Light" and "heavy" medium represents medium containing light and heavy labeled amino acids, respectively. The graphs depict the three possible MS results: a high intensity of light (left graph), both light and heavy (middle graph) or heavy (right graph) labeled amino acids. Only in the case of results as displayed in the right graph, meaning a higher intensity of heavy compared to light labeled amino acids, is the amino acid considered part of a potential BORIS-interacting protein. C) HT- EV and HT-BORIS pull-down followed by IB on TCE and HT-PD samples using anti-BORIS and anti-CTCF antibodies. Da: Dalton, m/z: ion mass to charge ratio.

The MS results revealed 83 BORIS-interacting proteins, excluding BORIS itself, by applying the following criteria in the Scaffold Viewer software: a minimum detection of 2 peptides, 99% protein threshold, 99% peptide threshold, $log_2FC>0.6$ and $p<0.05$ (**Appendix IV**). Proteins were ranked based on the log_2FC and *p* value, and an overview of the top 10 BORIS-interacting proteins can be found in **table III**. Known interacting proteins of BORIS that were identified using these criteria included HCFC1 [167] and UBF [183]. By reducing the stringency of the analysis, we were also able to detect the known partners BAG6 (BAT3), POGZ, CHD8, and ELF2 [167]. In addition, ASH2L was detected, which is known to interact with the previously identified BORIS protein partners HCFC1 ([284] and **Table III**) and SET1A [167, 285]. The identification of these known interacting partners indicates that HT-BORIS is a functional fusion protein. To our surprise, we did not identify CTCF among the BORIS-interacting proteins, while CTCF is known to interact with BORIS ([144, 181] and **Fig. 31C**). Next, we used MIST [229] to identify protein complexes among BORIS-interacting proteins, which revealed a large number of complexes that contain BORIS protein partners (**Appendix V**). Mainly, we observed BORIS-interacting partners among protein complexes involved in translation, splicing, DNA repair, and chromatin remodeling (**Fig. 32**). To determine if there is an overlap between BORIS and CTCF-interacting proteins, we first identified CTCF interacting partners via a literature search, resulting in a list of 32 proteins (**Appendix VI**). Then we compared them to the identified BORIS interacting partners, revealing only 3 overlapping protein partners: PARP, UBF and SMC1 (**Fig. 33A**). We validated the interaction of BORIS and CTCF with UBF and PARP using the HT-PD system (**Fig. 33B and C**). Overall, the SILAC-based MS approach combined with the HT-PD system, revealed 79 new BORIS-only interacting proteins.

Table III: Top 10 BORIS-interacting proteins

Protein ID	Ranking score	Log2FC	p-value	Coverage (%)
CTCFL	49.16	3.7	1.00E-04	37
SUPT16H	34.55	2.6	1.00E-04	31
HCFC1	32.81	2.67	2.00E-04	11
SND1	25.11	1.89	1.00E-04	22
SMARCA5	23.12	1.74	1.00E-04	21
SSRP1	20.06	1.51	1.00E-04	33
HSPD1	18.20	1.37	1.00E-04	55
IPO7	17.74	1.78	1.00E-03	14
PRKDC	16.74	1.26	1.00E-04	14
HSP90AA1	16.74	1.26	1.00E-04	30
RPL7A	16.61	1.25	1.00E-04	35

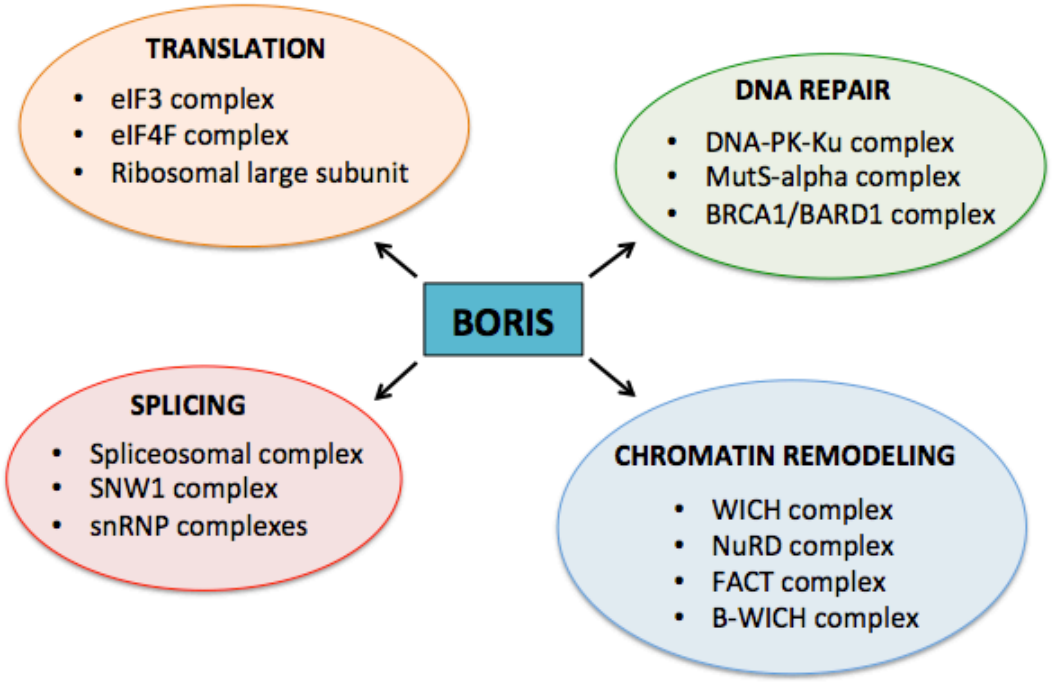


Figure 32. Identification of protein complexes that contain BORIS-interacting proteins.
Schematic representation of various protein complexes identified among BORIS protein partners using MIST [229]. Protein complexes were grouped according to known involvement in a cellular process.

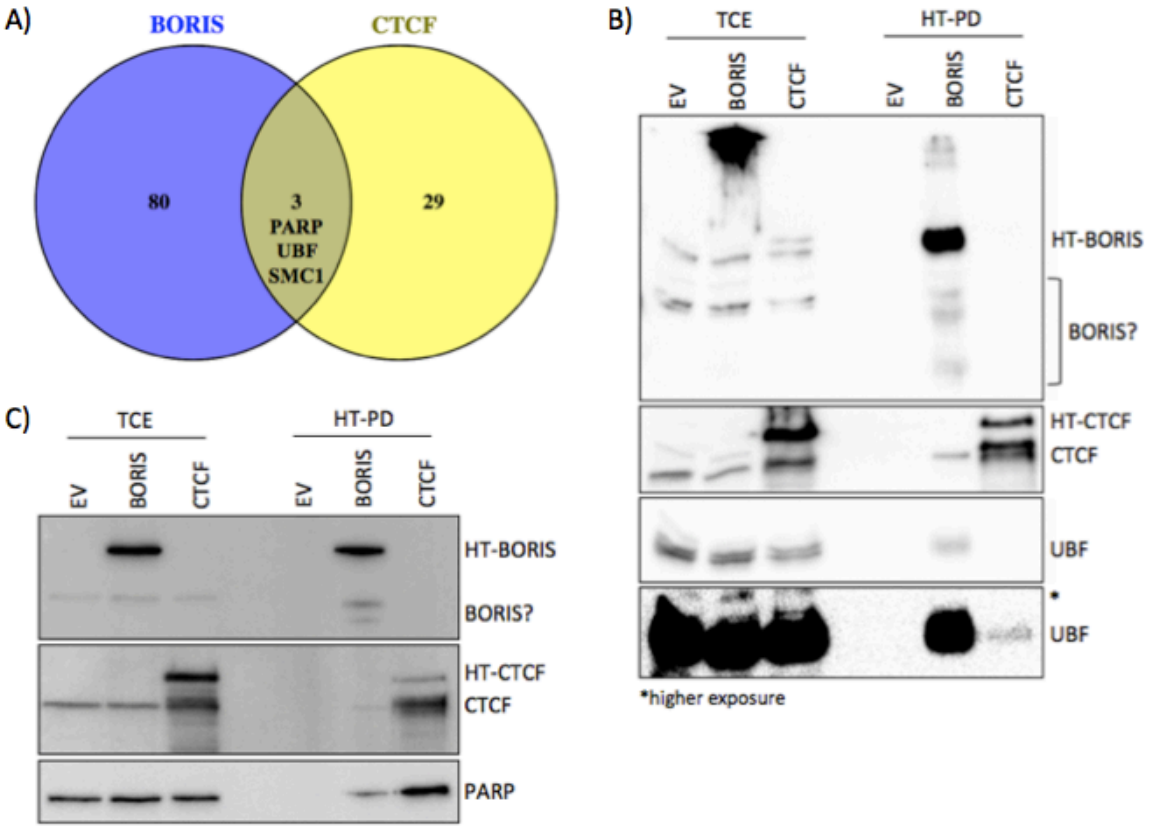


Figure 33. PARP and UBF are common protein partners of BORIS and CTCF.

A) Venn diagram [281] demonstrating the overlap between BORIS and CTCF-interacting proteins. B and C) HT-EV, HT-BORIS and HT-CTCF PD followed by IB on TCE and HT-PD samples using anti-BORIS, anti-CTCF, B) anti-UBF, and C) anti-PARP antibodies.

3.3.2 Overrepresentation analysis indicates a role for BORIS in translation, DNA repair, and mRNA and rRNA expression

Next, we wanted to identify biological processes and pathways that are overrepresented among our list of BORIS-interacting partners as this will allow us to gain insight into cellular processes in which BORIS may be involved. Overrepresentation analysis for biological processes was performed on the 83 identified BORIS protein partners in the PANTHER classification system [243]. This analysis revealed 27 biological processes significantly overrepresented (FDR<0.05) among the input list of proteins (**Appendix VII**). We used REViGO [244] to remove highly redundant processes based on their GO terms (**Appendix VII**). The output was manually annotated to obtain a comprehensive, clustered overview of the main overrepresented biological processes (**Fig. 34A**). Based on the clustering, the top 5 overrepresented biological processes are primary metabolism, mRNA processing, translation, DNA repair and rRNA metabolism (**Table IV**). In addition, we performed overrepresentation analysis for Reactome pathways in the CPDB [245], which revealed 41 significantly overrepresented pathways ($q< 0.01$) among the list of BORIS protein partners (**Appendix VII**). Next, we made use of the concept overlap graph function within CPDB to cluster, manually annotate and visualize the results (**Fig. 34B**). Based on the clustering, the top 5 overrepresented pathways are metabolism of RNA, translation, rRNA processing, DNA repair and infectious disease (**Table V**). Together, the overrepresentation analyses indicate a role for BORIS in 1) translation, 2) DNA repair, and 3) mRNA and rRNA processing.

Table IV: Top 5 overrepresented biological processes

Biological Process	ID	FDR
Primary metabolism	GO:0044238	2.75E-06
mRNA processing	GO:0006397	5.69E-06
Translation	GO:0006412	6.93E-05
DNA repair	GO:0006281	1.72E-03
rRNA metabolism	GO:0016072	2.32E-03

Table V: Top 5 overrepresented pathways

Pathway	ID	q-value
Metabolism of RNA	R-HSA-8953854	1.41E-14
Translation	R-HSA-72766	7.66E-11
rRNA processing	R-HSA-72312	8.94E-05
DNA Repair	R-HSA-73894	2.12E-04
Infectious disease	R-HSA-5663205	1.24E-03

Figure 34

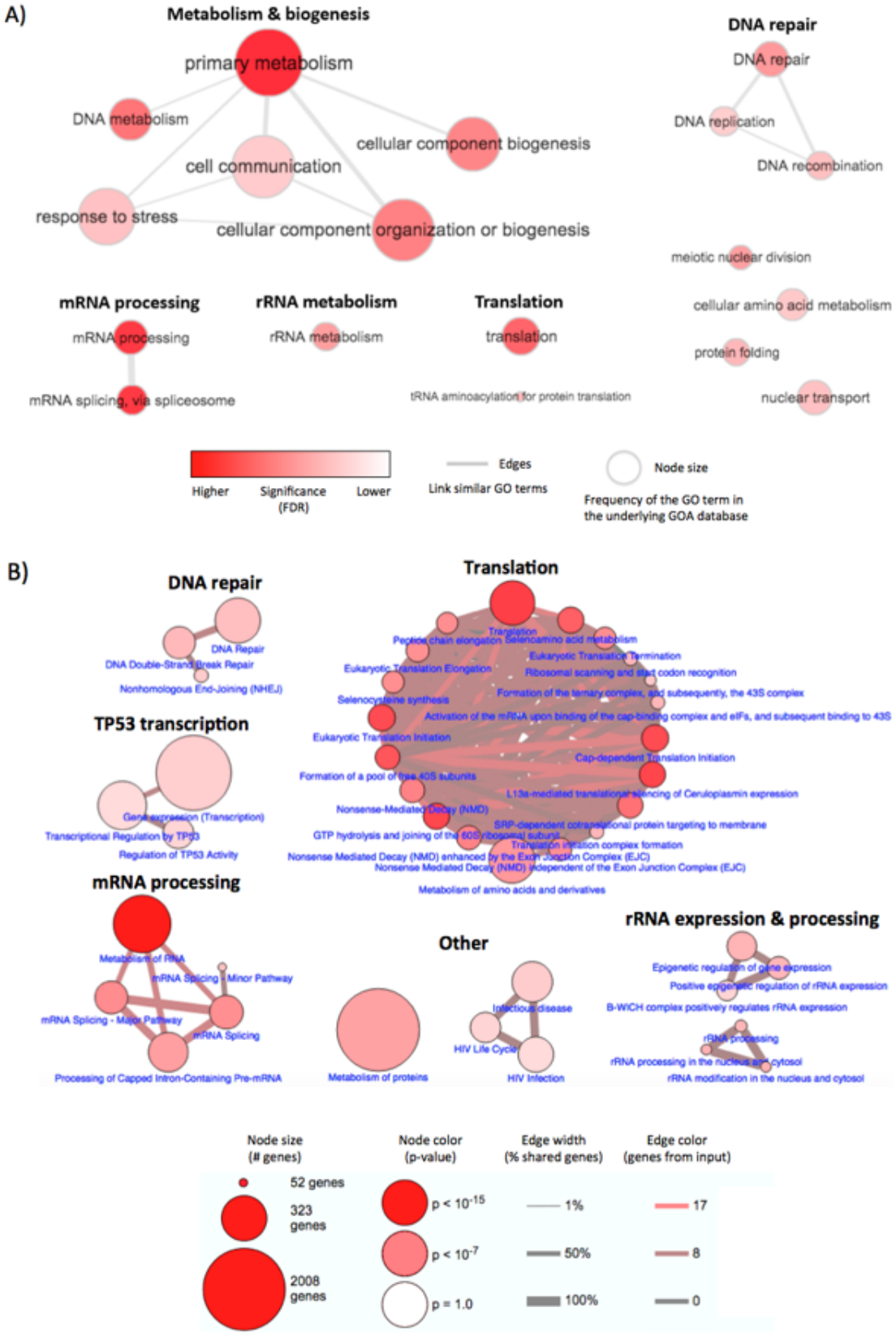

A)
Metabolism & biogenesis
primary metabolism
DNA metabolism
cellular component biogenesis
cell communication
response to stress
cellular component organization or biogenesis
DNA repair
DNA repair
DNA replication
DNA recombination
meiotic nuclear division
cellular amino acid metabolism
protein folding
nuclear transport
mRNA processing
mRNA processing
mRNA splicing, via spliceosome
rRNA metabolism
rRNA metabolism
Translation
translation
tRNA aminoacylation for protein translation
Higher
Significance (FDR)
Lower
Edges
Link similar GO terms
Node size
Frequency of the GO term in the underlying GOA database
B)
DNA repair
DNA Repair
DNA Double-Strand Break Repair
Nonhomologous End-Joining (NHEJ)
TP53 transcription
Gene expression (Transcription)
Transcriptional Regulation by TP53
Regulation of TP53 Activity
mRNA processing
Metabolism of RNA
mRNA Splicing - Minor Pathway
mRNA Splicing - Major Pathway
mRNA Splicing
Processing of Capped Intron-Containing Pre-mRNA
Translation
Translation
Peptide chain elongation
Selenocysteine synthesis
Eukaryotic Translation Elongation
Eukaryotic Translation Termination
Ribosomal scanning and start codon recognition
Formation of the ternary complex, and subsequently, the 43S complex
Activation of the mRNA upon binding of the cap-binding complex and eIFs, and subsequent binding to 43S
Eukaryotic Translation Initiation
Cap-dependent Translation Initiation
Formation of a pool of free 40S subunits
L13a-mediated translational silencing of Ceruloplasmin expression
Nonsense-Mediated Decay (NMD)
SRP-dependent cotranslational protein targeting to membrane
GTP hydrolysis and joining of the 60S ribosomal subunit
Translation initiation complex formation
Nonsense Mediated Decay (NMD) enhanced by the Exon Junction Complex (EJC)
Nonsense Mediated Decay (NMD) independent of the Exon Junction Complex (EJC)
Metabolism of amino acids and derivatives
Other
Infectious disease
HIV Life Cycle
HIV Infection
Metabolism of proteins
rRNA expression & processing
Epigenetic regulation of gene expression
Positive epigenetic regulation of rRNA expression
B-WICH complex positively regulates rRNA expression
rRNA processing
rRNA processing in the nucleus and cytosol
rRNA modification in the nucleus and cytosol
Node size (# genes)
52 genes
323 genes
2008 genes
Node color (p-value)
p < 10-15
p < 10-7
p = 1.0
Edge width (% shared genes)
1%
50%
100%
Edge color (genes from input)
17
8
0

Figure 34. Identification of cellular processes and pathways overrepresented among BORIS-interacting proteins.

Network maps of clustered A) biological processes (FDR<0.05) and B) reactome pathways (q< 0.01) revealed by overrepresentation analysis in the PANTHER classification system [243] and the CPDB [245], respectively. Clusters were manually annotated for enhanced comprehension. A) For biological processes, REViGO [244] was used to remove highly redundant processes and the interactive graph function for visualization of the results. B) The overlap graph function within the CPDB was used to generate the network map of overrepresented pathways.

3.3.3 BORIS does not alter rRNA expression in the MM057 melanoma cell line

Our overrepresentation analysis suggests a role for BORIS in rRNA metabolism (**Table IV and V**) and we, as well as others [183], have demonstrated an interaction between BORIS and the rDNA transcription factor UBF. In addition, among the list of BORIS-interacting proteins (**Appendix IV**) we observed SMARCA5/SNF2h/ISWI, BAZ1B/WSTF, MYBBP1A and SF3B1/SAP155 that belong to the B-WICH chromatin-remodeling complex [286], which is involved in the regulation of rRNA expression [287, 288]. Furthermore, BORIS was shown to bind rDNA and alter rRNA expression [157, 183]. Here, our aim was to investigate the role of BORIS in rRNA expression in melanoma cells.

First, we assessed if BORIS can localize to the nucleolus in melanoma cells. To this end, MM057 cells were transfected with the HK-BORIS plasmid (described in section 3.1.2), which allowed us to indirectly visualize BORIS via the fluorescent protein Katushka. Expression of HK-BORIS in the MM057 cells was validated by IB (**Fig. 35A**). IF revealed specific BORIS localization in DAPI-low areas of the nucleus that are known to correspond to the nucleoli (**Fig. 35B**). In addition, IF was performed for both HK-BORIS and UBF, revealing co-localization of the two proteins (**Fig. 35C**). This indicates that, in accordance with previous studies, BORIS localizes to the nucleoli in our melanoma cell line model. It remains unknown if BORIS expression in the nucleoli can alter UBF protein expression. Therefore, we performed IB for UBF upon ectopic expression of hBORIS-6xH in the MM057 cell line and observed no change in total UBF protein level compared to controls (**Fig. 35D**). In contrast to previous data [157], we observed no effect of ectopic hBORIS-6xH expression on pre-rRNA expression assessed by qPCR using different primer pairs (**Fig. 35E**). These data suggest that, while BORIS localizes to the nucleoli, BORIS expression does not alter UBF or rRNA expression in MM057 cells.

Figure 35

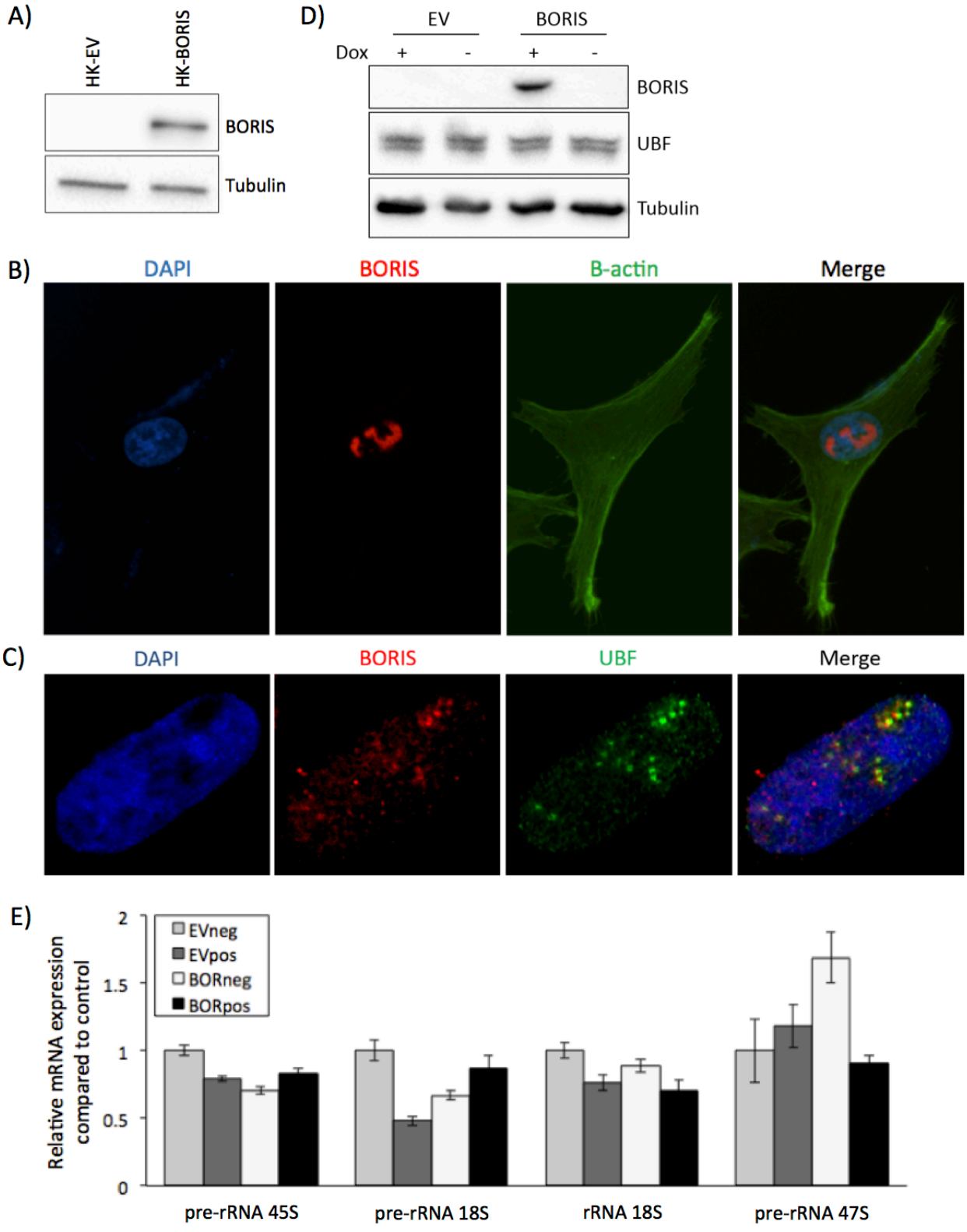

A)
HK-EV
HK-BORIS
BORIS
Tubulin
D)
EV
BORIS
Dox
+
-
+
-
BORIS
UBF
Tubulin
B)
DAPI
BORIS
B-actin
Merge
C)
DAPI
BORIS
UBF
Merge
E)
Relative mRNA expression compared to control
2
1.5
1
0.5
0
EVneg
EVpos
BORneg
BORpos
pre-rRNA 45S
pre-rRNA 18S
rRNA 18S
pre-rRNA 47S

Figure 35. BORIS and UBF co-localize in the nucleoli where BORIS expression does not alter rRNA expression.

A-C) MM057 cells were infected with the HK-EV or HK-BORIS vector. A) Expression of HK-BORIS protein was validated by IB using anti-BORIS antibody. IF for DAPI (blue), HK-BORIS (red), B) β-actin (green), or C) UBF (green). Images were acquired with a Nikon B) fluorescent microscope and C) C2+ confocal microscope. D and E) MM057 cells were genetically engineered to stably express EV-6xH or BORIS-6xH upon induction with dox. Expression of BORIS protein was induced for 5 days with 50ng/ml dox. D) BORIS and UBF protein level were determined by IB using anti-BORIS and anti-UBF antibodies. E) qPCR was performed in triplicate for the indicated rRNAs. Expression was normalized to *HPRT1* and *TBP*. Data was analyzed with the Bio-Rad CFX Manager™ Software. Error bars represent the standard error of mean. For IBs, anti-Tubulin was used as a loading control.

3.3.4 BORIS expression increases DNA damage and activates the DNA damage response in the MM057 melanoma cell line

The overrepresentation analysis on BORIS interacting proteins also revealed a potential role for BORIS in DNA repair. This was of special interest since recent studies have reported a role for CTCF in the DNA damage response (DDR) [289-291]. The DDR is a collective term for all cellular events and mechanisms involved in sensing, signaling and repair of damaged DNA. This response is mediated by a set of proteins known as DDR proteins, various DNA repair mechanisms as well as cell cycle checkpoints. The DDR is crucial for preventing accumulation of DNA damage, which leads to genomic instability. While CTCF is important for the maintenance of genome stability [291], ectopic BORIS expression was linked to genomic instability [157]. In addition, the ectopic expression of BORIS induced an accumulation of cells in the S-phase [157]. This is likely a result from delayed replication and not due to increased proliferation, since BORIS expression has been shown to reduce cell number by us (section 3.2.1) as well as others [157]. DNA replication can be delayed by activation of the replication checkpoint as a response to stress, or by activation of the replication-independent intra-S-phase checkpoint as a result of DNA damage [292]. This supports the idea that BORIS could play a role in the DDR and that this is related to activation of the intra-S-phase DNA damage checkpoint.

To further investigate if BORIS plays a role in the DDR, we first aimed to determine if BORIS expression in our cell model leads to an accumulation of cells in the S-phase. hBORIS-6xH expression was induced in the MM057 cell line using dox and the cell cycle profile was assessed by flow cytometry (**Fig. 36A**). We observed a decreased percentage of cells in the G1-phase, an increased number of cell in the S-phase and no significant difference in the G2/M-phase of the cell cycle compared to the controls (**Fig. 36B**). Interestingly, when looking at the proteins involved in the intra-S-phase checkpoint [292], multiple overlap with BORIS-interacting proteins as identified via our MS approach. These include all members of the MRN complex (Mre11, Rad50 and Nbs1), SMC1 and MDC1 (**Table VI**). Not all these proteins were considered significant interacting proteins according to the Scaffold software analysis, suggesting they may be indirect binding partners of BORIS. These data, together with our observation that BORIS does not increase cell proliferation (section 3.2.1, **Fig. 16A**), demonstrate a BORIS-induced

delay in S-phase and suggests that expression of BORIS leads to activation of the intra-S-phase checkpoint in the MM057 cell line.

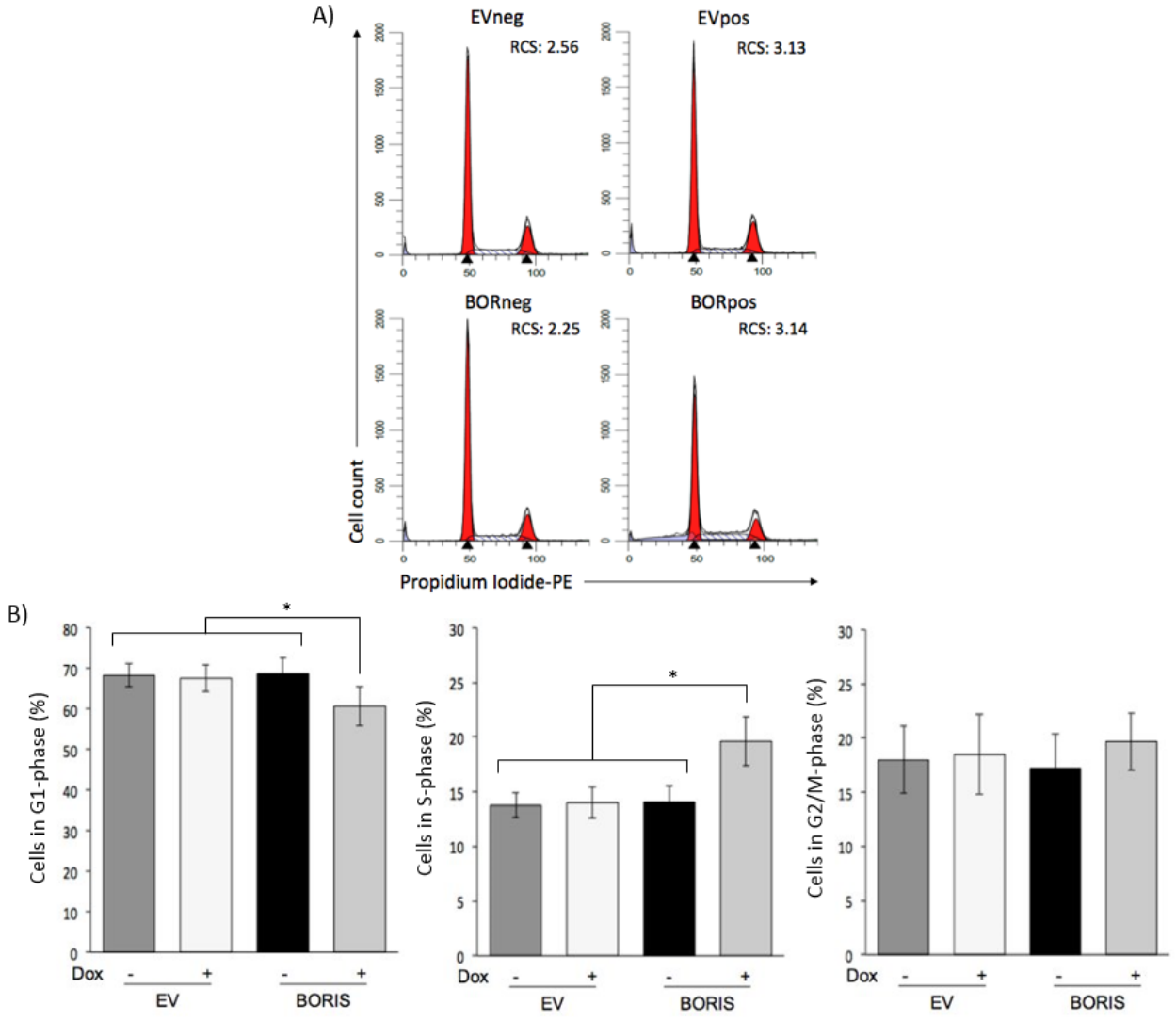


Figure 36. BORIS expression results in an accumulation of cells in the S-phase.

MM057 cells were genetically engineered to stably express EV-6xH or hBORIS-6xH upon induction with dox. Expression of BORIS was induced with 50ng/ml dox for 5 days and cell cycle was assessed by flow cytometry for PI incorporation. A) Representative image after analysis with ModFit LT™ (RCS: Reduced Chi-Square). B) Bar graph representing the percentage of live cells in the G1-, S- and G2/M-phase. * indicates significantly different (p<0.05).

Table VI: BORIS-interacting proteins involved in the intra-S-phase checkpoint

Protein	Exlusive peptides (#)	Coverage (%)
RAD50 (MRN)	8	9
SMC1A	12	13
MDC1*	4	3
MRE11 (MRN)*	2	5
NBS1 (MRN)*	2	5

*Not significant according to Scaffold software analysis

Since activation of the intra-S-phase checkpoint occurs upon DNA damage, we wanted to assess if BORIS expression changes the level of DNA damage in the MM057 cells. One marker that is commonly used to assess DNA damage is phosphorylation of histone H2AX at Serine 139, also known as γH2AX [293]. Compared to the controls, dox-induced hBORIS-6xH expression resulted in high γH2AX protein expression (**Fig. 37A**). Upon DNA damage in the form of double strand breaks (DSBs), γH2AX appears as nuclear foci at the site of damage [294]. Our preliminary results show that the number and intensity of nuclear γH2AX foci is higher in cells with BORIS expression compared to control cells (**Fig. 37B**). These data are in accordance with the idea that BORIS-induced DNA damage results in activation of the S-phase checkpoint. Activation of DNA damage checkpoints is mediated through a combined action of two main kinases, ataxia telangiectasia-mutated (ATM) and Ataxia telangiectasia and Rad3 related. In the case of the S-phase checkpoint, ATM is the main DNA damage-responsive kinase that can activate two downstream DDR pathways, either via phosphorylation of NBS1 or Checkpoint kinase 2 [292, 295-299]. To determine if BORIS expression leads to downstream ATM signaling, expression of phosphorylated Chk2 was assessed by IB and revealed increased phosphorylation in the presence of BORIS (**Fig. 37C**). Together, these results indicate that BORIS expression results in DNA damage, activation of the intra-S-phase checkpoint and downstream DDR signaling.

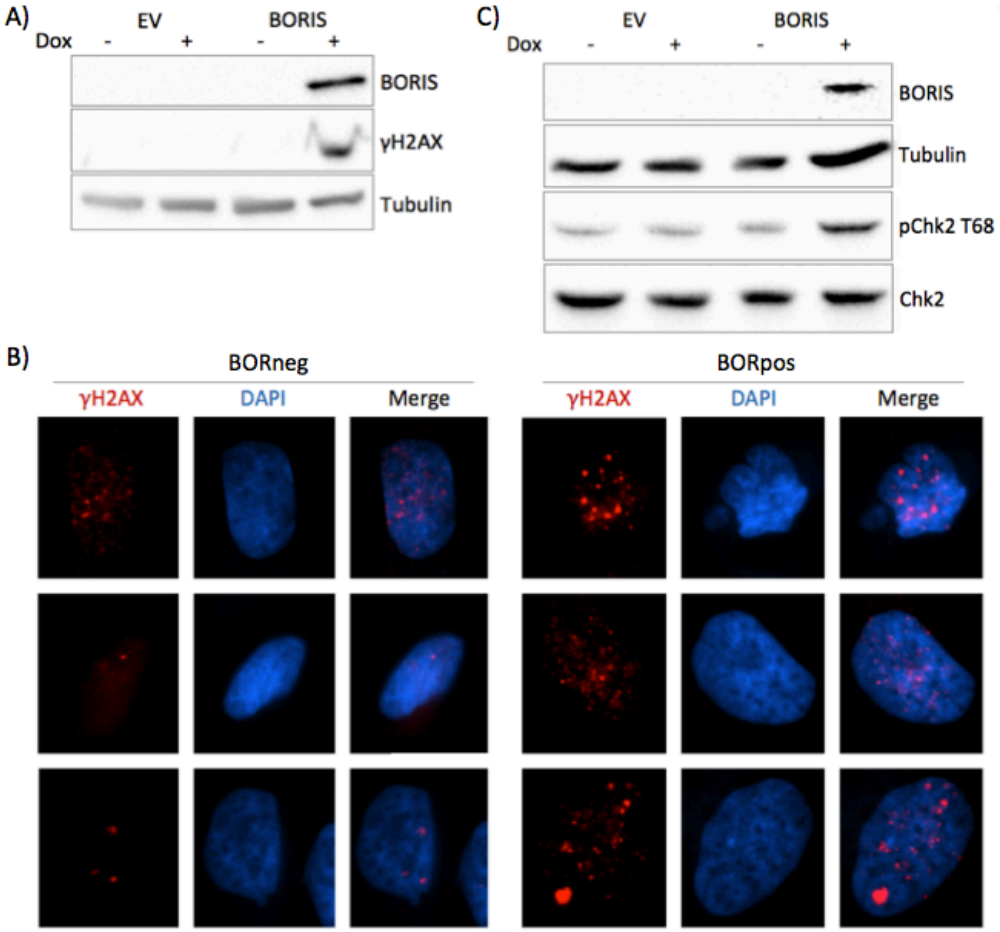


Fig 37. BORIS expression results in DNA damage and activation of the DDR.

MM057 melanoma cells were genetically engineered to stably express EV-6xH or hBORIS-6xH upon induction with dox. BORIS expression was induced with 50ng/ml dox for 5 days. A) BORIS and γH2AX protein level were determined by IB using anti-BORIS and anti-γH2AX antibodies. B) IF staining for DAPI (blue) and γH2AX (red). Images were acquired on a Nikon fluorescent microscope. C) IB using anti-BORIS, anti-phosphorylated Thr68 Chk2, and anti-Chk2 (total) antibodies. For IBs, anti-Tubulin was used as a loading control.

Knockdown of CTCF has been shown to increase sensitivity of cells to DNA damage-inducing agents [290]. We observed increased DNA damage upon BORIS expression, suggesting that BORIS in this context has an opposing role to CTCF. To further investigate this, the effect of BORIS expression on the sensitivity of cells to UVC damage was determined using flow cytometry analysis. Our preliminary results showed that, while control cells displayed ~10-20% cell death 24 hours after UVC irradiation, this percentage was increased to 40% in MM057 cells with ectopic hBORIS-6xH expression after correction for BORIS associated, UVC-independent cell death (**Fig. 38A**). In addition to UVC-induced DNA damage, we assessed the effect of PARP inhibition on MM057 cells with BORIS expression. Inhibition of PARP interferes with DNA repair of single strand DNA breaks, allowing these lesions to progress to DSBs. Therefore, cells lacking the ability to properly repair DSBs will 'overdose' on DNA damage in the presence of PARP inhibition, resulting in cell death [300]. Our preliminary results show a reduced number of metabolically active cells as determined by MTT assay (**Fig. 38B**) and an increase in cell death by flow cytometry (**Fig. 38C and D**) upon treatment with the PARP inhibitor Olaparib in BORIS expressing cells. This suggests that BORIS expression leads to a decreased ability of the cell to repair damaged DNA. To investigate the effect of BORIS expression on DNA repair, we induced DNA damage using hydrogen peroxide in the MM057 cells with hBORIS-6xH expression and followed γH2AX expression over time. While expression of γH2AX became undetectable after 1 hour in control cells, γH2AX was still detectable in BORIS expressing cells after 2 hours (**Fig. 38E**), suggesting a dealy in DNA repair. In sum, these preliminary data suggest that BORIS expression increases DNA damage sensitivity by negatively affecting repair kinetics and further support a role for BORIS in the DDR.

Figure 38

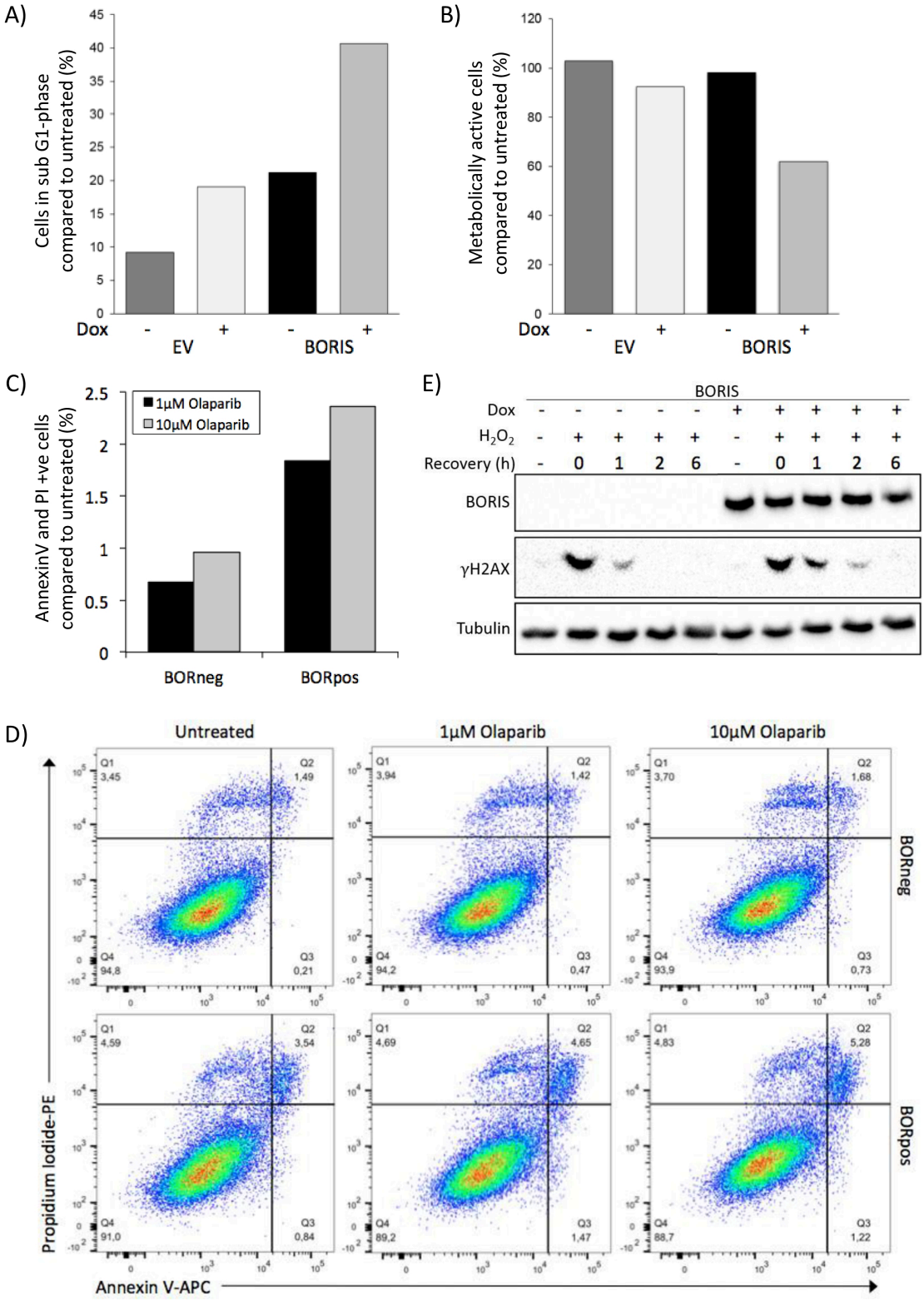

A)
Cells in sub G1-phase compared to untreated (%)
45
40
35
30
25
20
15
10
5
0
Dox - + - +
EV
BORIS
B)
Metabolically active cells compared to untreated (%)
120
100
80
60
40
20
0
Dox - + - +
EV
BORIS
C)
AnnexinV and PI +ve cells compared to untreated (%)
2.5
2
1.5
1
0.5
0
1µM Olaparib
10µM Olaparib
BORneg
BORpos
E)
BORIS
Dox - - - - - + + + + +
H2O2 - + + + + - + + + +
Recovery (h) - 0 1 2 6 - 0 1 2 6
BORIS
γH2AX
Tubulin
D)
Untreated
1µM Olaparib
10µM Olaparib
Q1 3,45
Q2 1,49
Q4 94,8
Q3 0,21
Q1 3,94
Q2 1,42
Q4 94,2
Q3 0,47
Q1 3,70
Q2 1,68
Q4 93,9
Q3 0,73
Q1 4,59
Q2 3,54
Q4 91,0
Q3 0,84
Q1 4,69
Q2 4,65
Q4 89,2
Q3 1,47
Q1 4,83
Q2 5,28
Q4 88,7
Q3 1,22
BORneg
BORpos
Propidium Iodide-PE
Annexin V-APC

Fig 38. BORIS expression increases DNA damage sensitivity and alters repair kinetics.
MM057 melanoma cells were genetically engineered to stably express EV-6xH or hBORIS-6xH upon induction with dox. A) BORIS expression was induced with 50ng/ml dox for 5 days. Cells were irradiated with 60 J/m^2 UVC and cell cycle was assessed after 24 hours by flow cytometry for PI incorporation. Bar graph represents the percentage of cells in the sub G1-phase compared to untreated cells. B) BORIS expression was induced with 50ng/ml dox for 3 days followed by treatment with a combination of 50 ng/ml dox and 10μM Olaparib for 3 days. The percentage of metabolically active cells was determined using the MTT assay with absorbance measured at 570nm. C) BORIS expression was induced with 50ng/ml dox for 3 days followed by treatment with a combination of 50ng/ml dox and 1μM or 10μM Olaparib for 3 days. Bar graph represents the percentage Annexin V and PI positive cells compared to untreated cells as analyzed by FlowJo®. D) Representative image of the percentage of apoptotic and necrotic cells as determined by flow cytometry analysis of Annexin V and PI staining. E) Expression of BORIS was induced for 2 days with 50ng/ml dox followed by DNA damage induction with 200μM H_2O_2 for 1.5 hours. Cells were allowed to recover for the indicated time points and IB was performed using anti-BORIS and anti-γH2AX antibodies. Anti-Tubulin was used as a loading control.

3.3.5 BORIS interacts with the histone chaperone FACT

Among the top 10 BORIS-interacting proteins we observed both Structure Specific Recognition Protein 1 (SSRP1) and SPT16, which together form the protein complex Facilitates Chromatin Transcription (FACT) [301]. FACT is a histone chaperone, which is a class of proteins that loosen higher order nucleosome structures and reorganize individual nucleosomes to allow access to the DNA for cellular machinery. Different studies have shown that FACT can bind H2A-H2B as well as H3-H4, which is a proposed mechanism on how FACT destabilizes nucleosomes and promotes histone dimer displacement [302-306]. FACT was shown to be important for retaining existing histones and reestablishing chromatin during transcription as well as replication [302, 307-309]. In addition to transcription and replication, FACT also plays a role in DNA damage repair [310, 311]. Recently, Piquet *et al.* revealed that FACT is required for new H2AX incorporation at the site of DNA damage [312]. This is a key event, since phosphorylation of H2AX is crucial for the recruitment of DDR proteins to the site of DNA damage. Interestingly, PARP-mediated ADP-ribosylation of the FACT subunit SPT16 was shown to inhibit the exchange reaction between H2A and H2AX [313]. Given that our data demonstrates a protein-protein interaction between BORIS and PARP and strongly indicates a role for BORIS in the DDR, we propose a possible interplay between BORIS and FACT in the DDR. As a first step to address this possibility, we set out to confirm the interaction between BORIS and the FACT complex.

First, the ability of BORIS and CTCF to interact with the SSRP1 subunit of the FACT complex was assessed by HT-PD. We were able to pull-down SSRP1 with HT-BORIS, but not with HT-CTCF (**Fig. 39A**). Next, we confirmed the interaction between BORIS and both the SSRP1 and SPT16 subunit using the HT-PD system and show that the interactions are maintained in the presence of Benzonase® nuclease (**Fig. 39B**), indicating that the interaction of BORIS with the FACT complex is nucleic acid independent. In addition, we performed co-IP in the MM057 melanoma cell line for BORIS and the FACT subunits. Due to the lack of a good BORIS antibody, we established stable integration of an expression vector encoding dox-inducible BORIS fused to a triple FLAG®-tag (hBORIS-3xF). The expression of hBORIS-3xF in the presence of dox was validated by IB (**Fig. 39C**). The results of the co-IP were in agreement with

the HT-PD experiments, showing that BORIS can be co-IP'ed with SSRP1 (**Fig. 39D**). Together, these data demonstrate that BORIS interacts with the FACT complex.

Figure 39

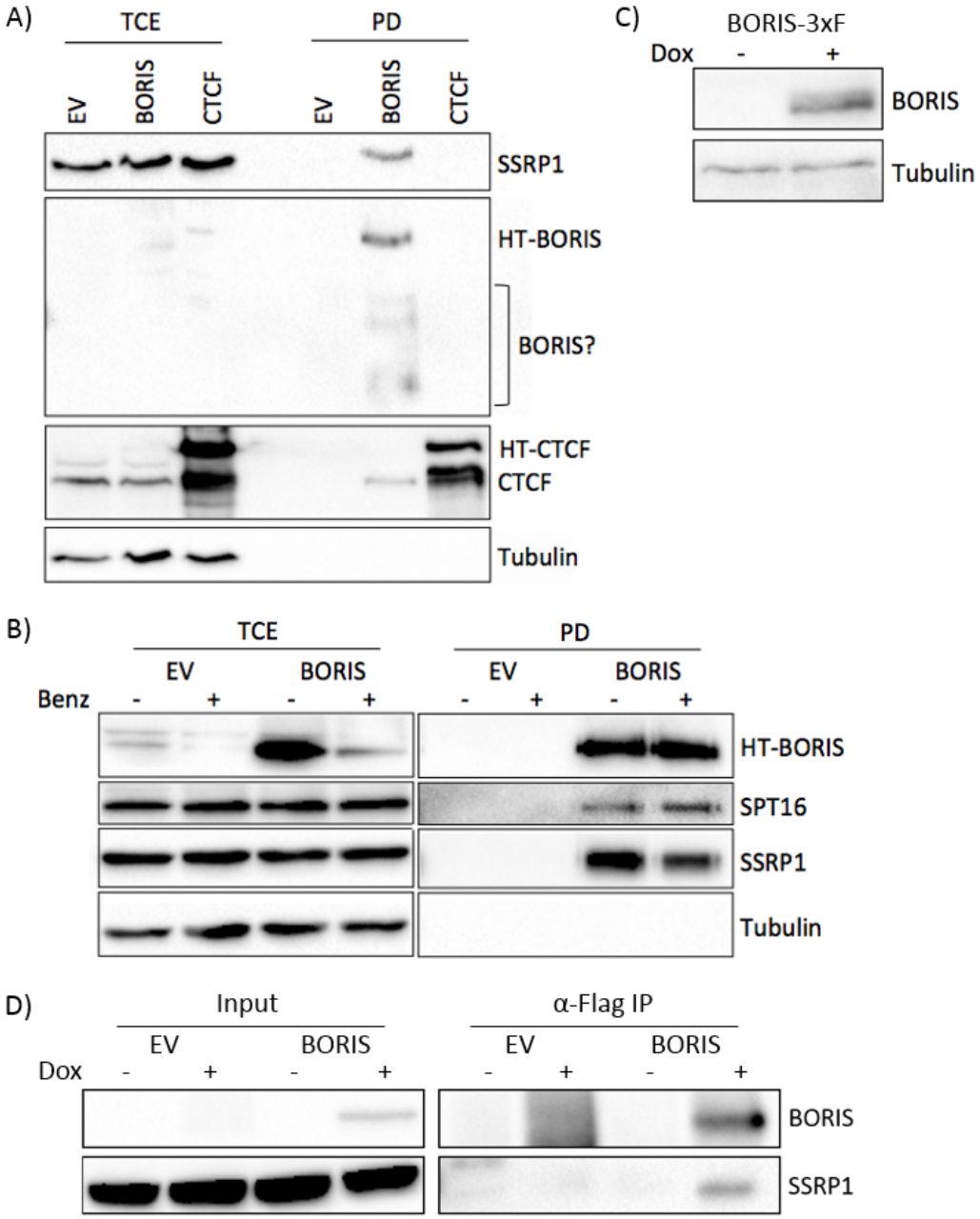


Figure 39. A nucleic acid-independent interaction between BORIS and the FACT subunits SPT16 and SSRP1.

A) HT-EV, HT-BORIS and HT-CTCF PD followed by IB on TCE and HT-PD samples using anti-BORIS, anti-CTCF, and anti-SSRP1 antibodies. B) HT-EV and HT-BORIS PD in the presence or absence of Benzonase® nuclease (Benz) followed by IB on TCE and HT-PD samples using anti-BORIS, anti-SPT16, and anti-SSRP1 antibodies. C and D) MM057 cells were genetically engineered to stably express EV-3xF or hBORIS-3xF upon induction with dox and BORIS expression was induced with 100ng/ml dox for 5 days. C) BORIS-3xF protein level was determined by IB using anti-BORIS antibody. D) Co-IP was performed using anti-Flag antibody followed by IB on TCE (Input) and IP'ed samples (α-Flag IP) with anti-BORIS and anti-SSRP1 antibodies. A-C) Anti-Tubulin antibody was used as a loading control for the TCE.

CHAPTER 3.4

–

Investigating the Role of BORIS in Gene Expression Regulation

Sanne M. Janssen, Mounib Elchebly, Roy Moscona, Sophia Borchers, Léon C.L. van Kempen, Eitan Rubin and Alan Spatz

Background and objectives

In the previous chapters we have demonstrated that BORIS expression alters the transcriptome of melanoma cells. In addition, our proteomics data suggests that BORIS interacts with nucleosome assembly proteins (WICH complex: BAZ1B/WSTF and SMARCA5/SNF2h), histone chaperone proteins (FACT complex) and chromatin remodeling proteins (B-WICH complex: BAZ1B, SMARCA5, MYBBP1A, SF3B1/SAP155). Furthermore, BORIS has previously been shown to regulate gene expression by binding within the gene promoter region and act as a transcriptional activator [151, 189-191] or recruit a transcriptional activator [182], which in some cases was associated with DNA demethylation [159, 189, 192] and altered chromatin modifications [189-193] that can lead to more open and accessible DNA. Together, these data indicate that BORIS can regulate gene expression at different levels. Here we set out to obtain a better understanding of how BORIS expression leads to altered gene expression by assessing BORIS-DNA binding in promoter regions and the effect of BORIS expression on DNA methylation and chromatin accessibility.

Results

3.4.1 BORIS binds within the promoter region of differentially expressed genes

To address the different ways in which BORIS could regulate gene expression, we first set out to address BORIS binding in the promoter region as a possible mechanism. To this end, we established a list of genes for which we aimed to investigate BORIS binding in the promoter region. The list of genes was based on differential expression in the RNA-seq dataset (Section 3.2.2) and putative BORIS binding sites in the promoter region derived from the ENCODE BORIS/CTCFL ChIP-seq track for the K562 cell line (**Fig. 40**). We observed putative binding sites for BORIS in both upregulated genes (*CLDN1, GDF6, PAX6,* and *CD274/PDL1*) and downregulated genes (*CDKN2C* and *RPL10*), suggesting that BORIS binding in the promoter region might be responsible for both gene activation and repression. Primers were designed for ChIP-qPCR based on the consensus DNA sequence surrounding the putative BORIS binding site. In addition, primers recognizing an intergenic region depleted of putative BORIS/CTCF binding sites were designed and served as negative control. To investigate binding of BORIS to the DNA, we used the inducible system described in section 3.3.5, which results in ectopic expression of hBORIS-3xF upon the addition of dox. BORIS expression was induced with 100ng/ml dox for 5 days and ChIP-qPCR was performed using an antibody directed towards the triple FLAG®-tag for the indicated promoter regions and intergenic control region. Compared to the control, our initial data revealed binding of BORIS at the putative binding site for all upregulated (**Fig. 41A**) and downregulated genes (**Fig. 41B**). The highest enrichment was observed for Paired-box 6 (*PAX6*) and the lowest enrichment for Programmed Death-Ligand 1 (*PDL1*). Together, we demonstrate that BORIS is able to bind in the promoter region of DEGs and propose that this is one mechanistic way in which BORIS expression contributes to an altered transcriptome.

A)

Gene	log2FC	FDR
CLDN1	2.51	3.04E-22
GDF6	8.81	1.19E-19
PAX6	2.94	0.00029
CD274/PDL1	2.2	7.14E-11
CDKN2C	-1.25	2.18E-07
RPL10	-0.89	3.90E-06

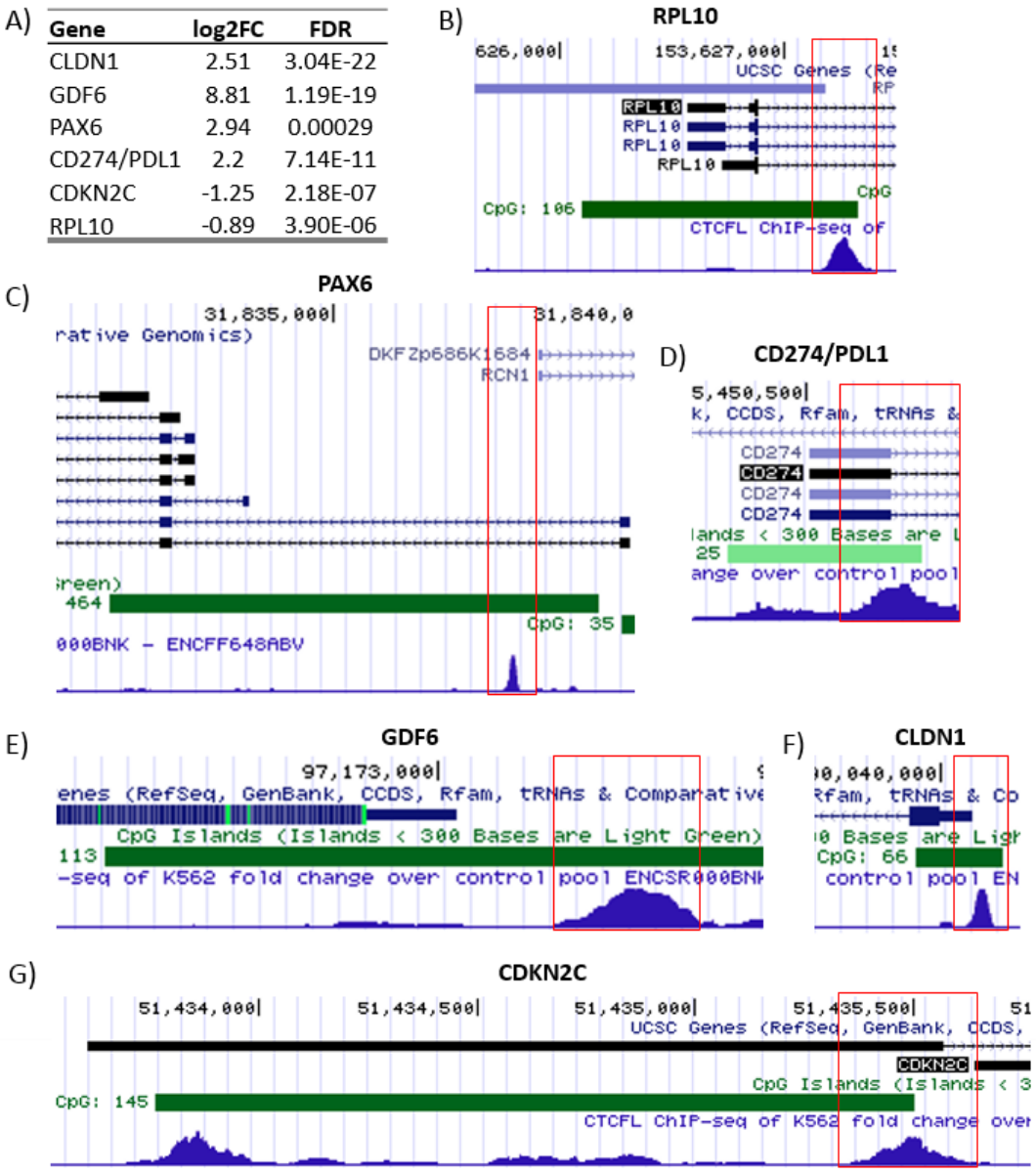


Figure 40. Identification of differentially expressed genes with a putative BORIS binding site.

Six genes were selected for assessment of BORIS binding in the promoter region based on A) DEGs as determined by RNA-seq and B-G) Screenshots from the UCSC genome browser demonstrating the gene of interest (top), CpG island (green) and the putative BORIS binding site (blue track, red box) identified from the ENCODE BORIS/CTCFL ChIP-seq track for the K562 cell line. Log2FC: $\log_2$ fold change, FDR: false discovery rate.

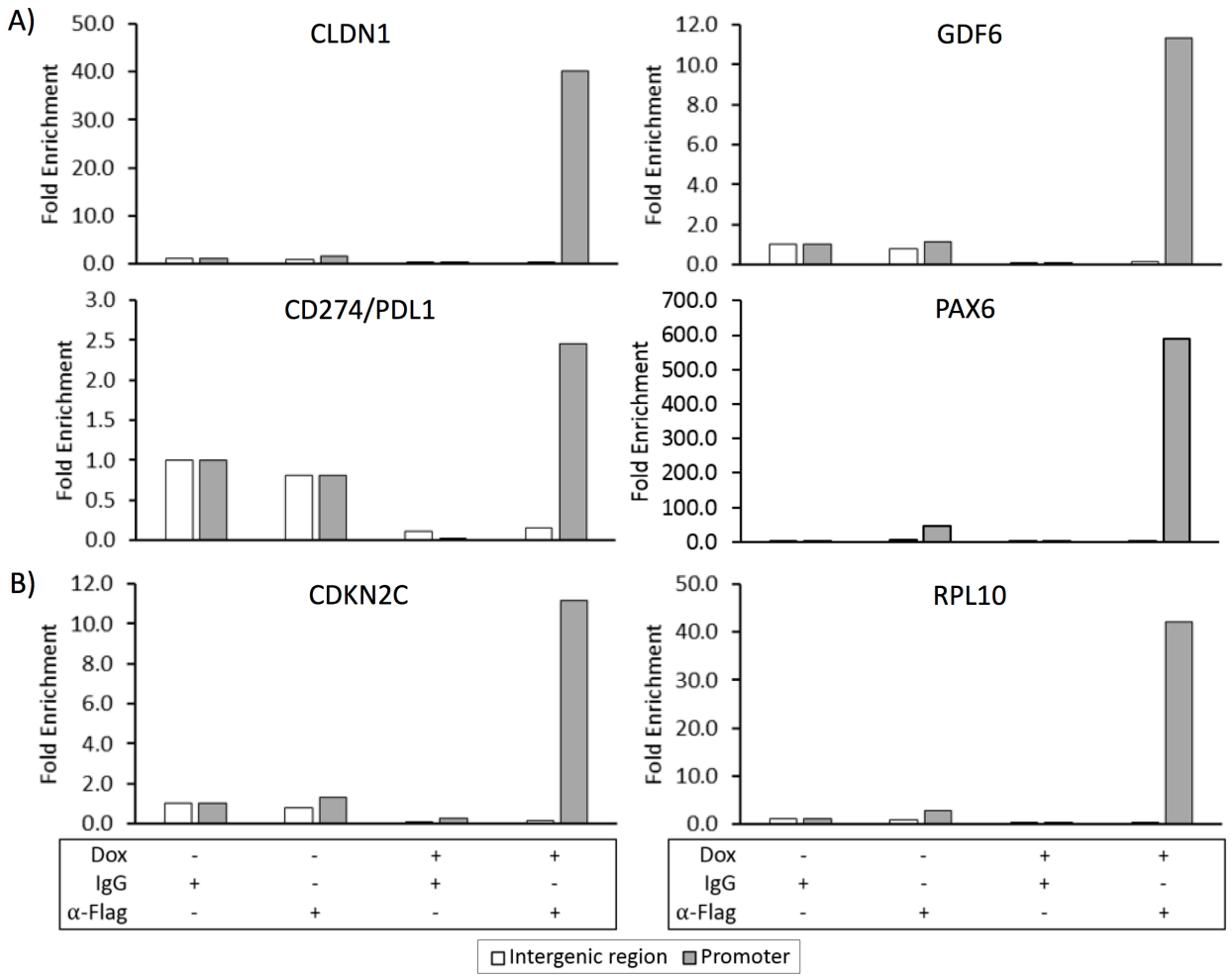


Figure 41. BORIS binds in the promoter region of both up- and downregulated genes.
MM057 cells were genetically engineered to stably express hBORIS-3xF upon induction with dox. BORIS expression was induced with 100ng/ml dox for 5 days and ChIP was performed using anti-Flag antibody or IgG control. gDNA was isolated from the IP'ed samples and qPCR was performed (technical duplicate) with primers directed at the intergenic control region and putative BORIS binding sites within A) upregulated and B) downregulated genes. Data was analyzed using the Bio-Rad CFX Manager™ Software and results are displayed as fold enrichment over the intergenic region.

3.4.2 BORIS binding within the *PAX6* promoter region does not require methylated DNA

One of the genes that is bound by BORIS in the promoter region and upregulated in the RNA-seq dataset is *PAX6*. PAX6 is a transcription factor that is important for early development. Interestingly, PAX6 expression is regulated by CTCF binding in the *PAX6* promoter region [314-318]. The interaction of CTCF with the *PAX6* promoter region occurs in a DNA methylation-dependent manner, whereby DNA methylation inhibits CTCF binding, leading to increased PAX6 expression [314]. PAX6 is overexpressed in various types of cancer [319]. In cancer, the role of PAX6 appears to be cell type specific, with a tumor suppressor-like role in glioblastoma and prostate cancer and tumor promoting role in retinoblastoma, colorectal, breast, pancreatic and non-small cell lung cancer [320-327]. While the role of PAX6 in cutaneous melanoma remains largely unknown, a recent study reported increased PAX6 expression during melanoma progression in some patients, though these results did not reach statistical significance [328]. This is in agreement with cBioPortal data [329, 330], which shows PAX6 overexpression in 6% of cutaneous melanomas present as obtained from the TCGA PanCanAtlas dataset (https://gdc.cancer.gov/about-data/publications/pancanatlas). PAX6 expression was increased in melanoma cell lines upon Adenosine Deaminase Acting On RNA 1 (ADAR1) knockdown-mediated invasion and reduced by inhibition of invasion, suggesting a potential role for PAX6 in melanoma progression [328]. Surprisingly, the *PAX6* promoter region is often found to be hypermethylated in melanoma cell lines [331], a phenomenon that is often associated with tumor suppressor genes. Here, we wanted to address if BORIS can regulate *PAX6* expression by altering the DNA methylation status around the BORIS binding site in the *PAX6* promoter region.

As a first step we aimed to validate the effect of BORIS expression on *PAX6* mRNA expression in the MM057, MM102 and MM117 cell lines. These cell lines were infected with a lentiviral expression vector encoding green fluorescent protein fused to an Internal Ribosome Entry Site fused to a FLAG®-tag fused to BORIS (GIF-BORIS) or used as an EV control (GIF-EV), and expression of *BORIS* was confirmed by qPCR (**Fig. 42A**). Compared to EV control, we observed increased *PAX6* expression in the presence of BORIS (**Fig. 42B**). Next, we compared *BORIS* and *PAX6* mRNA expression in 6 melanoma cell lines. While only 6 melanoma cell lines were assessed, these data suggest a correlation between *BORIS* and *PAX6* expression (Pearson's

correlation (r) = 0.83 and p=0.023; **Fig. 42C**). These results show that BORIS can upregulate *PAX6* expression and indicate a correlation between *BORIS* and *PAX6* mRNA levels.

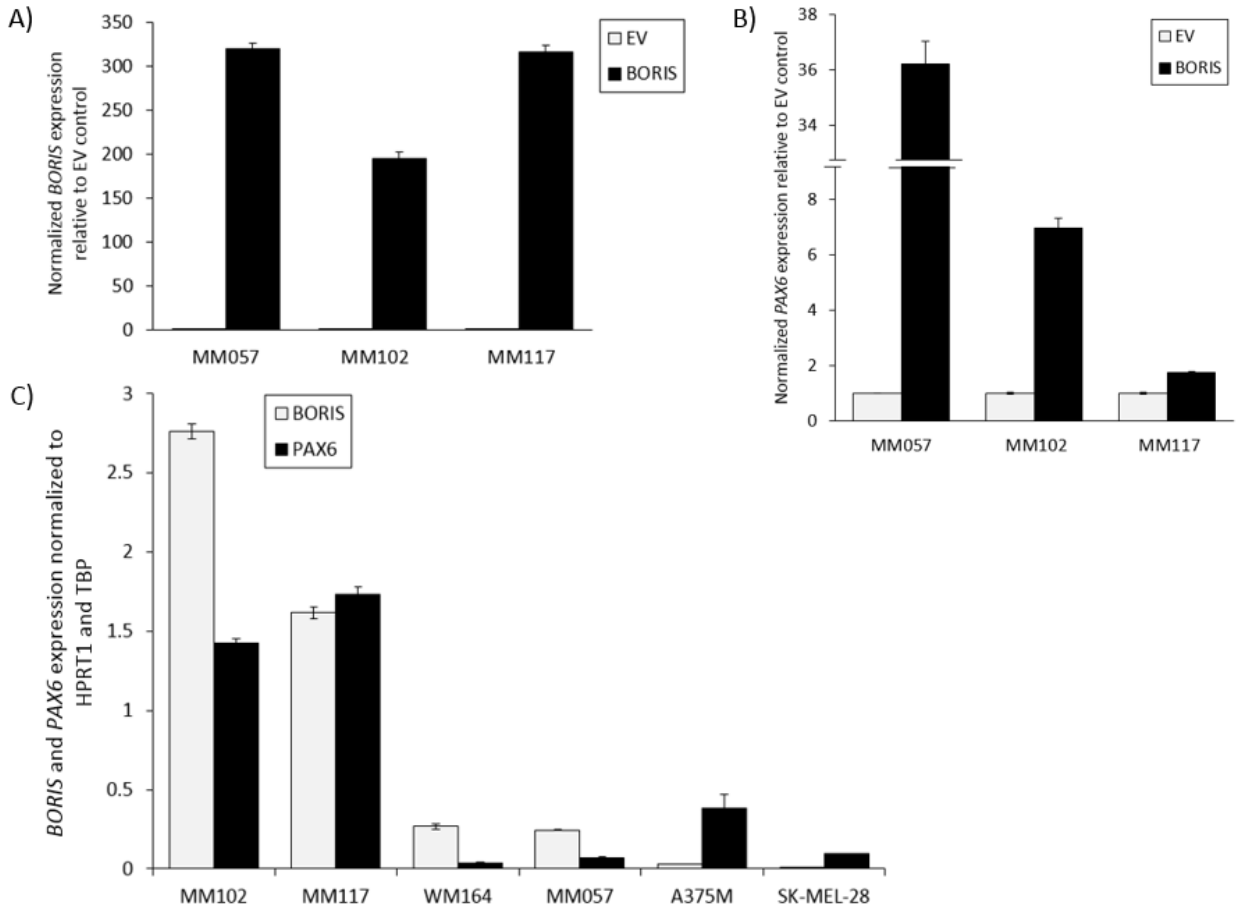


Figure 42. Ectopic BORIS expression results in increased *PAX6* expression.

MM057, MM102 and MM117 melanoma cell lines were infected with the expression plasmids GIF-EV or GIF-BORIS for 72 hours. qPCR) was performed (technical triplicate) for A) *BORIS* and B) *PAX6*. C) qPCR was performed (technical triplicate) on a panel of melanoma cell lines for *BORIS* and *PAX6*. Expression was normalized to *HPRT1* and *TBP* and data was analyzed with the Bio-Rad CFX Manager™ Software. Error bars represent the standard error of mean.

To address if *PAX6* expression is regulated by DNA methylation in melanoma cell lines, we treated the MM102 and MM117 cell lines with the demethylating compound 5-Aza-dC and assessed *PAX6* expression. We observed a large increase in *PAX6* expression (MM102, ~30 fold; MM117 ~9 fold), indicating a role for DNA methylation in the regulation of *PAX6* (**Fig. 43A**). The increase in *PAX6* expression was not due to increased BORIS or CTCF expression, since 5-Aza-dC treatment did not alter *BORIS* or *CTCF* mRNA (**Fig. 43A**). To determine the methylation status of the *PAX6* promoter region at the putative BORIS binding site, we performed bisulfite sequencing. This approach is based on the conversion of unmethylated cytosine to uracil upon treatment with sodium bisulfite, while leaving 5-methylcytosine unchanged. First, several potential 5-methylcytosines (CpG sites) were identified by analyzing the DNA sequence around the putative BORIS binding site. PCR primers specific for bisulfite converted DNA were designed to assess 24 CpG sites in this DNA region. We extracted gDNA from the MM102 cell line and performed bisulfite conversion. The converted DNA was amplified by PCR, resulting in PCR products in which unmethylated cytosines were converted into thymines. This product was cloned into a thymine and adenine cloning vector that was transformed into JM109 competent cells to obtain single colonies for sequencing. Three clones were sequenced and the 24 CpG sites in and around the putative BORIS binding site were assessed for their methylation status. To our surprise we observed no methylation of any of the investigated CpG sites (**Fig. 43B**). Since CTCF was shown to preferentially bind the unmethylated *PAX6* promoter and repress PAX6 expression [314] we assessed if ectopic CTCF expression represses PAX6 expression in melanoma cells. The MM102 melanoma cells were infected with the GIF-CTCF expression plasmid, leading to ectopic *CTCF* expression (**Fig. 43C**). Contrary to the reported repressive effect of CTCF, we observed increased *PAX6* levels as measured by qPCR compared to the EV control (**Fig. 43D**). Interestingly, upon ectopic CTCF expression we also measured increased *BORIS* expression (**Fig. 43D**), which could potentially contribute to the observed elevation in *PAX6*. We did not detect any effect of ectopic BORIS expression on *CTCF* expression (**Fig. 43E**). Altogether, these results demonstrate that both BORIS and CTCF can upregulate *PAX6* expression in melanoma cells. In addition, these data suggest that BORIS binding to the *PAX6* promoter region does not require DNA CpG methylation and that upregulation of *PAX6* in melanoma cells does not depend on BORIS-associated promoter demethylation at the BORIS binding site.

Figure 43

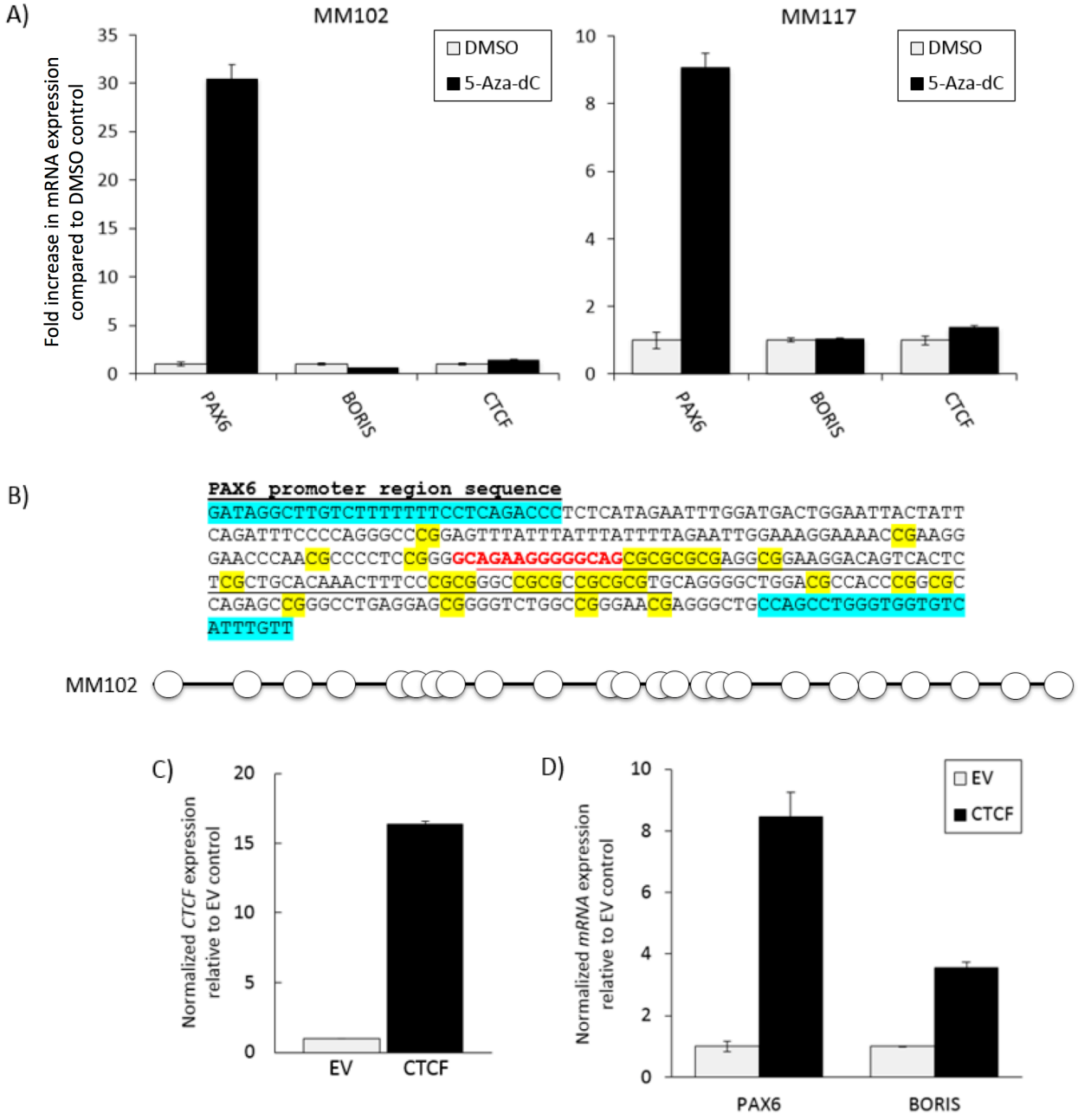

A)
MM102
MM117
DMSO
5-Aza-dC
Fold increase in mRNA expression compared to DMSO control
PAX6
BORIS
CTCF
B)
PAX6 promoter region sequence
GATAGGCTTGTCTTTTTTTCCTCAGACCCTCTCATAGAATTTGGATGACTGGAATTACTATT
CAGATTTCCCCAGGGCCCGGAGTTTATTTATTTATTTTAGAATTGGAAAGGAAAACCGAAGG
GAACCCAACGCCCCTCCGGGGCAGAAGGGGGCAGCGCGCGCGAGGCGGAAGGACAGTCACTC
TCGCTGCACAAACTTTCCCGCGGGCCGCGCCGCGCGTGCAGGGGCTGGACGCCACCCGGCGC
CAGAGCCGGGCCTGAGGAGCGGGGTCTGGCCGGGAACGAGGGCTGCCAGCCTGGGTGGTGTC
ATTTGTT
MM102
C)
Normalized CTCF expression relative to EV control
EV
CTCF
D)
Normalized mRNA expression relative to EV control
EV
CTCF
PAX6
BORIS

Figure 43. Increased *PAX6* expression by DNA demethylation, but not of the BORIS binding site.

A) MM102 and MM117 melanoma cell lines were treated with the demethylating agent 5-Aza-dC or DMSO (control) and *PAX6*, *BORIS* and *CTCF* expression was determined by qPCR (biological triplicate with technical triplicates). B) Top: DNA sequence highlighting the assessed CpG sites (yellow) around the BORIS binding site (underlined, core sequence in red) within the *PAX6* promoter (sequence primers in blue). Bottom: schematic representation of the bisulfite sequencing results from the MM102 melanoma cell line for each CpG site. C and D) MM102 cells were genetically modified to express GIF-CTCF or GIF-EV and C) *CTCF*, D) *PAX6* and *BORIS* expression was determined by qPCR (technical triplicate). E) MM102 cells were infected with the plasmid encoding GIF-BORIS or GIF-EV and *BORIS* expression was evaluated by qPCR (technical triplicate). Expression was normalized to *HPRT1* and *TBP* and data was analyzed with the Bio-Rad CFX Manager™ Software. Error bars represent the standard error of mean.

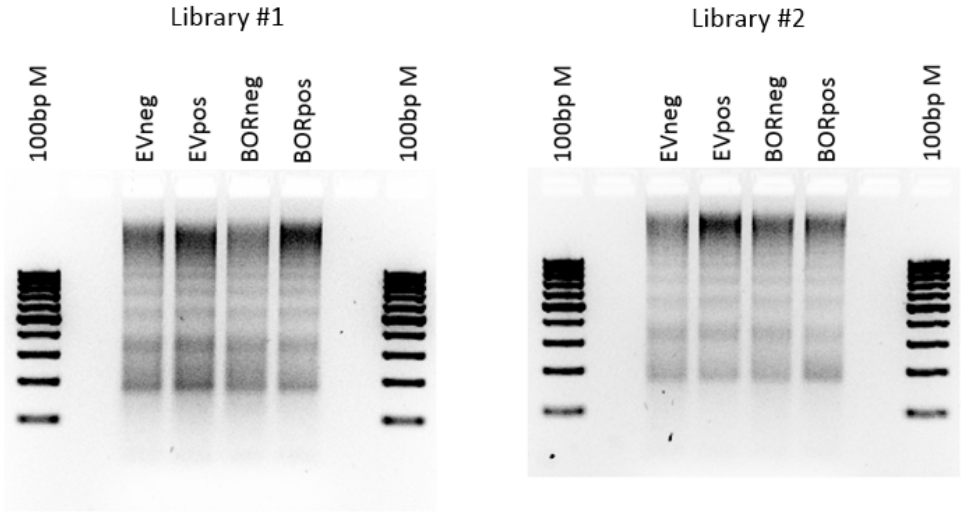
3.4.3 Preparation and validation of the ATAC-seq library

In previous studies BORIS was found to mainly bind at regions of open and active chromatin, including at TSS. In addition, active histone modifications (eg. H3K4me3 and H3K27ac) were associated with BORIS DNA binding [144, 153]. A recent study by Bergmaier *et al.* took advantage of available ENCODE datasets and found that, in contrast to CTCF, BORIS binding to the DNA is highly enriched at open chromatin. These sites of open chromatin were marked by various active histone modifications, including H3K27ac, H3K79me2, H3K4me2/me3, H3K9ac, and the histone variant H2A.Z. Furthermore, they observed BORIS binding at genomic regions bound by transcription factors and confirmed binding of BORIS at TSS [188]. Together these data strongly indicate that BORIS preferentially binds at regions of open chromatin and specifically at TSS, suggesting that BORIS' main way of regulating gene expression is via TSS binding. While these studies address the binding specificity of BORIS, they do not answer the question if altered BORIS expression changes the chromatin landscape resulting in more open chromatin, specifically at TSS.

To determine if BORIS expression increases chromatin accessibility we performed ATAC-seq and used the dox-inducible expression system established in the MM057 cells as described section 3.2.1 (EV-6xH and hBORIS-6xH) to alter BORIS expression. The reason for using this cell model is that it was used previously for the RNA-seq experiments and one of our aims is to project the ATAC-seq data onto the RNA-seq data. This would allow us to investigate the effect of potential changes in chromatin accessibility on the observed alterations in gene expression as described in section 3.2.2. To generate the ATAC-seq libraries BORIS expression was induced with 50ng/ml dox for 5 days and 50.000 nuclei were isolated from EVneg, EVpos, BORneg and BORpos. Library construction relies on the prokaryotic, hyperactive transposase Tn5, which cuts unprotected regions of DNA and simultaneously inserts sequencing adapters [332]. The nuclei were subjected to the Tn5 tagmentation reaction, resulting in tagmentation of accessible chromatin with the adapter sequences. These adapters were used for PCR to add Illumina primer sequences and a barcode to the 5' and 3' end of the PCR products. Next, the libraries were amplified with the established number of PCR cycles to obtain the required amount of DNA for sequencing without generating high numbers of amplification-related PCR duplicates. A successful Tn5 transposase reaction will cut and tagment the DNA in between nucleosomes,

resulting in a so-called nucleosome pattern [332-334]. As a quality control the ATAC-seq libraries were analyzed on agarose gel, which demonstrated the expected nucleosome pattern for all samples (**Fig. 44A**). Next, we performed qPCR on DNA from the libraries for known regions of open and closed chromatin and confirmed that the libraries are enriched for open chromatin (**Fig. 44B**). One of the biggest challenges with ATAC-seq is contamination of the sequencing library with mitochondrial DNA (mtDNA). mtDNA is present in the cytoplasm and can be highly abundant depending on the cell type, ranging from about 20-70% of the total DNA. We performed qPCR to approximate the level of mtDNA in our libraries (established from purified and washed nuclei) as well as libraries prepared from whole cell lysate using 2 sets of primers that are referred to as mtDNA and ND1-mtDNA. As expected, we were unable to entirely deplete mtDNA from the libraries. However, compared to the libraries produced from whole cell lysate, the quantitation cycle values for mtDNA in the libraries from isolated nuclei were higher, indicating lower levels of mtDNA. On average the quantitation cycle value for the mtDNA primers was increased by 4.8 in library #1 and 3.4 in library #2, and for the ND1-mtDNA primers by 4.7 in library #1 and 3.2 in library #2 (**Fig. 44C**). Next, sequencing libraries were size selected using magnetic beads (150bp - 750bp) and submitted for sequencing at the McGill University and Genome Québec Innovation Centre. As a final step before sequencing, all libraries were run on a bioanalyzer to determine DNA concentration and confirm both size selection and nucleosomal profile (**Fig. 45**). Sequencing results of the first biological replicate provided an average of ~170 million reads (100bp, paired-end) and demonstrated high quality sequencing reads with an average duplication rate of 32%. Quality controls were performed and upon data processing we obtained an average of ~85 million non-mtDNA, non-Y-chromosome, and non-duplicate paired-end reads with a mapping quality Q-score >20 (**Table VII**). Visualization of the fragment size distribution using the *CollectInsertSizeMetrics* tool demonstrated increased read lengths consistent with nucleosomes (eg ~200bp peak represents mononucleosome) and about 50% of the post-filtering reads as short fragments (<100bp) [333, 334], which represents nucleosome-free open chromatin (**Fig. 46**). Currently, analysis for differentially accessible regions is being performed in collaboration with Dr. Eitan Rubin to address the effect of BORIS expression on chromatin accessibility.

A)



B)

	EVneg	EVpos	BORneg	BORpos
GAPDH	21.345	21.495	22.125	21.915
KAT6B	21.275	21.02	21.27	21.115
SLC22A3	27.075	27.06	27.61	27.175
RHO	25.765	25.655	26.125	26.185

	EVneg	EVpos	BORneg	BORpos
GAPDH	21.285	20.485	20.65	20.255
KAT6B	20.425	19.795	20.075	20.025
SLC22A3	26.945	25.815	26.05	25.785
RHO	25.39	24.425	24.73	24.825

C)

mtDNA	EVneg	EVpos	BORneg	BORpos
WCL	8.83	9.38	10.24	9.64
Library #1	14.16	14.34	14.14	14.56
Library #2	13.03	12.51	12.46	13.54

ND1-mtDNA	EVneg	EVpos	BORneg	BORpos
WCL	8.45	8.88	9.49	9.27
Library #1	13.68	13.89	13.55	13.92
Library #2	12.06	11.99	11.92	12.99

Figure 44. ATAC-seq library quality control.

A) Part of the DNA of each ATAC-seq library was analyzed by agarose gel to verify the presence of a nucleosome pattern. B and C) Quantitative cycle (Cq) value for each target as assessed by qPCR on DNA from ATAC-seq libraries. B) Primers for *GAPDH* and *KAT6B* were used to amplify regions of open chromatin and *SLC22A3* and *RHO* for regions of closed chromatin. C) Two sets of primers (mtDNA and ND1-mtDNA) were used to amplify mtDNA. WCL: library prepared from whole cell lysate instead of isolated nuclei.

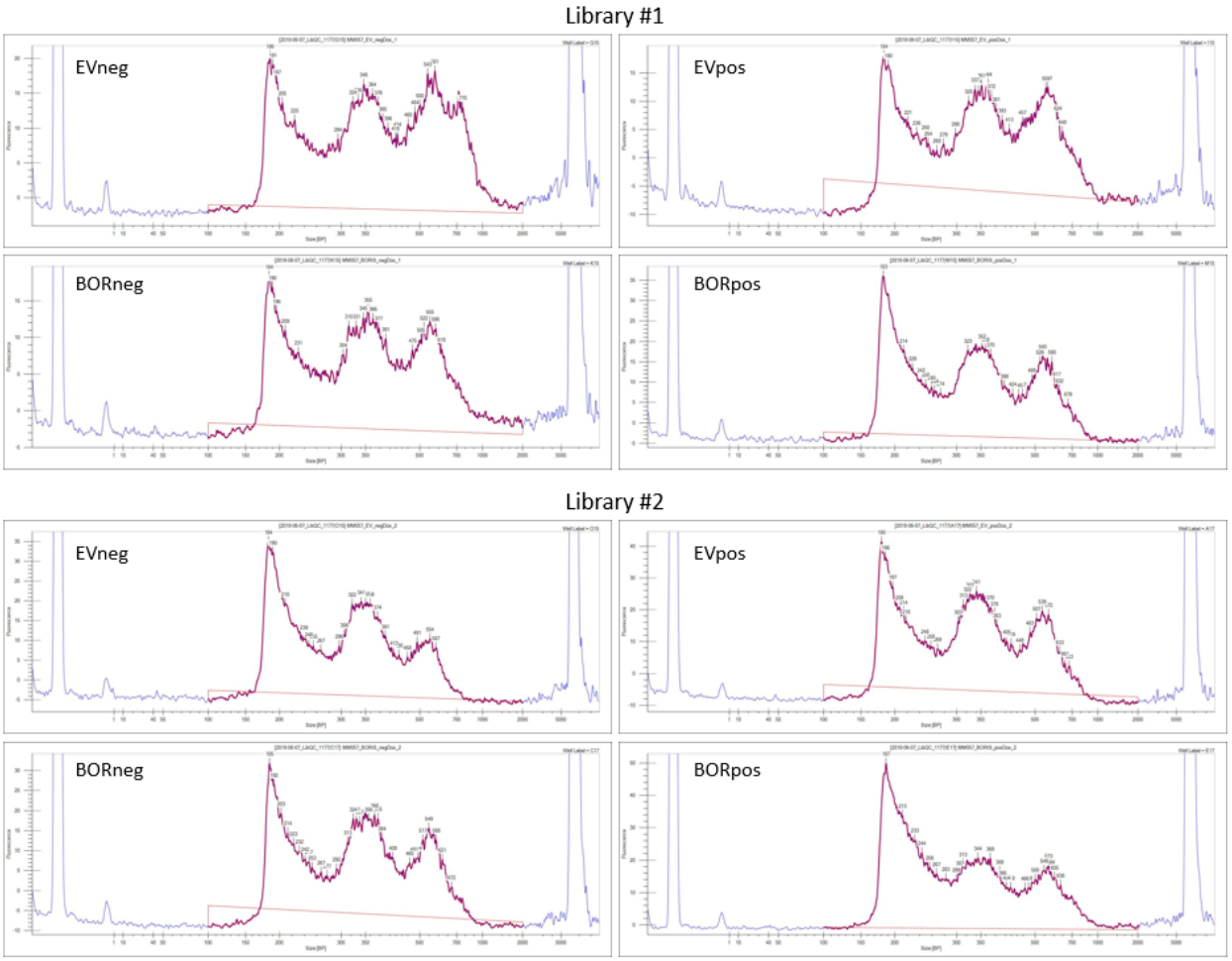


Figure 45. Bioanalysis profile of the ATAC-sequencing libraries.

As a quality control to rule out primer contamination and confirm the presence of a nucleosome pattern, the amount of DNA was measured (fluorescence) and plotted versus the fragment length (size in bp) for each ATAC-seq library using TapeStation.

Table VII: Overview of the ATAC-sequencing read statistics

Sample	**EVneg**	**EVpos**	**BORneg**	**BORpos**
Total reads	170,249,181	186,116,876	170,369,568	162,923,298
Mapped reads	154,804,648 (90.93%)	170,689,262 (91.71%)	158,221,560 (92.87%)	143,495,544 (88.08%)
Properly paired reads	144,974,306	161,207,938	150,015,666	135,571,842
Duplication rate	31.54%	30.21%	35.15%	30.62%
mtDNA reads	23,894,070	23,491,198	29,950,657	19,028,967
Y-chr reads	389,398	419,826	350,680	388,374
Final reads for peak calling	84,362,705	95,952,118	83,089,595	77,670,407

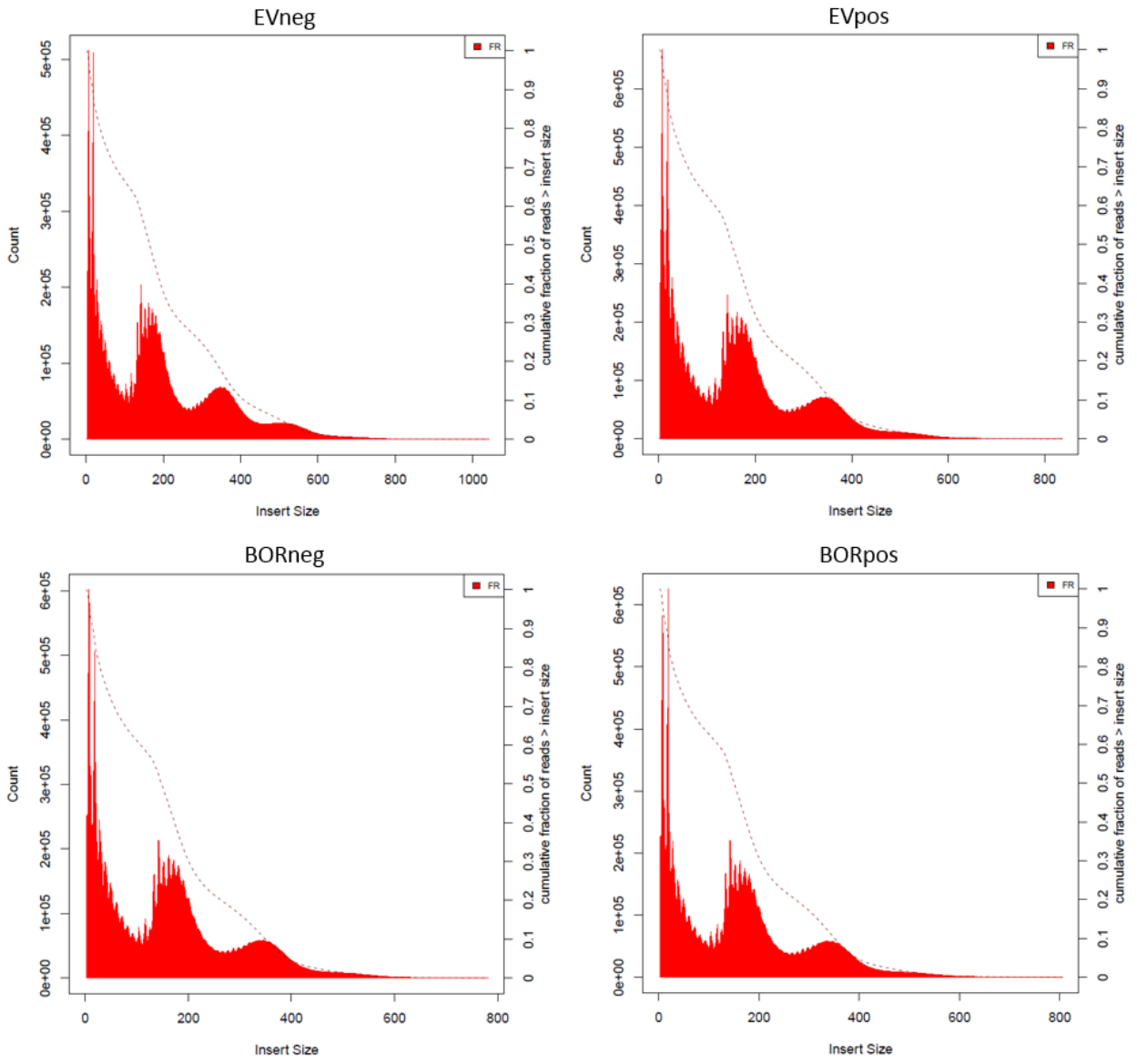


Figure 46. ATAC-sequencing insert size histograms.

For quality control, the number of sequencing reads (post-filtering) was plotted versus the insert size (bp) for each library to assess fragment size distribution using CollectInsertSizeMetrics tool from the Picard toolbox.

CHAPTER 4

–

DISCUSSION AND CONCLUSION

4.1 SUMMARY

Here we demonstrate that BORIS, by altering gene expression and interacting with a variety of protein partners, increases the invasive potential of melanoma cells and contributes to DNA damage. Furthermore, we provide insight into regulatory mechanisms involved in BORIS-induced gene expression changes. Taking into account our results together with existing literature, we propose a model that describes the effect of BORIS expression in melanoma cells and discusses how BORIS can promote melanoma progression (**Fig. 47**).

On the one hand BORIS expression promotes an invasive phenotype, which can be mainly attributed to BORIS-induced large-scale differential gene expression. We believe this is achieved through a combination of the following 3 mechanisms that are involved in the regulation of gene expression:

1) BORIS' ability to bind within gene promoters and function as a transcriptional activator [151, 189-191] and our data)

2) Recruitment of chromatin modifying proteins that promote open, active chromatin [189-193], which allows binding of transcriptional activators, like AP-1. In addition, BORIS itself functions as a recruitment factor for transcriptional activators, including Sp1 [182].

3) Recruitment of chromatin modifying proteins that repress transcription, like the Nucleosome Remodeling and Deacetylation (NuRD) complex of which subunits were identified as potential BORIS-interacting proteins, and promote binding of transcriptional repressors.

BORIS-induced changes in chromatin accessibility will be revealed upon analysis of the ATAC-seq data. Altogether, we propose that these regulatory mechanisms promote downregulation of proliferative genes and upregulation of invasive genes, resulting in an increased invasive potential of melanoma cells.

In addition to upregulation of the invasive genes from the Verfaillie gene signature [59], we find that ectopic BORIS expression leads to upregulation of the transcriptional regulator *ZBTB33*, also known as Kaiso. *ZBTB33* is located on the X-chromosome and its expression is repressed on the Xi as a result of XCI. We observe expression of *ZBTB33* from the Xi upon BORIS overexpression. The mechanisms by which BORIS allows escape of inactivated genes remain unknown, but it is likely related to the interplay with CTCF-DNA binding as discussed below.

Interestingly, increased ZBTB33 expression and localization to the nucleus is linked to migration and invasion [259, 260, 335-338], indicating that ZBTB33, possibly via its role as transcriptional regulator, can contribute to invasion and melanoma progression.

On the other hand, we propose that BORIS expression induces genomic instability, as demonstrated previously [157], via increased DNA damage and altered repair mechanics. This notion is supported by our observation that BORIS expressing cells have an increased sensitivity to DNA damage inducing agents. What causes the increase in DNA damage upon BORIS expression remains unknown. Possibly BORIS recruits chromatin modifying proteins that promote open, accessible DNA [189-193], which is more prone to DNA damage from, for example, reactive oxygen species. In addition, the increase in DNA damage may be a result from inhibition of DNA repair that allows accumulation of damage from replication errors. It is tempting to speculate that the interaction between BORIS and the FACT complex plays a role in genetic instability based on the implication of FACT in DNA replication and repair [313, 339, 340]. In agreement, knockdown of FACT increases the sensitivity of cells to UVC and ionizing radiation [310-312]. Possibly the interaction of BORIS with FACT interferes with FACT-mediated nucleosome remodeling at the replication fork [309], leading to DNA damage via replication fork stalling and collapse. In addition, FACT localizes at sites of DNA damage where it might inhibit repair [311]. However, findings related to the exact implication of FACT in DNA repair seem contradictory and possibly depend on the type of cell and DNA damage [312]. Overall, our findings suggest that BORIS expression in melanoma leads to a state where the rate of DNA damage is higher than the rate of DNA repair, which could contribute to the previously reported genomic instability that was observed upon ectopic BORIS expression [157].

It is important to note that we have not included a potential interplay between BORIS and CTCF in our model. This is done for two reasons. 1) We did not investigate the effect of BORIS expression on CTCF's DNA binding ability. 2) The interplay between these two proteins requires more insight. One of the major questions in the field at the moment is if BORIS and CTCF interact cooperatively or competitively, as mentioned in a recent review by Lobanenkov and Zentner [186]. It was the identification of two types of CTCF binding sites, 1xCTSes and 2xCTSes, that sparked this debate. In their study, Pugacheva *et al.* observed simultaneous

occupancy of 2xCTSes by BORIS and CTCF, though BORIS-BORIS and CTCF-CTCF binding was found at some sites as well. Interestingly, transcription factors and chromatin modifying proteins are enriched at 2xCTSes compared to 1xCTSes, suggesting that these sites play an important role in gene expression regulation [144]. Based on these findings we could speculate that the effects observed upon BORIS expression in melanoma cell lines are likely a result of interference with CTCF binding at 2xCTSes, leading to disruption of CTCF-mediated chromatin loops, boundaries between chromatin states, and alterations in the recruitment of transcription factors and chromatin modifiers.

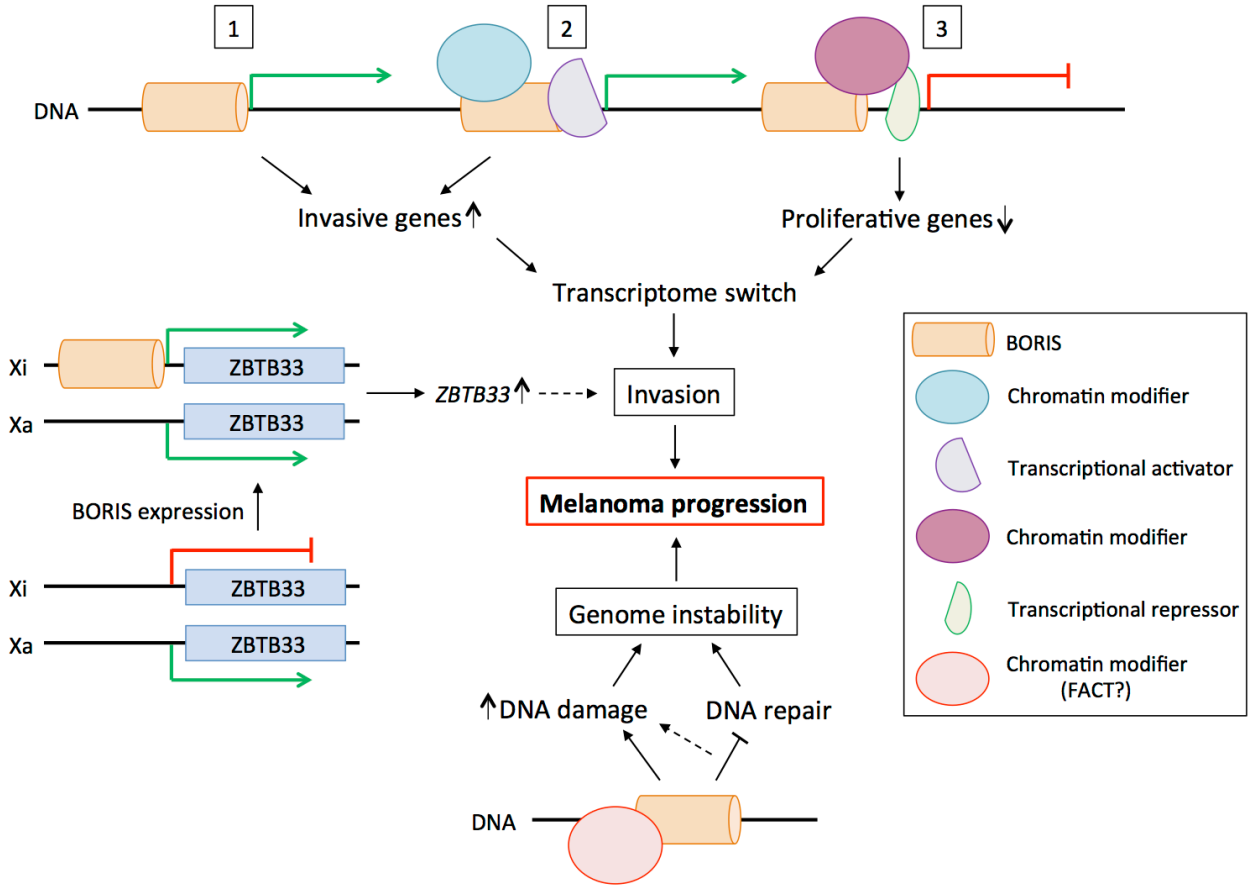


Figure 47. Model of BORIS' involvement in and regulation of processes related to melanoma progression. BORIS promotes melanoma progression through its DNA-binding abilities and the recruitment of proteins to the DNA, leading to increased genome instability and a higher invasive potential of melanoma cells. Top: Via various mechanisms that regulate gene expression (number 1 to 3) BORIS promotes the expression of genes belonging to the invasive gene signature, while genes from the proliferative gene signature are repressed. Left: In addition, the presence of BORIS results in expression of *ZBTB33* from the Xi via an unkown mechanism. Previous studies have demonstrated that ZBTB33 is involved in migration and invasion, suggesting increased ZBTB33 expression could promote melanoma progression (dotted arrow). Bottom: BORIS expression leads to increased DNA damage and slower DNA repair kinetics, which contribute to genome instability. Inhibition of DNA repair also contributes to increased DNA damage (dotted arrow). Chromatin remodelers and transcriptional activators/repressors that might be involved in the observed processes are indicated in the figure legend in between brackets. For further details see the text above.

4.2 BORIS AND CTCF EXPRESSION IN MELANOMA

We detected BORIS and CTCF expression at the RNA and protein level in melanoma cells. Our observations regarding BORIS expression in melanoma are in agreement with previous findings [138, 159, 162] in that we detect expression of *BORIS* in ~60% of the tested melanoma cell lines. In addition, using IB with anti-BORIS antibody, our data indicates that BORIS protein is expressed in 4 out of 12 metastatic melanoma cell lines. Unfortunately, due to the lack of a reliable commercial BORIS antibody [280], the detection of BORIS at the protein level remains a difficult task and often leads to questionable results. To confirm that the observed IB bands are indeed BORIS protein, experiments using siRNA targeting BORIS in the K562 cell line are ongoing. Another challenge is the existence of 23 alternative *BORIS* transcripts that encode 17 distinct protein isoforms with predicted molecular weights ranging from ~10 to 80-90kDa [146]. We detect BORIS at ~90kDa, which likely represents full length BORIS encoded by the isoform transcript family sf1 [146]. We believe this is true BORIS protein for three reasons: 1) the band for BORIS is at the same molecular weight as exogenous His- or Flag-tagged BORIS, 2) at this molecular weight we detect a strong band in the K562 cell line, which has been shown to express BORIS in various studies [144, 145, 195, 196], and 3) when performing a pull down with HT-BORIS followed by IB for BORIS we detect a band at ~90kDa. Interestingly, the latter result demonstrated an additional band migrating slightly lower at ~80kDa, which indicates the presence of at least two BORIS isoforms in the HEK293T cells that were used for this experiment.

In the tested melanoma cell lines, the presence of BORIS protein did not correlate with mRNA expression. This lack of correlation between BORIS mRNA and protein was also described for various cell lines and tissue types by Jones *et al.* [156]. It was proposed that this discrepancy could be due to a short half-life of the *BORIS* transcript and high protein stability [169]. However, others did observe a correlation between BORIS mRNA and protein level [158].

We observed CTCF mRNA and protein expression in all tested melanoma cell lines, which is consistent with previous observations [156]. Importantly, the expression of both BORIS and CTCF in melanoma cells suggests that a possible competition between CTCF and BORIS DNA binding [139, 141, 153] could takes place in BORIS-positive melanoma cell lines.

Renaud *et al.* showed that siRNA mediated knockdown of CTCF in normal human dermal fibroblasts induces *BORIS* expression [145], which would suggest higher BORIS levels in cells with low CTCF expression. On the contrary, we observed no inverse correlation between BORIS and CTCF expression at neither the mRNA nor protein level. Furthermore, instead of BORIS repression by CTCF we show increased *BORIS* level upon ectopic *CTCF* expression in the MM102 melanoma cell line and observe BORIS protein in cell lines with high CTCF expression. These data suggest a potential positive correlation between BORIS and CTCF expression in melanoma cell lines. However, this effect was not seen at the mRNA level. Furthermore, we find that ectopic BORIS expression does not alter *CTCF* expression. Therefore, it appears that there is no correlation between BORIS and CTCF expression in melanoma cells.

4.3 BORIS AND ALLELE-SPECIFIC GENE EXPRESSION

Our data provides the first evidence that BORIS can modulate gene expression in an allele-specific manner. Previous work from our laboratory demonstrated changes in X-linked gene expression upon BORIS knockdown in melanoma cells using NanoString gene expression profiling [220]. The X-chromosome is widely used as a model to study allele-specific expression due to the silencing of one allele, the Xi. These observations raise the question if BORIS expression can lead to aberrant expression of X-linked genes in melanoma cells as a result of escape from the Xi. To address this question, we altered BORIS expression in a female melanoma cell line with one Xa and one Xi.

We identified only one melanoma cell line, MM057, with one Xa and one Xi. While unfortunate, this was expected, since cancer cell lines are known to possess a high number of chromosomal aberrations and can have multiple X-chromosomes [341]. As we were unable to detect endogenous BORIS protein in the MM057 cells we generated a model system with stable ectopic BORIS expression. Consistent with our previous observation, we confirm altered expression of X-linked genes upon ectopic BORIS expression. A comparison of the differences in X-linked gene expression obtained with either ectopic BORIS expression or BORIS knockdown [220] shows no overall opposite effect on expression level. This may be due to the use of different cell lines as each contains a different number of X-chromosomes. We were able to confirm the

expression changes observed upon ectopic BORIS expression in our RNA-seq data for 7 out of 9 genes (*BEX2, RAP2C, ZBTB33, MST4, PIM2, SAGE1 and RBBP7*). In addition to the validation that was already done, this further validates the RNA-seq dataset. Furthermore, it confirms the effect of ectopic BORIS expression on X-linked genes in two different overexpression systems and highlights the sensitivity of qPCR over RNA-seq. Noteworthy, not all expression changes for the 7 genes are considered significant according to the criteria set for DEGs (FDR<0.05, not for *BEX2, RAP2C* and *PIM2*), which demonstrates that we used a stringent cut-off for the identification of DEGs in the RNA-seq data.

SNPs are commonly used to assess allele-specific expression. Out of the 9 differentially expressed X-linked genes we were able to validate the presence of SNPs in only 2 genes, the inactivated gene *ZBTB33* and escaping gene *PPP2R3B*. This is not surprising considering that a typical human genome, which comprises ~ 3 billion base pairs, contains only 4.1-5 million SNPs. In addition, many SNPs are located outside the coding region and 5'/3' UTRs that are transcribed into mRNA. By exploiting the SNP in *ZBTB33* we demonstrate that ectopic expression of BORIS induces *ZBTB33* expression from the Xi, suggesting that BORIS expression can lead to local loss of the repressed state of inactivated genes on the Xi. We propose that this could be due to a competition between CTCF and BORIS for DNA binding sites, resulting in loss of CTCF binding on the Xi that contributes to the loss of repressive chromatin marks. In support of this hypothesis, heterozygous loss of CTCF in endometrial tumors was shown to result in global loss of demethylation on the X-chromosome [212]. Furthermore, CTCF binding on the Xi is clustered at genes that are known to escape from the transcriptionally inactive status of the Xi [207, 211], indicating a role for CTCF as an insulator protein in maintaining boundaries on the Xi between active and repressive chromatin. Importantly, in cancer the Xi is often disrupted or lost, as detected by disappearance of the Barr body [342-345], including in melanoma as recently demonstrated by our laboratory [222]. In addition abnormalities of XCI have been reported in various cancers [221, 346]. Loss of the Barr body can be a result of Xa duplication or (partial) Xi reactivation. While most studies have focused on genetic instability as a cause for Barr body loss [345, 347], more recent studies have started to investigate epigenetic instability as an underlying mechanism. For example, it was shown that DNA methylation at the promoter of inactivated genes is important for maintenance

of gene repression and is reduced at the promoters of escaping genes [348]. Furthermore, Chaligne *et al.* demonstrated that the Xi in breast tumors and cell lines often displays increased euchromatic marks and disruption of repressive marks, like DNA methylation and H3K27me3. This epigenetic instability resulted in aberrant reactivation of tumor promoting genes from the Xi [341], which demonstrates the significance of the chromatin landscape for proper maintenance of XCI. These findings highlight the importance for a better understanding of factors and processes that interfere with the inactive state of the Xi.

To investigate if the effect of BORIS expression on activation from the Xi holds true for other X-linked genes, we established a model with known SNPs. Most research on allele-specific gene expression is performed with mouse-derived cell lines for two reasons: 1) XCI can be completely skewed either to the paternal or maternal allele by the use of *Xist* or *Hprt* mutations [211, 343] or 2) Crossing two mice with highly polymorphic backgrounds will generate offspring with a large number of SNPs that are, in the case of common mouse strains, known and listed in the Mouse Genome Informatics database. Upon selection of single clones, RNA-seq can be used to look at the expression of mono-allelic genes to determine if the maternal or paternal X-chromosome is inactivated [349-351]. For our model we crossed a male C57BL/6J and female FVB/NJ mouse to obtain MEF cell lines with known SNPs. According to the Mouse Genome Informatics database, offspring from this cross contains a SNP in *Xist*, which would allow us to identify if *Xist* is expressed from the maternal or paternal allele. Indeed, we were able to validate the *Xist* SNP in our female MEF clones and found that in two clones the maternal X-chromosome is inactivated (pXa, mXi) and in one clone the paternal X-chromosome is inactivated (mXa, pXi). To the best of our knowledge, we are the first to establish a model to study allele-specific expression based on the presence of a SNP in *Xist*. The advantage of this model is that there is no need to generate a mutant mouse or perform RNA-seq to determine from which mouse, the maternal or paternal, the X-chromosome is inactivated. Furthermore, immortalized MEFs are easy to maintain in culture. Finally, this model allows us to study the effect of BORIS expression on allele-specific X-linked gene expression in the context of a maternal Xi and paternal Xi. Due to the difference between mouse and human BORIS protein we are currently establishing dox-inducible mBORIS expression in the newly generated MEF cell lines.

Many questions remain to be answered to obtain a better understanding of the involvement of CTCF and BORIS in XCI. It will be important to determine DNA binding sites of BORIS on the Xa and Xi and compare them with CTCF binding sites in order to identify preferential BORIS sites (gene promoters versus insulator sites) and address competition between BORIS and CTCF for DNA binding on the X-chromosome. In addition, the effect of BORIS expression on promoter DNA methylation and chromatin marks that determine chromatin accessibility remains unknown. We anticipate that a better understanding of the interplay between BORIS and CTCF on the Xi and the effect on expression of tumor promoting genes will also provide insight into a potential role for BORIS in the regulation of autosomal genes involved in tumor progression.

4.4 THE FUNCTIONAL ROLE OF BORIS IN MELANOMA

According to Kholmanskikh *et al. BORIS* expression is higher in metastatic melanoma compared to primary tissue samples, which led to the proposal of a possible role for BORIS in melanoma progression [138]. To the best of our knowledge, no studies have tried to elucidate the biological functions of BORIS in melanoma progression. As a first step to obtain insight into a potential role for BORIS, we investigated the effect of aberrant BORIS expression on the transcriptome of melanoma cells. In addition, we identified BORIS-interacting proteins.

4.4.1 Identification of BORIS' functions based on BORIS-induced transcriptional changes

Here we demonstrated large-scale differential gene expression upon the induction of low level BORIS expression in the BORIS negative MM057 melanoma cell line. Previously, a few studies have investigated the effect of altered BORIS expression on genome wide gene expression [144, 153, 195, 196], with only one these studies reporting differential gene expression of a large number of genes [144]. The difference between our observation and that of others reporting very few DEGs might be due to the use of different cell types and techniques. In addition, the sensitivity of a RNA-seq experiment is dependent on the obtained read depth, which can alter the number of detected DEGs. The large number of DEGs that we detected was expected, since BORIS can act as a transcription factor [151, 189-191] and is involved in a large number of cellular processes [142, 149]. Despite the difference in the number of identified DEGs, we did identify some overlap in genes between our study and previous studies, including *FGFR4, GAL3ST1* and *FOXA3* that were upregulated in MCF7 cells with ectopic BORIS expression. In

addition, *SDC2* is downregulated in our dataset and upregulated upon BORIS knockdown in K562 cells. We also observe upregulation of *AKR1B1* and *CA2,* which were upregulated upon BORIS knockdown as well [144, 195]. Furthermore, we identify *PRSS50*, a known BORIS target gene [144, 153], as differentially expressed. However, this gene is not in our list of DEGs as it is not expressed in the control samples and therefore does not pass the criteria of CPM>0.5. While the use of CPM>0.5 excludes genes that are only expressed in the BORIS positive sample, we use this filter to prevent skewing the fold change in gene expression towards low expressed genes, which would interfere with GSEA analysis. The robustness of our RNA-seq dataset is highlighted by our validation approach, which demonstrates that genes throughout the list of DEGs, including the last gene of the list, are differentially expressed between BORIS positive and negative control cells as assessed by qPCR.

Consistent with the number of different functions for BORIS reported previously, GSEA on our RNA-seq data reveals enrichment for many biological processes. In addition, we identify many overrepresented pathways among up- and downregulated genes. This was expected given the large number of DEGs. While we observed overlap between the identified biological processes and pathways among upregulated genes, this was not the case for downregulated genes. One striking observation was the presence of multiple EMT-related processes and pathways, including chemotaxis and migration, cell cycle, ECM organization, and various signaling pathways involved in EMT, like the MAPK and WNT/Ca^{2+} pathways.

In agreement with the identified processes and pathways, we demonstrate a switch from an epithelial-like state (proliferative phenotype) to a mesenchymal-like state (invasive phenotype) upon ectopic BORIS expression in melanoma cells. We confirm this switch at the transcriptome level as well as phenotypically by an increase in migration of melanoma cells in the presence of BORIS expression. At the transcriptome level, previously reported gene signatures [51, 59] were used for each phenotypic state to look for enrichment of BORIS-induced DEGs. This demonstrates an extraordinarily high enrichment for upregulated DEGs among the Verfaillie [59] invasive gene set (NES=5.99) as well as a negative enrichment for the proliferative gene set (NES=-3.84). The overlap with the Hoek gene signature [51] was less significant, likely due to the small number of genes within this gene set. In addition, we validate the upregulation of

invasive genes and downregulation of proliferative genes by qPCR. Consistent with our findings, previous work has reported a link between BORIS expression and EMT in a neuroblastoma cell line [215].

Specific to melanoma is the switch from ZEB2 expression in melanocytes to ZEB1 expression in melanoma cells. ZEB1 functions as a transcriptional repressor of *CDH1* [34] and is transcriptionally activated by SLUG (*SNAI2*) [37]. We find that ectopic BORIS expression results in the downregulation of *SLUG* (log2FC -1.1, FDR 2.41E-07) without an effect on *ZEB1* expression (log2FC -0.2, FDR 0.52). Perhaps, we observe no increase in *ZEB1* expression, because *CDH1* expression in the MM057 cells is already very low. Interestingly, *SPARC* is upregulated upon ectopic BORIS expression. SPARC is a matricellular protein involved in cell-matrix interactions that promotes EMT-like changes [352] and is associated with the metastatic potential of melanomas [353]. The increase in *SPARC* expression level and its reported role in melanoma are consistent with the increase in migration that we observe upon ectopic BORIS expression. SPARC is also a transcriptional activator of SNAIL (*SNAI1*) [283], which is upregulated in our RNA-seq dataset. Increased SNAIL expression in melanoma cells was shown to induce EMT via upregulation of invasive markers and reduced immunosuppression, which is in part driven by increased secretion of TGF-beta (*TGFB*) and thrombospondin (*THBS*) [282]. In agreement, we observe increased expression of *TGFB1* and *THBS2* (log2FC 1.49, FDR 1.39E-08), suggesting that the upregulation of SNAIL is involved in the BORIS-induced phenotype switch. Another factor that is used to differentiate proliferative from invasive melanoma cells is MITF. Expression of MITF is downregulated in invasive cells compared to proliferative cells [32] and is regulated by the switch in ZEB expression [34, 35]. In our model, BORIS expression does not result in reduced *MITF* expression. This might be due to the fact that MM057 cells do not require increased ZEB1 to repress CDH1, which is already low in the MM057 cells. Based on our findings, we speculate that BORIS expression induces partial transcriptional phenotype switching, via the upregulation of *SPARC* and *TGFB1* that leads to *SNAI1* expression. The idea of partial phenotype switching comes from Zhentao *et al.* who proposed that the observations related to phenotype switching in melanoma are not always consistent due to fluctuating expression of EMT-associated genes that leads to a spectrum of EMT states [354]. Consistent with this idea, Noguchi *et al.* showed that interleukin-like EMT inducer ILEI regulates partial

phenotype switching. They state: “EMT is not a simple on/off switch but a long gradient with multiple paths between the epithelial and mesenchymal states” [355]. A similar observation of partial EMT is found in other types of cancer as well [356, 357].

Another possible explanation for the discrepancy in ‘benchmark’ gene expression switches upon BORIS expression is related to the cell line that was used. To our surprise, the RNA-seq data reveals that the MM057 cell line, being described as a proliferative cell line, demonstrates no detectable *CDH1* (E-cadherin) and high *CDH2* (N-cadherin) expression. This is in contrast with the described gene expression profile for proliferative cell lines, which is high *CDH1* and low *CDH2* expression [59, 354]. Therefore, the use of another cell line with a known proliferative transcriptome, like the MM074 cell line, could have demonstrated changes in cadherin, ZEB and MITF expression as shown previously [51, 59]. Despite not observing a cadherin or ZEB switch upon BORIS expression, the decrease in proliferation of MM057 cells and increase in migration of MM057 and MM074 cells upon BORIS expression confirm the contribution of BORIS to a switch from a proliferative to invasive state. To strengthen these findings we are currently performing matrigel invasion assays that allow us to assess the invasive capacity of these cell lines in the presence and absence of BORIS expression.

Various transcriptional regulators involved in the switch towards an invasive state have been identified, including AP-1, NFIB, NFATC2 and the TEAD family [59, 61]. These are likely involved in the transcriptome switch that we observe, since *in silico* data analysis reveals that only 7.6% of the upregulated invasive genes contain a putative BORIS binding site in the promoter region. In agreement with the reported regulators, we observe upregulation of the AP-1 family members *FOSL2* (*FRA2*; log2FC 0.53, FDR 0.05)*, FOSB* (log2FC 1.83, FDR 0.01) and *JUNB* (log2FC 0.67, FDR 3.0E-04) and downregulation of the transcriptional repressors *NFIB* (log2FC -0.88, FDR 6.16E-06) and *NFATC2* (log2FC -1.36, FDR 5.12E-13) upon ectopic BORIS expression. On the contrary, none of the TEAD family members was differentially expressed, nor was the AP-1 subunit *FOSL1* (*FRA1*). Possibly, the different transcriptional regulators are involved at different phases of the transition from a proliferative to invasive state. Regarding regulation of *FOSL2, FOSB, JUNB, NFIB* and *NFATC2* expression by BORIS, ENCODE data tracks for BORIS binding and CTCF-mediated chromatin looping indicate a role

for direct BORIS binding in the promoter region as well as the formation of BORIS-BORIS or BORIS-CTCF chromatin loops. While STAT1/2 has previously been associated with the proliferative state [61], we identify STAT1/2 binding motifs as enriched in the TSS of invasive genes and observe upregulation of *STAT1* (log2FC 0.69, FDR 0.06). Perhaps, this is due to the partial transcriptional phenotype switch induced by BORIS expression. Future ChIP experiments using the newly established MM057 cell line model with inducible BORIS-3xF will provide more insight into BORIS-mediated regulation of transcriptional regulators of the invasive state. From a clinical perspective, it is important to identify factors involved in partial EMT, since this state is associated with a more aggressive state of the tumor [29].

It is important to note that the RNA-seq data reveals no differential expression for most of the known CTAs upon ectopic BORIS expression, including *MAGE-A1-6*, *NY-ESO-1* (*CTA1B*) and *SSX-2*. This observation is in agreement with findings from Kholmanskikh *et al.* [138]. Interestingly, we do observe activation of one X-CTA cluster, CT45, in cells with BORIS expression compared to controls. The CT45 cluster includes *CT45A1-6* and *SAGE1* [358]. We have validated the increase in *SAGE1* expression upon ectopic BORIS expression in the MM057 cell line by qRT-PCR, since *SAGE1* was among the X-linked genes in our gene list. According to putative BORIS binding sites identified from the ENCODE ChIP-seq track for K562 cells some of these CTAs contain a BORIS binding site in the promoter region (*CTA45A1*, *CT45A6* and *SAGE1*). In addition, the ENCODE CTCF ChIA-PET track for K562 cells demonstrates a potential BORIS-CTCF-mediated chromatin loop encompassing *CT45A4-6* and *SAGE1*, indicating that BORIS is likely involved in activation of these CTAs via both regulatory mechanisms. Importantly, CT45A1 expression in a breast cancer model demonstrated that this CTA functions as a proto-oncogene and promotes EMT [359]. On the other hand, CT45(A1) knockdown inhibited migration and invasion in a lymphoma, breast cancer and fibrosarcoma cell line [359, 360]. These findings suggest that BORIS-induced upregulation of the *CT45* cluster might also contribute to the phenotype switch and promote melanoma progression.

4.4.2 Identification of BORIS' functions based on BORIS-interacting proteins

Considering the large number of cellular processes that were identified by GSEA on BORIS-induced DEGs and the importance of BORIS-specific proteins partners for divergence from

CTCF functions, it was likely that BORIS interacts with additional and different protein partners than the 25 reported so far [144, 167, 181-184]. With the availability of more advanced techniques to study protein-protein interactions, we investigated BORIS-interacting proteins and identify 79 new BORIS protein partners. A combination of SILAC and the HT-PD technology was used followed by MS. The advantage of using SILAC is that the PD and MS are performed on a pooled sample that includes lysate of cells transfected with the control and protein of interest (BORIS) to avoid bias in PD efficiency and MS experiment inter-variability. To isolate control and BORIS-interacting proteins we used the HT system, which relies on a fusion between the HT protein and BORIS covalently linked to the HT ligand that binds the HT resin used for PD. The main advantage of this system is the high affinity and specificity, it being a bacterial protein, resulting in low background binding [361].

As mentioned before, HT-PD with HT-BORIS followed by IB with anti-BORIS antibody revealed two bands at the predicted molecular weight of endogenous BORIS isoforms. Assuming that these bands are truly BORIS protein, this finding is interesting for three reasons. Firstly, it demonstrates the presence of endogenous BORIS protein, while being undetectable in the total cell extract. Secondly, it reveals BORIS expression in the non-cancerous cell line HEK293T, which has been reported previously [169, 191]. Though care needs to be taken with this statement, since the HEK293T cell line is known to harbor a highly altered genome. This could result in duplication of the BORIS gene or epigenetic alterations leading to demethylation and activation of the BORIS gene promoter. In addition, BORIS was demonstrated to upregulate its own expression, meaning that the expression of HT-BORIS could induce endogenous BORIS expression. Thirdly, it suggests that BORIS is able to dimerize with itself, which has been suggested by Jones et *al.* [156], but to the best of our knowledge has not been demonstrated previously.

Among the identified BORIS-interacting proteins we observe multiple proteins that were identified previously [167, 183], demonstrating that HT-BORIS is a functional fusion protein, our approach worked and the data is consistent with previous studies. To our surprise, we did not identify CTCF in the MS results, while CTCF is a known BORIS partner [144, 181], verified by us via HT-PD. Why we were unable to detect CTCF remains unknown. Previous studies

investigating CTCF interacting partners unfortunately do not mention if they were able to detect CTCF protein by MS [183, 362-364]. Our MS approach reveals three common interacting proteins between BORIS and CTCF (UBF, PARP and SMC1), of which only UBF has been reported previously [183]. While we had anticipated the identification of more common protein partners between BORIS and CTCF as a result of their highly similar zinc finger domain, our findings are consistent with the idea that BORIS and CTCF have divergent functions due to different interacting proteins. In addition, *in silico* analysis predicted very little overlap between CTCF and BORIS protein partners [185]. Another possible reason for the low number of common protein partners is that CTCF interacting proteins described in the literature are almost all experimentally confirmed and do not include complete lists of putative CTCF protein partners as identified by MS.

Overrepresentation analysis for biological processes and pathways among the 79 newly identified BORIS-interacting proteins revealed enrichment of proteins involved in mRNA and rRNA processing. This is in agreement with BORIS' ability to bind RNA and its suggested role in mRNA splicing [169]. In addition, BORIS was shown to localize to the nucleolus, bind rDNA, interact with the rDNA transcription factor UBF, and alter rRNA expression [157, 183]. Importantly, an increase in nucleolar number and size is considered a hallmark of cancer. Furthermore, extensive proteomic analysis has revealed a role for nucleolar proteins in various processes other than ribosome biogenesis, including mitosis, cell cycle progression, and DNA replication and repair [365, 366]. These finding highlight the importance of proper functioning of the nucleolus. In an effort to further address the role of BORIS in the nucleolus, we were able to confirm previous studies reporting BORIS' presence in the nucleoli and interaction with UBF. To our surprise and in contrast to the reported decrease in pre-rRNA expression upon BORIS knockdown [157], we did not observe an effect of ectopic BORIS expression on rRNA level in the MM057 melanoma cell line. The reason for this discrepancy, other than potential cell type dependency, remains unknown and warrants further investigations into the role of BORIS in the nucleolus. Besides BORIS, CTCF was also reported to localize to the nucleolus [367], bind rDNA, change nucleoli number and size [183] and alter rRNA expression [368]. This suggests that there may be competition between BORIS and CTCF to bind rDNA and interact with UBF to regulate rRNA expression. The ability of both BORIS and CTCF to localize to the nucleolus

suggests involvement of the zinc finger domain. Indeed, two putative nucleolar localization sequences have been identified in the BORIS/CTCF zinc finger domain [156]. Wether these sequences are required for BORIS and CTCF localization remains to be determined.

The overrepresentation analysis also revealed enrichment of proteins involved in DNA repair. This sparked our interest, since recent evidence points to a role for CTCF in the DDR [289-291]. Consistent with a potential role for BORIS in the DDR, we demonstrate an increase in H2AX phosphorylation and activation of the DDR upon ectopic BORIS expression. In addition, we show that ectopic BORIS expression results in an increased cell number in the S-phase, which is in agreement with the findings from Rosa-Garrido *et al.* [157]. Based on these data, we propose that BORIS plays a role in the intra-S-phase DNA damage checkpoint. This idea is supported by the identification of proteins that play a role in this checkpoint among BORIS-interacting partners. Interestingly, loss of CTCF increased the sensitivity of cells to DNA damage-inducing agents [290] and we observe the same effect with increased BORIS expression, indicating an opposite role for BORIS and CTCF in the DDR. Furthermore, it suggests that BORIS expression sensitizes cells to DNA damaging agents, which could be exploited therapeutically. These findings are consistent with the observed genomic instability upon ectopic BORIS expression in HEK293T cells [157]. To better understand the role of BORIS in the DDR, it needs to be confirmed that BORIS, like CTCF, can localize to the site of DNA damage. Related to the role of CTCF in the DDR, Hilmi *et al.* propose an interesting role for CTCF in DNA repair, whereby CTCF-mediated chromatin loops facilitate the recruitment of repair factors. In addition, CTCF as a boundary protein maintains a defined chromatin domain in a DNA repair-promoting chromatin state [290]. If true, this opens another possibility for how BORIS might interfere with CTCF's function in the DDR. Perhaps BORIS competes for DNA binding sites at CTCF-CTCF chromatin loops, resulting in BORIS-CTCF chromatin loops that alter the chromatin conformation and hinder DNA repair.

Compared to potential cellular functions of BORIS based on transcriptional changes, we did not identify processes related to EMT or invasion among BORIS interacting proteins. This indicates that the role of BORIS in the switch from a proliferative to invasive phenotype is mainly driven by its DNA binding capacity and function as transcriptional regulator. Nevertheless, BORIS-

interacting partners that have an effect on transcriptional regulation may contribute to the transcriptional switch. These may be transcription factors or proteins that can modulate chromatin accessibility. Indeed, we observe proteins involved in histone and chromatin remodeling as well as nucleosome assembly in the list of BORIS-interacting proteins, which is in agreement with previous findings [167, 184]. Whereas BORIS expression could modulate the expression of transcription factors such as AP-1, NFIB and NFATC2, we identified only one transcription factor, the transcriptional repressor GATAD2A (p66α), out of a curated list of 1639 human transcription factors [369] among BORIS binding partners. Interestingly, GATAD2A belongs to the NuRD complex, an ATP-dependent chromatin remodeler involved in gene repression via histone deacetylation [370]. These findings suggest that, besides BORIS' function as transcription factor, BORIS' ability to recruit chromatin modifying proteins to the DNA likely contributes to the BORIS-induced transcriptional switch.

Here we demonstrate for the first time that BORIS interacts with both subunits of the FACT complex, a histone chaperone that plays an important role in transcription, replication and DNA repair [302, 307-310]. We revealed that the interaction with SSRP1 and SPT16 is not mediated by nucleic acids. In addition, co-IP confirmed the interaction between BORIS and SSRP1. Unfortunately, we were unable to verify the interaction with SPT16 by co-IP, which is likely due to the poor quality of the SPT16 antibody. Interestingly, a recent study demonstrated the importance of FACT for new H2AX incorporation at sites of DNA damage [312], which is crucial for DNA repair. This exchange reaction is dependent on PARP-mediated ADP-ribosylation of STP16, which is mediated by the interaction of FACT with PARP [313]. Furthermore, FACT localized to sites of DNA damage and inhibited DNA repair upon ionizing radiation [311], though repair efficiency did not change upon UVC radiation [312]. Since we showed that BORIS interacts with both FACT and PARP, and plays a role in DNA damage repair, we propose an interplay between BORIS and FACT in the DDR. Besides the role of SSRP1 and SPT16 as subunits of the FACT complex, these proteins also have FACT-independent functions. We cannot rule out that BORIS-mediated recruitment of SSRP1 or SPT16 plays a role in FACT-independent processes. Besides BORIS, CTCF was identified as a potential interacting protein of SSRP1, suggesting that FACT is not a BORIS-only interacting protein complex. Since SPT16 was not identified among CTCF protein partners in this study

[364], it is unknown if CTCF interacts with SSRP1 as part of the FACT complex or assists SSRP1 in FACT-independent functions. Furthermore, if and how the interplay between BORIS, CTCF and FACT affects the functional outcome of this histone chaperone complex remains to be investigated.

The gene expression and proteomics data both reveal involvement of BORIS in mRNA processing. We identify a number of proteins among BORIS-interacting partners that are part of various splicing related protein complexes, strongly indicating a role for BORIS in RNA splicing. In agreement with these findings, a functional link between BORIS and RNA splicing has been reported. In addition, Ogunkolade *et al.* also revealed that BORIS binds mRNA and is associated with actively translating ribosomes, which was the first indication of BORIS involvement in protein synthesis [169]. We show here that BORIS interacts with many proteins involved in translation, including proteins of the eukaryotic initiation factor (eIF) complexes eIF3 and eIF4F as well as multiple ribosomal proteins. Interestingly, melanoma progression is associated with phosphorylation of one of the eIF4F subunits, eIF4E [371] that promotes EMT-induced invasion in transformed MEFs when phosphorylated [372]. Perhaps BORIS alters the expression of serine/threonine kinases or phosphatases, resulting in increased phosphorylation of eIF4E. In agreement with this idea, ectopic BORIS expression leads to downregulation of *PPP2R3B*, a subunit of the protein phosphatase 2A that alters eIF4E phosphorylation (our laboratory, unpublished). Further investigation is needed to better understand the effect of BORIS expression on translation. This will be an intriguing avenue to explore, since tumor cells rely on increased protein synthesis and many studies are looking to target protein synthesis as a therapeutic intervention for melanoma [373].

4.5 GENE EXPRESSION REGULATION BY BORIS

Analysis of our RNA-seq data revealed large-scale differential gene expression upon ectopic BORIS expression. Though it remains unknown how aberrant BORIS expression contributes to altered gene expression. Our *in silico* analysis showed that about 38% of the DEGs are predicted BORIS target genes, indicating that BORIS likely alters gene expression via a combination of regulatory mechanisms. To obtain a better understanding of the means by which BORIS regulates the DEGs from our RNA-seq data we investigated BORIS-DNA binding in the

promoter region, the DNA methylation status at a specific BORIS-DNA binding site, and the effect of BORIS expression on chromatin accessibility.

Here we demonstrated BORIS binding in the promoter region of six DEGs (*CLDN1, GDF6, PDL1, PAX6, CDKN2C* and *RPL10*) for which BORIS binding has not been reported previously. We identified putative BORIS binding sites for these genes based on the ENCODE BORIS ChIP-seq data track for K562 cells. Simultaneously, we verified the presence of a putative CTCF binding site at the same location, since BORIS and CTCF binding sites display high similarity. This was important for us, because the ENCODE project contains only one BORIS ChIP-seq dataset that can be used to identify putative BORIS binding sites. It is important to mention that this dataset contains various audits (flags) that ENCODE uses to provide additional information about the quality of the dataset, which in this case includes an audit for extremely low read depth and poor library complexity. In addition, the antibody that was used (sc-98982) has not been validated according to ENCODE standards. To circumvent the lack of a reliable commercial antibody we generated a dox-inducible system to express low levels of BORIS protein fused to a triple flag-tag that allowed us to assess BORIS-DNA binding by ChIP-qPCR with an anti-Flag antibody. To our surprise, we observe BORIS binding in the promoter region of not just upregulated genes, but downregulated genes as well. How BORIS binding in the promoter region leads to gene repression remains unknown. We speculate that this occurs via BORIS-altered chromatin looping, since the BORIS binding sites identified for the downregulated genes are also part of a CTCF ChIA-PET data track for the K562 cell line. If BORIS is indeed involved in the formation of chromatin loops, either by BORIS-BORIS dimers or BORIS-CTCF dimers, is an important question that remains to be answered. In addition, future experiments are needed to address if BORIS itself acts as a transcription factor at the identified binding sites or if BORIS is involved in the recruitment of chromatin modifying proteins.

To investigate the relation between BORIS expression and DNA methylation status we focused on *PAX6* for the following reasons: 1) *PAX6* is upregulated upon BORIS expression in the RNA-seq data, 2) BORIS binds in the promoter region of *PAX6*, 3) the *PAX6* promoter region is often methylated in melanoma cell lines [331], 4) CTCF binds to the *PAX6* promoter region when demethylated, resulting in repression of PAX6 expression [314, 316-318]. We postulated that, in

contrast to CTCF, BORIS can bind within a methylated *PAX6* promoter region, leading to DNA demethylation and increased gene expression. In agreement, we confirm that DNA demethylation by treatment with the DNA demethylating agent 5-Aza-dC results in increased *PAX6* expression. To our surprise, the region surrounding the putative BORIS/CTCF binding site was not methylated, suggesting that BORIS binding in the *PAX6* promoter does not require DNA methylation. We observe that ectopic CTCF expression results in increased *PAX6* levels in the MM102 melanoma cell line, which is in contrast to previous studies [316, 318]. We find that ectopic CTCF expression also upregulates *BORIS* expression. Perhaps increased BORIS level upon ectopic BORIS expression, leads to replacement of CTCF at the unmethylated DNA binding site, relieving the repressive effect of CTCF and allowing PAX6 expression. Another possibility is that BORIS expression leads to demethylation of CpG sites in the *PAX6* promoter region that were not part of the assessed region, which would allow for a transcriptional activator to bind and induce PAX6 expression. Overall, the link between BORIS expression and DNA demethylation remains a topic of debate and requires further investigation.

Experimental and *in silico* studies have revealed that BORIS expression is associated with active histone modification [189-193], BORIS can recruit chromatin modifiers to the DNA [167, 184], and BORIS binding is enriched at regions of open chromatin [144, 153, 188]. These findings indicate a link between BORIS expression and chromatin accessibility and point to the possibility that altered chromatin accessibility underlies the BORIS-associated changes in gene expression observed in our RNA-seq data. This prompted us to investigate the effect of BORIS expression on genome-wide chromatin accessibility.

In this age of next-generation-sequencing, the techniques and bioinformatics pipelines to assess open, accessible chromatin have improved rapidly. Currently, there are 3 techniques that are commonly used, including Deoxyribonuclease I-hypersensitivity site-sequencing (DNase-seq), Formaldehyde-Assisted Isolation of Regulatory Elements-sequencing (FAIRE-seq), and ATAC-seq. All these techniques are based on the principle of chemical or enzymatic separation and isolation of either inaccessible or accessible chromatin, meaning chromatin that is or is not wrapped around histones or bound by transcription factors, respectively. We use ATAC-seq to investigate the effect of BORIS expression on chromatin accessibility, since this techniques

requires less cells and experimental steps compared to DNase-seq and produces lower background signal compared to FAIRE-seq (For a thorough review on these techniques see [374]). ATAC-seq was first introduced by Buenonostro *et al.* in 2013 [333] and its use has since increased rapidly over the past few years. This has resulted in optimized protocols as well as better bioinformatics pipeline standards and requirements [375-378].

One of the advantages of ATAC-seq is that, in addition to analyzing chromatin accessibility, the data can also be used to assess nucleosome positioning and perform transcription factor footprinting, though the latter does require a large number of exclusively mapped sequencing reads (> 200 million) [332, 333]. Nucleosome positioning is of particular interest, as previous studies implicated BORIS in altered nucleosome occupancy at gene promoter regions [189, 379]. One of the main challenges with ATAC-seq is contamination of the sequencing library with mtDNA. While new methods have been developed to circumvent this problem, including clustered regularly interspaced short palindromic repeats-based targeting of mtDNA [376], we find that optimization of the lysis steps decreases the number of mtDNA reads to ~25 million out of a total number of ~170 million reads (~15%). In agreement with described ATAC-seq quality guidelines, we have obtained a high quality ATAC-seq dataset. We anticipate that ongoing analysis will provide insight into BORIS-mediated changes in chromatin accessibility as a mechanism of gene expression regulation by BORIS.

4.6 CONCLUSION

To conclude, we here demonstrate for the first time that BORIS expression in melanoma cells can induce allele-specific expression of an X-linked gene, promotes phenotype switching to the invasive state, and increases DNA damage accompanied by activation of the DDR. In addition, we have set the first steps in an effort to elucidate the mechanisms by which BORIS alters the transcriptional landscape of melanoma cell. Furthermore, we identified new BORIS-interacting proteins and show a nucleic-acid independent interaction with the histone chaperone FACT. Finally, we also generated a new model to study allele-specific gene expression in mouse embryonic fibroblast cell lines with a known XCI status based on a SNP in *Xist*. Taken together, the results presented in this work indicate a role for BORIS in melanoma progression via both increased genome instability and enhanced invasive capacity. We acknowledge that the

presented work has its limitations, including the lack of a BORIS knockdown modelsystem, due to difficulties detecting endogeous BORIS with the commercially availalbe antibodies, and the use of few cell lines. Future work aimed at deciphering BORIS-DNA binding and its role in transcription regulatory mechanisms, specifically in the context of CTCF, will be extremely valuable towards understanding how BORIS regulates the expression of genes that are central to the observed phenotypic alterations.

It is important to note that during the writing of this thesis Loukinov presented new data regarding BORIS vaccination as anti-cancer treatment. Their most recent work, currently in preparation for publication, demonstrates promising results with the use of anti-BORIS immunotherapy in a rodent breast cancer model, which, together with previously acquired data, has prompted the initiation of a phase 1 clinical trial for breast cancer patients [280]. While these are very exciting developments, it stresses the importance for a better understanding of BORIS' regulation of and role in cellular processes in both normal and cancer cells to prevent or anticipate potential side-effects and possibly enhance therapeutic benefits from targeting BORIS expressing cells in humans.

APPENDICES

Appendix I. Primers

Cloning

Primer	Sequence (5'-3')
BORIS-1F_NheI	GAGGCTAGCACCATGGCAGCCACTGAGATCTCTGTCCTTTCTGAG
BORIS-663R_EcoRI	CGCGAATTCCTTATCCATCGTGTTGAGGAGCATTTCAC
CTCF_1F-NheI	CGGGCTAGCCGCCATGGAAGGTGATGCAGTCGAAGCCATTG
CTCF_727R-EcoRI	GACGGAATTCCCGGTCCATCATGCTGAGGATCATCTC
HK-F	TCCAAAGCTTGCCACCATGCATCATCACCATCACCATATCGATGT GGGTGAGGATAGCGTGCTGATCACC
HK-R	GGCCGGATCCGCTGTGCCCCAGTTTGCTAGGCAGGTCGCAG
mBORIS-F-NcoI	CCATGGCTGCCGCTGAGGT
mBORIS-636R-AgeI	CTTATACCGGTCTTATCCATCATGTTAAAG
hBORIS-F-NcoI	CACCATGGCAGCCACTGAGATCTC
hBORIS-663R-AgeI	TCCACCGGTCTTATCCATCGTGTTG
attB-BORIS-1F	GGGGACAAGTTTGTACAAAAAAGCAGGCTCCACCATGGCAGCCA CTGAGATCTCTGTCC
attB-mBORIS-MF_R-Stop	GGG GAC CAC TTT GTA CAA GAA AGC TGG GTC TCA AAC CTT ATC GTC GTC ATC CTT GTA ATC
attB-3Falg_R-Stop	GGGGACCACTTTGTACAAGAAAGCTGGGTCTCACTAGTCACGCTA GCATCTGACGCGCTA
BORIS-2F-BamHI	CAGCGGATCCACTGAGATCTCTGTCCTTTCTGAG
BORIS-663R-Stop_MluI	GCGGACGCGTCTCACTTATCCATCGTGTTGAGGAGCATTTCACAG G

MEF genotyping

Primer	Sequence (5'-3')
SRY-mouse_F	TTGTCTAGAGAGCATGGAGGGCCATGTCAA
SRY-mouse_R	CCACTCCTCTGTGACACTTTAGCCCTCCGA
FABP2-mouse_F	CCTCCGGAGAGCAGCGATTAAAAGTGTCAG
FABP2-mouse_R	TAGAGCTTTGCCACATCACAGGTCATTCAG

SNP profiling

Primer	Sequence (5'-3')	SNP
BEX2_Ex3F	CCTTGGCCCTACCTTTGAAT	rs7557
BEX2_Ex3R	GAGCATCTTTCCATGCAATAGG	
ZBTB33_Ex2F	CTGAGCCTTTACCTCCTGATTCT	rs201958171
ZBTB33_Ex2R	TCAGATGTGTTGGAGCTGAAGA	
ZBTB33_rs201958171_R	TGCTAGCACTAGGTGCAACATC	
ZBTB33_3UTR_1F	GGCTTTCATTGCCATAGTCTGT	rs7882480
ZBTB33_3UTR_1R	CATTTACAGACGTTCAACAATTCTG	rs11537755

ZBTB33_3UTR_2F	GGCAAAAGTCCTTCATTCCTG	rs1044413
ZBTB33_3UTR_2R	GTACTCGGAGGAATGAACAAGG	
MST4_Ex4F	CCAGGTTGTGAGCTCCTTCAT	rs5933055
MST4_Ex4R	CCTGAATGGCAAAGATCTAGCA	
MST4_3UTR_F	TGTTGAGACACCGTTTTGCT	rs995246
MST4_3UTR_R	TGAGCTGCCTGGAATATCATAC	rs201033420
SAGE1_Ex2F	CGCCAATTTATGTGCCAGTTAC	rs5930805
SAGE1_Ex2R	CAACCAATCCACTCTTGGATCA	
SAGE1_Ex18F	GACAGGACATTGTATGCCACC	rs4829799
SAGE1_Ex18R	AAGAGGCAGCTCAATTCTGTTC	rs4829584
SAGE1_Ex4F	ACTTACCTCACAGCTCAACCACTT	rs12014884
SAGE1_Ex4R	TCATGGATGAAATTCCTGCTG	
PPP2R3B_Ex13_3F	TGTGGAGCAGAAGCTGAGTG	rs1133530
PPP2R3B_EX13_3R	CATTCCGTACAAACGCACTC	rs14127
		rs1133535
PPP2R3B_rs1133530_F2	GCTCTCGGACTGGGAGAAGTA	rs1133530
PPP2R3B_rs1133530_R	TCATTCCGTACAAACGCACTC	rs14127
		rs1133535
PPP2R3B_Ex12F	TGTACGTAACCCGTGTGCTTC	rs1133520
PPP2R3B_EX12R	GCACACACACCCGTCCTC	rs2738376
PPP2R3B_Ex2F	CAGGTTCAGACACGGAAAGAA	rs3813594
PPP2R3B_Ex2R	CACTCACGCTCGTGTGACA	rs3813593
		rs17855192
PPP2R3B_Ex13F	AGCAGCCTGTGTTTCCCG	rs1133523
PPP2R3B_Ex13R	AGGCTCCTGTCCAGGACTGA	
mXIST_rs31391216F	CAGGAACCATTCTTGCCTACCTT	
mXIST_rs31391216R	TTCTGGCAGTTGGTCCTGTACA	
mXist_rs31391216_cDNA_F	GAGTTGGATTGCCACCTTTTAC	
mXist_rs31391216_cDNA_R	CAGACTCCAAATTCATCCTGTAGTC	

qPCR

Primer	**Sequence (5'-3')**
HPRT1_267F	AGGGTGTTTATTCCTCATGGAC
HPRT1_367R	CACAGAGGGCTACAATGTGAT
CTCF-1056F	GTGTTCCATGTGCGATTACG
CTCF-1195R	TCATGTGCCTTTTCAGCTTG
BORIS-1590F	CGCTCACTTCAGGAAATACCAC
BORIS-1696R	CCGAATGTCTGTGCAGGTTAAT
TBP_562F	GAAACGCCGAATATAATCCCAA
TBP_672R	GGCTCCTGTGCACACCATT
SAGE1_26F	TCAACCAACTCCACCTGAAGA
SAGE1_127R	GCCCTGTTGCAACCAGATT

PIM2_155F	CACAGATCGACTCCAGGTGG
PIM2_239R	CATGTGACTGAGTCTGACAAGGG
MST4_922F	GCTCTGATTCGGAATCTACCAG
MST4_1026R	TGCCCCATTCTGTACTTTCTTT
PPP2R3B_310F	CAGGGACGAGAGTAGTTCAGACA
PPP2R3B_449R	CTCTCGATCTTGCTGATGACG
RAP2C-187F	AGTTTGCCTCCATGAGAGATCT
RAP2C-290R	CTCATTGGCTTGATATCCTGAA
FGF13-431F	ACCACTGAAAGTGGCCATGTA
FGF13-542R	TTCAGCACGCCAGAGACA
RBBP7-397F	CGCAGAATCCTCACATCATT
RBBP7-493R	CTCCACTTGGGTCTGGTTTAG
ZFX-1079F	GGCAGCAGCTTATGGTAATAATT
ZFX-1170R	GAGGCCAGCAGACTCATCTAT
ZFX-46F	TTGATGCAACAGGAGCTGAT
ZFX-146R	ATGTCTGAATCCACAACATCTGA
ZBTB33-217F	GGGAAGCAGTGTCTGTTGACT
ZBTB33-334R	CCATGCCCTTTCCTCTTTC
KDM5C-763F	AGACTCTGCGGAAGAAAGATAAG
KDM5C-852R	GAAGGTCTTAGGCGATGTTGA
BMX-803F	GAATCACACCACCTCAAAGATTT
BMX-903R	GATGTTACCAGCAAACCAGTCA
BEX2-526F	AAGGATAGGCCCAGGAGTAATG
BEX2-671R	ATTCAAAGGTAGGGCCAAGG
CTNNA2_1031F	CCTGCTCAGCGAGTACATGA
CTNNA2_1147R	CTTTCCGAAGCTGTCTCCTTAG
HELT_2F	GTCAGACAAGCTCAAGGAACG
HELT_137R	GAACTCTGCTTCGCCAAGG
HERC6_951F	GACACAAGCAAGCCAACTCAT
HERC6_1036R	CATCCACGAAGTCTTCAGCA
IKBKE_611F	CTGGAGCATTGGAGTGACCT
IKBKE_711R	GATCCGGTACATGATCTCCTTG
MCTP1_1063F	CCACAGATGTGACCCTTACTCTG
MCTP1_1185R	CATTAGCATTGTCACATCCCTG
STON1_2045F	CTCAAGGACAGAGGTCAGGTCT
STON1_2143R	CTATTTCAACCTGGATGTTGTAGC
ABCA5_458F	GGATTCAAGAGCTGGCTGTT
ABCA5_549R	GGCAGCATCTATGGATGCTT
ECHDC3_182F	GAGCAATCCCAAGAAGAGGAA
ECHDC3_299R	CCCTCAGCCGAGATGATAAT
PMM1_458F	CTCCGAACTGGACAAGAAAGAG

PMM1_577R	GGAAGACGTCAAAGCTGATCA
S100A10_84F	GAGGACCTGAGAGTACTCATGGA
S100A10_189R	TCTACACTGGTCCAGGTCCTTC
DKK1_437F	CCGAGGAGAAATTGAGGAAAC
DKK1_557R	CCTTCTTGTCCTTTGGTGTGATA
FBN1_433F	GTGGACAACCTGTTTGTGAAAG
FBN1_553R	GGCCTGTCCTGTAATCTCTTTC
INHBA_374F	GTTTGCCGAGTCAGGAACAG
INHBA_524R	AGGCGGATGGTGACTTTG
VEGFC_1131F	CACCAAACATGCAGCTGTTAC
VEGFC_1228R	GGACACAACGACACACTTCTTC
GYG2_696F	GCCTTCAAGCAATTCGGTT
GYG2_788R	GAGCCACTCTGTGGATTGTACTT
LZTS1_238F	CAGATTACACGGCACTGTCC
LZTS1_353R	TCGGAGCCCATTTCTAGCT
MLANA_155F	TGGATACAGAGCCTTGATGGAT
MLANA_299R	GCATTGGGAACCACAGGTT
TFAP2A_871F	CACTAGTAGAGGGAGAAGCTGTCC
TFAP2A_979R	CGGAATGTTGTCGGTTGAG
TGFB1_848F	CTATTGCTTCAGCTCCACGG
TGFB1_961R	AGTTGGCATGGTAGCCCTT
SPARC_216F	CAGAACCACCACTGCAAACA
SPARC_340R	TGCTGCACACCTTCTCAAAC
47S_pre-rRNA-F	GCTGACACGCTGTCCTCTGG
47S_pre-rRNA-R	GAGAACGCCTGACACGCACG
45S_pre-rRNA-F	GCCTTCTCTAGCGATCTGAGAG
45S_pre-rRNA-R	CCATAACGGAGGCAGAGACA
hs_pre18S_F	ACCCACCCTCGGTGAGA
hs_pre18S_R	CAAGGCACGCCTCTCAGAT
hs_18Sreg_F	CTCAACACGGGAAACCTCAC
hs_18Sreg_R	CGCTCCACCAACTAAGAACG

ChIP-qPCR

Primer	**Sequence (5'-3')**
CLDN1-ChIP-F-P01	ACT GAG ACG CAG AAC CGC
CLDN1-ChIP-R-P01	GTG AGC CGC CCT GAA AC
CLDN1-ChIP-F-N01	TGA AGC TCC TTG GAG TCT TTG
CLDN1-ChIP-R-N01	CAA GGT TAC TTT AAG TGG CAG TAA TG
PDL1-ChIP-F-P01	CAG TTC TGC GCA GCT TCC
PDL1-ChIP-R-P01	GTG CCC AGC GAG CTA GC
GDF6-ChIP-F-P01	AGC ACA GCT GCT TCC AGG
GDF6-ChIP-R-P01	TCC CTT GCC TTT CGC C

CDKN2C-ChIP-F-P01	CTC TCC CAG CAG CGG AG
CDKN2C-ChIP-R-P01	TTG TTC ACG GTG GTG ACT TG
RPL10-ChIP-F-P01	CAG TCC GGG AAA AAC GG
RPL10-ChIP-R-P01	TCC ACC TCC AAG CCG AG
PAX6-ChIP-F-P01	GCG AGA GTG ACT GTC CTT CC
PAX6-ChIP-R-P01	AGG AAA ACC GAA GGG AAC C
PAX6-ChIP-F-N01	GAT TCT CCA CCG CAC AGT G
PAX6-ChIP-R-N01	TTC CCC TAC TTC AAG GCC TC

Bisulfite sequencing

Primer	**Sequence (5'-3')**
Pax6-M_F	GATAGGTTTGTTTTTTTTTTTTAGATTT
Pax6-M_R	AACAAATAACACCACCCAAACTAA

ATAC-seq

Primer	**Sequence (5'-3')**
KAT6B_213F_ATAC	CAGCGGTCTGTTGATTGGTAT
KAT6B_381R_ATAC	GCCTACGGGTCTGTCTGTCA
GAPDH_130F_ATAC	ACGTAGCTCAGGCCTCAAGA
GAPDH_237R_ATAC	GGCTCAATTTATAGAAACCGGG
ATACNC_RHO-F	TGGGTGGTGTCATCTGGTAA
ATACNC_RHO-R	GGATGGAATGGATCAGATGG
ATACNC_SLC22A3-F	GGAGAGGGTGGACAGATTGA
ATACNC_SLC22A3-R	TCAGCCTTGCTGCTACAGTG
ND1-mtDNA_F	CCT AAA ACC CGC CAC ATC T
ND1-mtDNA_R	GAG CGA TGG TGA GAG CTA AGG
mtDNA-ATAC_F	CTT CTG GCC ACA GCA CTT AAA C
mtDNA-ATAC_R	AAT CTG GTT AGG CTG GTG TTA GG
Ad1_noMX	AAT GAT ACG GCG ACC ACC GAG ATC TAC ACT CGT CGG CAG CGT CAG ATG TG
Ad2.3_AGGCAGAA	CAA GCA GAA GAC GGC ATA CGA GAT TTC TGC CTG TCT CGT GGG CTC GGA GAT GT
Ad2.4_TCCTGAGC	CAA GCA GAA GAC GGC ATA CGA GAT GCT CAG GAG TCT CGT GGG CTC GGA GAT GT
Ad2.5_GGACTCCT	CAA GCA GAA GAC GGC ATA CGA GAT AGG AGT CCG TCT CGT GGG CTC GGA GAT GT
Ad2.6_TAGGCATG	CAA GCA GAA GAC GGC ATA CGA GAT CAT GCC TAG TCT CGT GGG CTC GGA GAT GT

Appendix II. Putative SNPs between the C57BL/6J and FVB/NJ mouse genetic background according to the MGI database

dbSNP Build 137

SNP ID	Map Position (GRCm38)	Gene: dbSNP Function Class	C57BL /6J	Allele Summary
rs4232458	ChrX:6081673	AU022751 : Coding-Synonymous, Gm26618 : within coordinates of	T	C/T
rs4232457	ChrX:6081787	AU022751 : Coding-Synonymous, Gm26618 : within coordinates of	C	C/T
rs4232456	ChrX:6081892	AU022751 : Coding-NonSynonymous, Gm26618 : within coordinates of	A	A/T
rs4232455	ChrX:6081904	AU022751 : Coding-Synonymous, Gm26618 : within coordinates of	T	C/T
rs4232454	ChrX:6081927	AU022751 : Coding-NonSynonymous, Gm26618 : within coordinates of	T	C/T
rs4232461	ChrX:6314288	Bmp15 : mRNA-UTR	C	C/T
rs3672113	ChrX:6629220	Shroom4 : Coding-NonSynonymous	T	C/T
rs4135915	ChrX:6629240	Shroom4 : Coding-NonSynonymous	A	A/C
rs4232452	ChrX:8062418	Gm14820 : within coordinates of, Suv39h1 : mRNA-UTR	G	A/G
rs4232451	ChrX:8062456	Gm14820 : within coordinates of, Suv39h1 : mRNA-UTR	G	G/T
rs4232450	ChrX:8062613	Gm14820 : within coordinates of, Suv39h1 : mRNA-UTR	A	A/C
rs33390776	ChrX:8240033	Ftsj1 : mRNA-UTR	C	C/T
rs31057885	ChrX:8276811	Slc38a5 : Coding-Synonymous	A	A/G
rs33866966	ChrX:8454416	Ssxb2 : mRNA-UTR	C	C/T
rs33872067	ChrX:8454447	Ssxb2 : mRNA-UTR	C	A/C
rs31385959	ChrX:9310568	Xk : mRNA-UTR	A	A/G
rs31060723	ChrX:9435560	Cybb : mRNA-UTR	T	C/T
rs31257704	ChrX:9436019	Cybb : mRNA-UTR	T	G/T
rs3684547	ChrX:9436145	Cybb : mRNA-UTR	G	A/G
rs33444507	ChrX:9436716	Cybb : mRNA-UTR	T	C/T
rs4232462	ChrX:9436981	Cybb : mRNA-UTR	A	A/G
rs4232463	ChrX:9437062	Cybb : mRNA-UTR	A	A/C
rs4232464	ChrX:9437245	Cybb : mRNA-UTR	G	A/G
rs4232465	ChrX:9437262	Cybb : mRNA-UTR	C	C/T
rs4232466	ChrX:9654354	Dynlt3 : mRNA-UTR	T	C/T
rs4232467	ChrX:9654456	Dynlt3 : mRNA-UTR	C	C/T
rs4232468	ChrX:9654484	Dynlt3 : mRNA-UTR	C	C/T
rs4232469	ChrX:9654525	Dynlt3 : mRNA-UTR	C	C/T
rs4232470	ChrX:9654566	Dynlt3 : mRNA-UTR	C	A/C
rs4232471	ChrX:9654581	Dynlt3 : mRNA-UTR	T	C/T

rs4232472	ChrX:9654585	Dynlt3 : mRNA-UTR	G	A/G
rs33459047	ChrX:9655269	Dynlt3 : mRNA-UTR	C	C/T
rs31238837	ChrX:9963664	Sytl5 : Coding-NonSynonymous	G	A/G
rs30730568	ChrX:10038935	Srpx : Coding-Synonymous	C	A/C
rs30580058	ChrX:10039060	Srpx : Coding-NonSynonymous	G	C/G
rs30925594	ChrX:10039061	Srpx : Coding-Synonymous	G	G/T
rs33873425	ChrX:10158483	Rpgr : mRNA-UTR	A	A/G
rs33492631	ChrX:10206224	Rpgr : Coding-NonSynonymous	T	C/T
rs31120189	ChrX:10207400	Rpgr : mRNA-UTR, Rpgr : Intron	T	C/T
rs31270554	ChrX:10207403	Rpgr : Intron, Rpgr : mRNA-UTR	T	C/T
rs6256688	ChrX:10596248	Tspan7 : mRNA-UTR	C	C/G
rs4232475	ChrX:10596429	Tspan7 : mRNA-UTR	A	A/T
rs13483717	ChrX:10718421	Cpgi22605 : within coordinates of, Gm14493 : within coordinates of, Mid1ip1 : Coding-Synonymous	G	A/G
rs4223179	ChrX:12652320	Cpgi22627 : within coordinates of, Gm4984 : Coding-Synonymous	G	A/G
rs4223180	ChrX:12652592	Gm4984 : Coding-NonSynonymous	C	C/G
rs4223181	ChrX:12652642	Gm4984 : Coding-NonSynonymous	A	A/G
rs6407077	ChrX:17572166	Fundc1 : Coding-NonSynonymous	G	A/G
rs33879491	ChrX:21061933	Zfp182 : mRNA-UTR	T	C/T
rs13483723	ChrX:21061954	Zfp182 : mRNA-UTR	C	C/T
rs4232488	ChrX:21488245	Agtr2 : mRNA-UTR	A	A/T
rs4232487	ChrX:21488368	Agtr2 : mRNA-UTR	C	C/T
rs4232486	ChrX:21488431	Agtr2 : mRNA-UTR	T	C/T
rs4232485	ChrX:21488443	Agtr2 : mRNA-UTR	G	A/G
rs4232484	ChrX:21488514	Agtr2 : mRNA-UTR	A	A/T
rs4232483	ChrX:21488517	Agtr2 : mRNA-UTR	T	C/T
rs30855408	ChrX:26043145	Gm5168 : Coding-NonSynonymous	T	A/T
rs13483724	ChrX:36008085	Dock11 : Coding-Synonymous	C	A/C
rs6350693	ChrX:36126664	Il13ra1 : Coding-Synonymous	G	A/G
rs33870231	ChrX:36743822	Cpgi22664 : within coordinates of, Slc25a43 : Coding-NonSynonymous	C	C/T
rs8238456	ChrX:36798689	Fnld : within coordinates of, Slc25a5 : mRNA-UTR	T	A/T
rs8238457	ChrX:36798724	Fnld : within coordinates of, Slc25a5 : mRNA-UTR	C	C/T
rs6378292	ChrX:38772816	6030498E09Rik : mRNA-UTR, Fnld : within coordinates of	A	A/G
rs29944664	ChrX:41796488	Fnld : within coordinates of, Thoc2 : mRNA-UTR	T	A/T
rs31309753	ChrX:42276376	Fnld : within coordinates of, Stag2 : mRNA-UTR	A	A/G
rs31300344	ChrX:42276571	Fnld : within coordinates of, Stag2 : mRNA-UTR	A	A/G

rs6386855	ChrX:42535954	Fnld : within coordinates of, Tenm1 : Coding-NonSynonymous	A	A/G
rs6282846	ChrX:44790175	Cpgi22687 : within coordinates of, Dcaf12l1 : Locus-Region, Dcaf12l1 : mRNA-UTR, Fnld : within coordinates of	C	C/G
rs6310503	ChrX:46329245	4930515L19Rik : Locus-Region, Actrt1 : Coding-NonSynonymous, Fnld : within coordinates of	A	A/G
rs6290295	ChrX:47884066	Fnld : within coordinates of, Smarca1 : Coding-NonSynonymous	A	A/T
rs31161735	ChrX:48025806	Apln : mRNA-UTR, Fnld : within coordinates of	T	G/T
rs8276051	ChrX:48109074	Fnld : within coordinates of, Xpnpep2 : mRNA-UTR	T	C/T
rs8276050	ChrX:48109077	Fnld : within coordinates of, Xpnpep2 : mRNA-UTR	T	C/T
rs8276049	ChrX:48109089	Fnld : within coordinates of, Xpnpep2 : mRNA-UTR	T	G/T
rs8276099	ChrX:48116984	Fnld : within coordinates of, Xpnpep2 : Coding-NonSynonymous	C	C/G
rs8276114	ChrX:48117298	Fnld : within coordinates of, Xpnpep2 : Coding-NonSynonymous	C	C/T
rs8276113	ChrX:48117300	Fnld : within coordinates of, Xpnpep2 : Coding-Synonymous	G	A/G
rs8276211	ChrX:48118350	Fnld : within coordinates of, Xpnpep2 : Coding-NonSynonymous	A	A/T
rs8276213	ChrX:48118447	Fnld : within coordinates of, Xpnpep2 : Coding-Synonymous	C	C/T
rs8276274	ChrX:48121803	Fnld : within coordinates of, Xpnpep2 : Coding-NonSynonymous	A	A/G
rs8276276	ChrX:48121848	Fnld : within coordinates of, Xpnpep2 : Coding-NonSynonymous	C	C/T
rs8276374	ChrX:48122555	Fnld : within coordinates of, Xpnpep2 : Coding-NonSynonymous	A	A/G
rs8276375	ChrX:48122597	Fnld : within coordinates of, Xpnpep2 : Coding-NonSynonymous	A	A/G
rs8276376	ChrX:48122605	Fnld : within coordinates of, Xpnpep2 : Coding-Synonymous	C	C/T
rs8276377	ChrX:48122606	Fnld : within coordinates of, Xpnpep2 : Coding-NonSynonymous	A	A/G
rs8276384	ChrX:48122910	Fnld : within coordinates of, Xpnpep2 : Coding-NonSynonymous	T	C/T
rs8276385	ChrX:48122998	Fnld : within coordinates of, Xpnpep2 : Coding-NonSynonymous	G	C/G

rs8276386	ChrX:48123035	Fnld : within coordinates of, Xpnpep2 : Coding-Synonymous	A	A/G
rs8276349	ChrX:48123061	Fnld : within coordinates of, Xpnpep2 : Coding-NonSynonymous	C	C/G
rs8276350	ChrX:48125354	Fnld : within coordinates of, Xpnpep2 : Coding-NonSynonymous	C	A/C
rs8276325	ChrX:48130055	Fnld : within coordinates of, Xpnpep2 : Coding-Synonymous	T	C/T
rs8276326	ChrX:48130064	Fnld : within coordinates of, Xpnpep2 : Coding-Synonymous	G	C/G
rs16797337	ChrX:48130689	Fnld : within coordinates of, Xpnpep2 : Coding-NonSynonymous	T	C/T
rs8276389	ChrX:48130729	Fnld : within coordinates of, Xpnpep2 : Coding-NonSynonymous	T	C/T
rs8276410	ChrX:48135788	Fnld : within coordinates of, Xpnpep2 : Coding-Synonymous	G	A/G
rs8276411	ChrX:48135832	Fnld : within coordinates of, Xpnpep2 : Coding-NonSynonymous	A	A/G
rs8276412	ChrX:48136013	Fnld : within coordinates of, Xpnpep2 : mRNA-UTR	C	A/C
rs8276413	ChrX:48136072	Fnld : within coordinates of, Xpnpep2 : mRNA-UTR	G	A/G
rs8276414	ChrX:48136073	Fnld : within coordinates of, Xpnpep2 : mRNA-UTR	T	C/T
rs8276415	ChrX:48136076	Fnld : within coordinates of, Xpnpep2 : mRNA-UTR	G	A/G
rs8276416	ChrX:48136099	Fnld : within coordinates of, Xpnpep2 : mRNA-UTR	A	-/A
rs8276417	ChrX:48136100	Fnld : within coordinates of, Xpnpep2 : mRNA-UTR	T	-/T
rs8276418	ChrX:48136101	Fnld : within coordinates of, Xpnpep2 : mRNA-UTR	A	-/A
rs8276419	ChrX:48136102	Fnld : within coordinates of, Xpnpep2 : mRNA-UTR	C	-/C
rs8276420	ChrX:48136229	Fnld : within coordinates of, Xpnpep2 : mRNA-UTR	A	A/C
rs8276421	ChrX:48136370	Fnld : within coordinates of, Xpnpep2 : mRNA-UTR	T	C/T
rs8276422	ChrX:48136453	Fnld : within coordinates of, Xpnpep2 : mRNA-UTR	C	C/T
rs8276423	ChrX:48136471	Fnld : within coordinates of, Xpnpep2 : mRNA-UTR	G	A/G
rs8276424	ChrX:48136548	Fnld : within coordinates of, Xpnpep2 : mRNA-UTR	G	A/G

rs8276425	ChrX:48136585	Fnld : within coordinates of, Xpnpep2 : mRNA-UTR	C	C/T
rs8276426	ChrX:48136709	Fnld : within coordinates of, Xpnpep2 : mRNA-UTR	G	C/G
rs8276427	ChrX:48136752	Fnld : within coordinates of, Xpnpep2 : mRNA-UTR	A	A/G
rs8276428	ChrX:48136757	Fnld : within coordinates of, Xpnpep2 : mRNA-UTR	C	C/T
rs8276429	ChrX:48136760	Fnld : within coordinates of, Xpnpep2 : mRNA-UTR	T	G/T
rs8276430	ChrX:48136766	Fnld : within coordinates of, Xpnpep2 : mRNA-UTR	T	C/T
rs8276431	ChrX:48136787	Fnld : within coordinates of, Xpnpep2 : mRNA-UTR	A	A/T
rs8276432	ChrX:48136857	Fnld : within coordinates of, Xpnpep2 : mRNA-UTR	T	C/T
rs8276433	ChrX:48136867	Fnld : within coordinates of, Xpnpep2 : mRNA-UTR	C	C/T
rs30156027	ChrX:48160757	Fnld : within coordinates of, Sash3 : mRNA-UTR	G	A/G
rs30156030	ChrX:48160869	Fnld : within coordinates of, Sash3 : mRNA-UTR	A	A/G
rs30156939	ChrX:48161458	Fnld : within coordinates of, Sash3 : mRNA-UTR	C	C/T
rs31116843	ChrX:48406366	Bcorl1 : mRNA-UTR, Fnld : within coordinates of	G	G/T
rs33873159	ChrX:48412538	Elf4 : mRNA-UTR, Fnld : within coordinates of	T	C/T
rs13483739	ChrX:48416800	Elf4 : Coding-NonSynonymous, Fnld : within coordinates of	T	C/T
rs30176917	ChrX:48542466	Fnld : within coordinates of, Zfp280c : mRNA-UTR	A	A/G
rs31336254	ChrX:49497218	Arhgap36 : Coding-NonSynonymous	G	A/G
rs31214442	ChrX:49498632	Arhgap36 : Coding-NonSynonymous	T	A/T
rs31171729	ChrX:49499859	Arhgap36 : mRNA-UTR	A	A/G
rs30226154	ChrX:49727525	Olfr1321 : Coding-Synonymous	C	C/T
rs6299349	ChrX:49886102	Olfr1322 : Coding-Synonymous	T	C/T
rs31075039	ChrX:50573246	Firre : Noncoding-Transcript-Variant, Gm35612 : within coordinates of	T	C/T
rs33872269	ChrX:50573261	Firre : Noncoding-Transcript-Variant, Gm35612 : within coordinates of	G	C/G
rs31413341	ChrX:50573280	Firre : Noncoding-Transcript-Variant, Gm35612 : within coordinates of	T	A/T
rs31158503	ChrX:50573281	Firre : Noncoding-Transcript-Variant, Gm35612 : within coordinates of	C	C/G

rs33871613	ChrX:50893926	Frmd7 : mRNA-UTR	C	C/T
rs31114590	ChrX:50894865	Frmd7 : mRNA-UTR	T	C/T
rs31300327	ChrX:51118200	Mbnl3 : mRNA-UTR	A	A/C
rs31079278	ChrX:51118867	Mbnl3 : mRNA-UTR	A	A/G
rs30282288	ChrX:51389074	Hs6st2 : mRNA-UTR	A	A/G
rs31325877	ChrX:51969727	1700080O16Rik : Coding-Synonymous	G	C/G
rs31292784	ChrX:52772698	Gm14586 : Noncoding-Transcript-Variant	A	A/C
rs30422414	ChrX:52955497	Phf6 : mRNA-UTR	G	C/G
rs30437299	ChrX:53244368	Fam122b : mRNA-UTR	G	C/G
rs33880297	ChrX:53282571	Fam122c : Coding-NonSynonymous	G	A/G
rs13483759	ChrX:53775441	Zfp36l3 : Coding-NonSynonymous	T	C/T
rs6270558	ChrX:56485340	Ddx26b : Coding-Synonymous	T	C/T
rs6200595	ChrX:56501787	Ddx26b : Coding-NonSynonymous	G	G/T
rs6200613	ChrX:56501801	Ddx26b : Coding-Synonymous	T	A/T
rs33876751	ChrX:56779794	Fhl1 : Intron, Fhl1 : Coding-NonSynonymous	A	A/G
rs29060172	ChrX:57101932	Vgll1 : Coding-NonSynonymous	A	A/G
rs13475263	ChrX:57233200	Arhgef6 : mRNA-UTR	A	A/C
rs29060140	ChrX:58035069	Zic3 : mRNA-UTR	A	A/G
rs16814457	ChrX:60028515	F9 : Coding-Synonymous	A	A/G
rs16814469	ChrX:60029114	F9 : Coding-Synonymous	G	C/G
rs16814516	ChrX:60029189	F9 : Coding-Synonymous	A	A/G
rs16814515	ChrX:60029190	F9 : Coding-Synonymous	T	G/T
rs16814514	ChrX:60029473	F9 : mRNA-UTR	C	C/T
rs16814513	ChrX:60029672	F9 : mRNA-UTR	A	A/G
rs16814473	ChrX:60029781	F9 : mRNA-UTR	A	A/C
rs16814472	ChrX:60029791	F9 : mRNA-UTR	A	A/G
rs31119550	ChrX:60891647	Sox3 : mRNA-UTR	A	A/C
rs29051076	ChrX:66654243	Slitrk2 : Coding-Synonymous, Stol : within coordinates of	A	A/G
rs29051075	ChrX:66654921	Slitrk2 : Coding-Synonymous, Stol : within coordinates of	C	C/T
rs29050690	ChrX:66659942	Slitrk2 : mRNA-UTR, Stol : within coordinates of	C	C/G
rs6371072	ChrX:66660875	Slitrk2 : mRNA-UTR, Stol : within coordinates of	T	C/T
rs29051494	ChrX:66799228	Mir463 : Noncoding-Transcript-Variant, Stol : within coordinates of	A	A/G
rs29047498	ChrX:69544423	Aff2 : Coding-Synonymous, Stol : within coordinates of	T	C/T
rs13483811	ChrX:69544913	Aff2 : Coding-NonSynonymous, Stol : within coordinates of	A	A/G
rs29047497	ChrX:69545029	Aff2 : Coding-Synonymous, Stol : within coordinates of	T	C/T

rs29044836	ChrX:69830745	Aff2 : Coding-NonSynonymous, Stol : within coordinates of	C	C/T
rs13483812	ChrX:69834798	Aff2 : Coding-NonSynonymous, Stol : within coordinates of	G	A/G
rs6161348	ChrX:69989050	1700111N16Rik : Noncoding-Transcript-Variant, Stol : within coordinates of	G	A/G
rs30570993	ChrX:70386121	1110012L19Rik : mRNA-UTR, Stol : within coordinates of	A	A/C
rs31340548	ChrX:70461013	BC023829 : mRNA-UTR, Stol : within coordinates of	G	C/G
rs29043047	ChrX:71197548	Gm14730 : Noncoding-Transcript-Variant, Stol : within coordinates of	A	A/T
rs29043046	ChrX:71197577	Gm14730 : Noncoding-Transcript-Variant, Stol : within coordinates of	A	A/T
rs6151989	ChrX:71421042	Cd99l2 : mRNA-UTR, Stol : within coordinates of	i	-/AA/ AAAAAA
rs4232507	ChrX:71421085	Cd99l2 : mRNA-UTR, Stol : within coordinates of	G	C/G
rs3161045	ChrX:71421158	Cd99l2 : mRNA-UTR, Stol : within coordinates of	A	A/T
rs4232506	ChrX:71421370	Cd99l2 : mRNA-UTR, Stol : within coordinates of	G	G/T
rs4232505	ChrX:71421431	Cd99l2 : mRNA-UTR, Stol : within coordinates of	C	C/T
rs4232504	ChrX:71421434	Cd99l2 : mRNA-UTR, Stol : within coordinates of	C	C/T
rs3161046	ChrX:71421675	Cd99l2 : mRNA-UTR, Stol : within coordinates of	A	A/G
rs3157108	ChrX:71422216	Cd99l2 : mRNA-UTR, Stol : within coordinates of	G	A/G
rs3161047	ChrX:71422267	Cd99l2 : mRNA-UTR, Stol : within coordinates of	C	C/T
rs31261214	ChrX:71939593	Gm1141 : Noncoding-Transcript-Variant, Stol : within coordinates of	G	A/G
rs31233614	ChrX:71968425	Prrg3 : mRNA-UTR, Stol : within coordinates of	A	A/T
rs31384593	ChrX:71968623	Prrg3 : mRNA-UTR, Stol : within coordinates of	A	A/T
rs33870377	ChrX:71988653	Fate1 : Noncoding-Transcript-Variant, Stol : within coordinates of	T	C/T
rs29038226	ChrX:71991876	Cnga2 : mRNA-UTR, Stol : within coordinates of	G	A/G
rs33877614	ChrX:72222162	Magea4 : Coding-NonSynonymous, Stol : within coordinates of	A	A/T

rs33877489	ChrX:72222733	Magea4 : Coding-NonSynonymous, Stol : within coordinates of	G	C/G
rs31345732	ChrX:72222763	Magea4 : Coding-Synonymous, Stol : within coordinates of	T	C/T
rs33875390	ChrX:72222892	Magea4 : Coding-Synonymous, Stol : within coordinates of	C	C/T
rs3161142	ChrX:72257804	Gabre : Coding-NonSynonymous, Stol : within coordinates of	C	C/T
rs3161127	ChrX:72270448	Gabre : Coding-Synonymous, Stol : within coordinates of	T	G/T
rs29041437	ChrX:72382557	Magea10 : Coding-NonSynonymous, Stol : within coordinates of	A	A/G
rs4232510	ChrX:72913580	Cetn2 : mRNA-UTR, Stol : within coordinates of	G	A/G
rs4232509	ChrX:72913581	Cetn2 : mRNA-UTR, Stol : within coordinates of	G	A/G
rs4232508	ChrX:72913964	Cetn2 : mRNA-UTR, Stol : within coordinates of	G	A/G
rs13473968	ChrX:72957627	Nsdhl : mRNA-UTR, Stol : within coordinates of	C	C/T
rs29038928	ChrX:72987401	Cpgi22764 : within coordinates of, Stol : within coordinates of, Zfp185 : mRNA-UTR	A	A/C
rs33870890	ChrX:73012574	Stol : within coordinates of, Zfp185 : Coding-NonSynonymous	A	A/T
rs13483822	ChrX:73015776	Stol : within coordinates of, Zfp185 : Coding-NonSynonymous	A	A/G
rs31256282	ChrX:73018414	Stol : within coordinates of, Zfp185 : Coding-NonSynonymous	G	C/G
rs29040658	ChrX:73030003	Stol : within coordinates of, Zfp185 : mRNA-UTR	A	A/G
rs13467684	ChrX:73031144	Stol : within coordinates of, Zfp185 : mRNA-UTR	C	C/G
rs8266058	ChrX:73096992	Stol : within coordinates of, Xlr3a : mRNA-UTR	C	C/T
rs8266057	ChrX:73097044	Stol : within coordinates of, Xlr3a : mRNA-UTR	G	A/G
rs8266056	ChrX:73097066	Stol : within coordinates of, Xlr3a : mRNA-UTR	G	A/G
rs29037634	ChrX:73353417	Stol : within coordinates of, Zfp275 : Coding-Synonymous	C	C/T
rs4232514	ChrX:73358347	Stol : within coordinates of, Zfp275 : mRNA-UTR	A	A/C
rs4232513	ChrX:73358348	Stol : within coordinates of, Zfp275 : mRNA-UTR	A	A/T

rs4232512	ChrX:73358392	Stol : within coordinates of, Zfp275 : mRNA-UTR	G	A/G
rs4232511	ChrX:73358516	Stol : within coordinates of, Zfp275 : mRNA-UTR	T	C/T
rs6202124	ChrX:73411133	Cpgi22768 : within coordinates of, Gm26726 : within coordinates of, Stol : within coordinates of, Zfp92 : mRNA-UTR	A	A/C
rs33876282	ChrX:73640629	Dusp9 : Coding-Synonymous, Stol : within coordinates of	G	A/G
rs3166614	ChrX:73682502	Slc6a8 : mRNA-UTR, Stol : within coordinates of	A	A/G
rs8248894	ChrX:73729117	Abcd1 : Coding-Synonymous, Stol : within coordinates of	T	C/T
rs8248912	ChrX:73737499	Abcd1 : Coding-NonSynonymous, Stol : within coordinates of	G	C/G
rs8248913	ChrX:73737518	Abcd1 : Coding-NonSynonymous, Stol : within coordinates of	G	A/G
rs8248914	ChrX:73737519	Abcd1 : Coding-Synonymous, Stol : within coordinates of	A	A/C
rs8248915	ChrX:73737520	Abcd1 : Coding-NonSynonymous, Stol : within coordinates of	G	A/G
rs31200739	ChrX:73762804	Plxnb3 : Coding-Synonymous, Stol : within coordinates of	G	A/G
rs8245177	ChrX:73862969	L1cam : Coding-Synonymous, Stol : within coordinates of	G	A/G
rs8245193	ChrX:73869854	L1cam : mRNA-UTR, Stol : within coordinates of	A	A/G
rs29037445	ChrX:74177399	Stol : within coordinates of, Tktl1 : Coding-NonSynonymous	A	A/G
rs29037404	ChrX:74197051	Stol : within coordinates of, Tktl1 : Coding-Synonymous	G	A/G
rs13474284	ChrX:74224810	Flna : Coding-Synonymous, Stol : within coordinates of	G	C/G
rs3166685	ChrX:74227502	Flna : Coding-NonSynonymous, Stol : within coordinates of	A	A/G
rs13474281	ChrX:74227949	Flna : Coding-Synonymous, Stol : within coordinates of	A	A/G
rs4232518	ChrX:74273270	Dnase1l1 : mRNA-UTR, Rpl10 : Intron, Stol : within coordinates of	C	C/T
rs4232519	ChrX:74273439	Dnase1l1 : mRNA-UTR, Rpl10 : Intron, Stol : within coordinates of	G	A/G
rs13463859	ChrX:74289448	Stol : within coordinates of, Taz : Noncoding-Transcript-Variant, Taz : mRNA-UTR	C	C/T

rs4232520	ChrX:74289880	Stol : within coordinates of, Taz : mRNA-UTR, Taz : Noncoding-Transcript-Variant	C	C/G
rs13461164	ChrX:74305318	Cpgi22795 : within coordinates of, Gdi1 : mRNA-UTR, Stol : within coordinates of	A	A/G
rs29035837	ChrX:74306865	Gdi1 : Coding-NonSynonymous, Stol : within coordinates of	A	A/G
rs31120470	ChrX:74311209	Fam50a : Locus-Region, Gdi1 : mRNA-UTR, Stol : within coordinates of	T	C/T
rs31072145	ChrX:74311223	Fam50a : Locus-Region, Gdi1 : mRNA-UTR, Stol : within coordinates of	C	C/T
rs31126977	ChrX:74311225	Fam50a : Locus-Region, Gdi1 : mRNA-UTR, Stol : within coordinates of	C	C/T
rs33876511	ChrX:74311231	Fam50a : Locus-Region, Gdi1 : mRNA-UTR, Stol : within coordinates of	C	C/T
rs31416952	ChrX:74311245	Fam50a : Locus-Region, Gdi1 : mRNA-UTR, Stol : within coordinates of	C	C/T
rs13473170	ChrX:74317666	Fam50a : Coding-Synonymous, Stol : within coordinates of	G	A/G
rs31153326	ChrX:74370862	Gm44504 : within coordinates of, Slc10a3 : Coding-NonSynonymous, Stol : within coordinates of	T	C/T
rs31211824	ChrX:74370893	Gm44504 : within coordinates of, Slc10a3 : mRNA-UTR, Stol : within coordinates of	T	G/T
rs31310267	ChrX:74392952	Fam3a : mRNA-UTR, Ikbkg : Locus-Region, Stol : within coordinates of	G	G/T
rs8238277	ChrX:74429374	G6pdx : Locus-Region, Ikbkg : Coding-Synonymous, Ikbkg : Locus-Region, Ikbkg : Intron, Stol : within coordinates of	A	A/G
rs8238275	ChrX:74429715	G6pdx : Locus-Region, Ikbkg : Intron, Ikbkg : mRNA-UTR, Stol : within coordinates of	C	C/T
rs33873163	ChrX:75084828	Gab3 : mRNA-UTR, Stol : within coordinates of	T	G/T
rs31348332	ChrX:75084831	Gab3 : mRNA-UTR, Stol : within coordinates of	T	G/T
rs31258799	ChrX:75109401	Dkc1 : mRNA-UTR, Mpp1 : Locus-Region, Stol : within coordinates of	G	C/G
rs31410567	ChrX:75109448	Dkc1 : mRNA-UTR, Mpp1 : Locus-Region, Stol : within coordinates of	T	C/T
rs6218208	ChrX:75453374	Brcc3 : Intron, Brcc3 : mRNA-UTR, Stol : within coordinates of	C	C/T
rs6218707	ChrX:75453424	Brcc3 : Intron, Brcc3 : mRNA-UTR, Stol : within coordinates of	G	A/G

rs6218774	ChrX:75453475	Brcc3 : mRNA-UTR, Brcc3 : Intron, Stol : within coordinates of	C	C/T
rs6219247	ChrX:75453521	Brcc3 : mRNA-UTR, Brcc3 : Intron, Stol : within coordinates of	A	A/T
rs6219276	ChrX:75453538	Brcc3 : mRNA-UTR, Brcc3 : Intron, Stol : within coordinates of	A	A/G
rs6233247	ChrX:75453798	Brcc3 : mRNA-UTR, Brcc3 : Intron, Stol : within coordinates of	G	A/G
rs6233252	ChrX:75453803	Brcc3 : mRNA-UTR, Brcc3 : Intron, Stol : within coordinates of	A	A/G
rs6233742	ChrX:75453858	Brcc3 : mRNA-UTR, Brcc3 : Intron, Stol : within coordinates of	T	C/T
rs31144974	ChrX:75572142	Rab39b : mRNA-UTR, Stol : within coordinates of	G	A/G
rs29035084	ChrX:75572388	Rab39b : mRNA-UTR, Stol : within coordinates of	A	A/G
rs13468851	ChrX:75785725	Pls3 : mRNA-UTR, Stol : within coordinates of	A	A/C
rs13468852	ChrX:75786136	Pls3 : mRNA-UTR, Stol : within coordinates of	A	A/T
rs6182380	ChrX:75799440	Pls3 : Coding-Synonymous, Stol : within coordinates of	C	C/T
rs31397961	ChrX:75805253	Pls3 : Coding-Synonymous, Stol : within coordinates of	A	A/G
rs31110528	ChrX:75874367	Pls3 : Intron, Pls3 : mRNA-UTR, Stol : within coordinates of	C	A/C
rs29034883	ChrX:75987297	Gm8692 : Noncoding-Transcript-Variant, Stol : within coordinates of	C	C/T
rs29034882	ChrX:75987397	Gm8692 : Noncoding-Transcript-Variant, Stol : within coordinates of	G	A/G
rs31371703	ChrX:75987455	Gm8692 : Noncoding-Transcript-Variant, Stol : within coordinates of	T	G/T
rs29034881	ChrX:75987475	Gm8692 : Noncoding-Transcript-Variant, Stol : within coordinates of	T	C/T
rs31330053	ChrX:76393351	Cldn34b4 : mRNA-UTR, Stol : within coordinates of	G	A/G
rs31350613	ChrX:76582865	Cldn34d : Coding-NonSynonymous, Stol : within coordinates of	T	G/T
rs31272979	ChrX:76583044	Cldn34d : Coding-NonSynonymous, Stol : within coordinates of	A	A/C
rs13483835	ChrX:76583396	Cldn34d : Coding-NonSynonymous, Stol : within coordinates of	C	C/T
rs29123688	ChrX:81419931	4930595M18Rik : Coding-Synonymous, Stol : within coordinates of	G	A/G

rs29123685	ChrX:81420560	4930595M18Rik : Coding-NonSynonymous, Stol : within coordinates of	T	C/T
rs29063711	ChrX:83487797	Dmd : Intron, exma : within coordinates of, Stol : within coordinates of, Tsga8 : mRNA-UTR	T	C/T
rs29063705	ChrX:83489911	Dmd : Coding-Synonymous, exma : within coordinates of, Stol : within coordinates of, Tsga8 : Locus-Region	C	C/T
rs33872868	ChrX:85704280	exma : within coordinates of, Gk : mRNA-UTR, Stol : within coordinates of	A	A/G
rs29065341	ChrX:85888930	5430427O19Rik : mRNA-UTR, exma : within coordinates of, Stol : within coordinates of	T	A/T
rs31041980	ChrX:85891165	5430427O19Rik : mRNA-UTR, exma : within coordinates of, Stol : within coordinates of	G	A/G
rs31362809	ChrX:85891266	5430427O19Rik : mRNA-UTR, exma : within coordinates of, Stol : within coordinates of	A	A/G
rs8240235	ChrX:86191957	Cpgi22818 : within coordinates of, exma : within coordinates of, Nr0b1 : Coding-NonSynonymous, Stol : within coordinates of	T	C/T
rs8240234	ChrX:86191979	Cpgi22818 : within coordinates of, exma : within coordinates of, Nr0b1 : Coding-NonSynonymous, Stol : within coordinates of	C	A/C
rs8240233	ChrX:86192016	Cpgi22818 : within coordinates of, exma : within coordinates of, Nr0b1 : Coding-Synonymous, Stol : within coordinates of	A	A/G
rs8240236	ChrX:86195790	exma : within coordinates of, Nr0b1 : mRNA-UTR, Stol : within coordinates of	A	A/G
rs8240237	ChrX:86195791	exma : within coordinates of, Nr0b1 : mRNA-UTR, Stol : within coordinates of	A	A/G
rs8240238	ChrX:86195792	exma : within coordinates of, Nr0b1 : mRNA-UTR, Stol : within coordinates of	G	A/G
rs8240239	ChrX:86195826	exma : within coordinates of, Nr0b1 : mRNA-UTR, Stol : within coordinates of	C	C/T
rs31224595	ChrX:86345704	exma : within coordinates of, Gm41 : Noncoding-Transcript-Variant, Stol : within coordinates of	T	C/T

rs31118897	ChrX:86345871	exma : within coordinates of, Gm41 : Noncoding-Transcript-Variant, Stol : within coordinates of	C	A/C
rs31209037	ChrX:86345878	exma : within coordinates of, Gm41 : Noncoding-Transcript-Variant, Stol : within coordinates of	C	C/T
rs16797825	ChrX:93304780	exma : within coordinates of, Pola1 : mRNA-UTR, Stol : within coordinates of	G	A/G
rs8278048	ChrX:93461480	exma : within coordinates of, Pola1 : Coding-NonSynonymous, Stol : within coordinates of	A	A/G
rs29071660	ChrX:93764688	exma : within coordinates of, Pdk3 : mRNA-UTR, Stol : within coordinates of	T	A/T
rs29071659	ChrX:93764818	exma : within coordinates of, Pdk3 : mRNA-UTR, Stol : within coordinates of	G	G/T
rs4232582	ChrX:93764956	exma : within coordinates of, Pdk3 : mRNA-UTR, Stol : within coordinates of	C	C/T
rs4232581	ChrX:93764979	exma : within coordinates of, Pdk3 : mRNA-UTR, Stol : within coordinates of	A	A/C
rs4232580	ChrX:93765031	exma : within coordinates of, Pdk3 : mRNA-UTR, Stol : within coordinates of	C	C/G
rs4232579	ChrX:93765056	exma : within coordinates of, Pdk3 : mRNA-UTR, Stol : within coordinates of	T	G/T
rs4232578	ChrX:93765057	exma : within coordinates of, Pdk3 : mRNA-UTR, Stol : within coordinates of	C	C/T
rs4232577	ChrX:93765082	exma : within coordinates of, Pdk3 : mRNA-UTR, Stol : within coordinates of	?	A/G
rs13475475	ChrX:93765137	exma : within coordinates of, Pdk3 : mRNA-UTR, Stol : within coordinates of	A	A/G
rs13459181	ChrX:94536135	Maged1 : Coding-Synonymous, Stol : within coordinates of	G	A/G
rs29069907	ChrX:94636845	Gspt2 : Coding-NonSynonymous, Stol : within coordinates of	T	G/T
rs6407378	ChrX:94637006	Gspt2 : Coding-Synonymous, Stol : within coordinates of	T	C/T
rs33901002	ChrX:94727633	Stol : within coordinates of, Zxdb : mRNA-UTR	T	A/T
rs33901007	ChrX:94791730	Stol : within coordinates of, Zxda : Noncoding-Transcript-Variant	A	A/C
rs33900758	ChrX:94791733	Stol : within coordinates of, Zxda : Noncoding-Transcript-Variant	C	A/C
rs31361377	ChrX:95049451	Arhgef9 : mRNA-UTR, Stol : within coordinates of	C	C/T

rs29088596	ChrX:95050676	Arhgef9 : mRNA-UTR, Stol : within coordinates of	C	C/T
rs13483887	ChrX:95051578	Arhgef9 : mRNA-UTR, Stol : within coordinates of	A	A/G
rs13483893	ChrX:96307632	Hsf3 : mRNA-UTR, Stol : within coordinates of	A	A/G
rs29084788	ChrX:96338532	Hsf3 : Coding-NonSynonymous, Stol : within coordinates of	T	G/T
rs13483894	ChrX:96499054	Heph : Coding-NonSynonymous, Stol : within coordinates of	A	A/G
rs13483897	ChrX:97076232	Pgr15l : Coding-Synonymous, Stol : within coordinates of	T	C/T
rs6365064	ChrX:97336265	Eda2r : mRNA-UTR, Stol : within coordinates of	C	C/G
rs6366690	ChrX:97336569	Eda2r : mRNA-UTR, Stol : within coordinates of	A	A/T
rs6367841	ChrX:97336764	Eda2r : mRNA-UTR, Stol : within coordinates of	T	G/T
rs29087480	ChrX:98558145	Ophn1 : mRNA-UTR, Stol : within coordinates of	G	A/G
rs13483900	ChrX:98726113	Ophn1 : Coding-Synonymous, Stol : within coordinates of	T	C/T
rs29084025	ChrX:99066758	Stard8 : Coding-NonSynonymous, Stol : within coordinates of	A	A/G
rs29087988	ChrX:99073225	Stard8 : mRNA-UTR, Stol : within coordinates of	T	C/T
rs8278082	ChrX:99466877	Pja1 : Coding-Synonymous, Stol : within coordinates of	G	A/G
rs8278081	ChrX:99466880	Pja1 : Coding-Synonymous, Stol : within coordinates of	T	C/T
rs8278079	ChrX:99467050	Pja1 : Coding-NonSynonymous, Stol : within coordinates of	C	C/T
rs8278078	ChrX:99467227	Pja1 : Coding-NonSynonymous, Stol : within coordinates of	G	C/G
rs8278077	ChrX:99467308	Pja1 : Coding-Synonymous, Stol : within coordinates of	C	A/C
rs8278076	ChrX:99467320	Pja1 : Coding-NonSynonymous, Stol : within coordinates of	C	A/C
rs8278083	ChrX:99467459	Pja1 : Coding-Synonymous, Stol : within coordinates of	C	C/T
rs16797834	ChrX:99467724	Pja1 : Coding-NonSynonymous, Pja1 : Intron, Stol : within coordinates of	C	C/T
rs8278084	ChrX:99468011	Pja1 : Coding-Synonymous, Pja1 : Intron, Stol : within coordinates of	G	A/G

rs31119769	ChrX:100594802	P2ry4 : mRNA-UTR, Stol : within coordinates of	C	C/G
rs29080867	ChrX:100605562	Arr3 : mRNA-UTR, Stol : within coordinates of	C	C/T
rs8278111	ChrX:100623889	Pdzd11 : Coding-Synonymous, Stol : within coordinates of	G	C/G
rs8236407	ChrX:100626224	Cpgi22859 : within coordinates of, Kif4 : mRNA-UTR, Pdzd11 : Locus-Region, Stol : within coordinates of	G	G/T
rs8236408	ChrX:100626282	Cpgi22859 : within coordinates of, Kif4 : mRNA-UTR, Pdzd11 : Locus-Region, Stol : within coordinates of	C	A/C
rs16783859	ChrX:100626374	Cpgi22859 : within coordinates of, Kif4 : mRNA-UTR, Pdzd11 : Locus-Region, Stol : within coordinates of	C	C/T
rs8278322	ChrX:100677308	Kif4 : Coding-Synonymous, Stol : within coordinates of	A	A/T
rs8278413	ChrX:100679851	Kif4 : Coding-Synonymous, Stol : within coordinates of	G	A/G
rs8278414	ChrX:100679895	Kif4 : Coding-NonSynonymous, Stol : within coordinates of	G	A/G
rs8278787	ChrX:100718580	Kif4 : Coding-Synonymous, Stol : within coordinates of	G	A/G
rs13483910	ChrX:100726551	Kif4 : Coding-Synonymous, Stol : within coordinates of	G	A/G
rs4232591	ChrX:100726638	Kif4 : Coding-Synonymous, Stol : within coordinates of	T	C/T
rs4232592	ChrX:100726661	Kif4 : mRNA-UTR, Stol : within coordinates of	T	C/T
rs4232593	ChrX:100726676	Kif4 : mRNA-UTR, Stol : within coordinates of	T	C/T
rs4232594	ChrX:100726689	Kif4 : mRNA-UTR, Stol : within coordinates of	G	C/G
rs4232595	ChrX:100726709	Kif4 : mRNA-UTR, Stol : within coordinates of	A	A/C
rs4232596	ChrX:100726740	Kif4 : mRNA-UTR, Stol : within coordinates of	A	A/T
rs4232597	ChrX:100726862	Kif4 : mRNA-UTR, Stol : within coordinates of	C	C/T
rs4232598	ChrX:100726890	Kif4 : mRNA-UTR, Stol : within coordinates of	A	A/G
rs4232599	ChrX:100726981	Kif4 : mRNA-UTR, Stol : within coordinates of	A	A/G
rs4232600	ChrX:100726984	Kif4 : mRNA-UTR, Stol : within coordinates of	G	A/G

rs8278800	ChrX:100727166	Kif4 : mRNA-UTR, Stol : within coordinates of	G	G/T
rs16795674	ChrX:100727250	Kif4 : mRNA-UTR, Stol : within coordinates of	A	A/C
rs31256336	ChrX:100796705	Dlg3 : Coding-Synonymous, Dlg3 : mRNA-UTR, Stol : within coordinates of	C	C/T
rs4232601	ChrX:101083875	Slc7a3 : Coding-Synonymous, Stol : within coordinates of	T	A/T
rs3090974	ChrX:101264736	Gm614 : Locus-Region, Gm20489 : within coordinates of, Il2rg : mRNA-UTR, Stol : within coordinates of	C	C/T
rs8238299	ChrX:101384452	Gjb1 : Coding-Synonymous, Stol : within coordinates of	A	A/T
rs8238300	ChrX:101385011	Cpgi22867 : within coordinates of, Gjb1 : Coding-NonSynonymous, Stol : within coordinates of	G	C/G
rs31253166	ChrX:102739727	Gm14839 : Coding-NonSynonymous, Stol : within coordinates of	G	G/T
rs31072009	ChrX:102859544	1700031F05Rik : mRNA-UTR, Stol : within coordinates of	G	C/G
rs31060216	ChrX:102864679	1700031F05Rik : Coding-NonSynonymous, Stol : within coordinates of	T	A/T
rs29082015	ChrX:102903356	Dmrtc1a : mRNA-UTR, Stol : within coordinates of	C	A/C
rs13483918	ChrX:102906863	Dmrtc1a : Coding-Synonymous, Dmrtc1a : Locus-Region, Stol : within coordinates of	A	A/T
rs29080554	ChrX:103391677	Chic1 : mRNA-UTR, Stol : within coordinates of	A	A/G
rs31372548	ChrX:103474157	Gm26992 : within coordinates of, Stol : within coordinates of, Tsix : Intron, Xist : Noncoding-Transcript-Variant	A	A/C
rs31391216	ChrX:103474838	Gm26992 : within coordinates of, Stol : within coordinates of, Tsix : Intron, Xist : Noncoding-Transcript-Variant	C	C/T
rs31308397	ChrX:103474914	Gm26992 : within coordinates of, Stol : within coordinates of, Tsix : Intron, Xist : Noncoding-Transcript-Variant	A	A/C
rs31295127	ChrX:103480851	Gm26992 : within coordinates of, Stol : within coordinates of, Tsix : Intron, Xist : Noncoding-Transcript-Variant	G	C/G
rs29082536	ChrX:103570693	Ftx : Noncoding-Transcript-Variant, Ftx : Intron, Stol : within coordinates of	G	A/G

rs4232614	ChrX:103699233	Slc16a2 : mRNA-UTR, Stol : within coordinates of	G	A/G
rs4232613	ChrX:103699303	Slc16a2 : mRNA-UTR, Stol : within coordinates of	A	A/G
rs4232612	ChrX:103699418	Slc16a2 : mRNA-UTR, Stol : within coordinates of	A	A/C
rs4232611	ChrX:103699520	Slc16a2 : mRNA-UTR, Stol : within coordinates of	A	A/G
rs31217367	ChrX:104078549	C77370 : mRNA-UTR, Stol : within coordinates of	A	A/G
rs8268102	ChrX:104300804	Abcb7 : Coding-Synonymous, Stol : within coordinates of	G	A/G
rs6291220	ChrX:104537229	Stol : within coordinates of, Zdhhc15 : mRNA-UTR	G	A/G
rs31186434	ChrX:104833063	1700121L16Rik : Noncoding-Transcript-Variant, Stol : within coordinates of	G	A/G
rs29075712	ChrX:105776224	Fgf16 : mRNA-UTR, Stol : within coordinates of	A	A/G
rs29074368	ChrX:105830896	Atrx : Coding-NonSynonymous, Stol : within coordinates of	C	C/T
rs6207541	ChrX:105876866	Atrx : Coding-NonSynonymous, Stol : within coordinates of	C	C/T
rs29076359	ChrX:105971333	Magt1 : Intron, Magt1 : mRNA-UTR, Stol : within coordinates of	T	A/T
rs6256328	ChrX:105992452	Magt1 : Coding-NonSynonymous, Stol : within coordinates of	G	A/G
rs4232616	ChrX:106021975	Cox7b : mRNA-UTR, Stol : within coordinates of	T	C/T
rs4232617	ChrX:106022221	Cox7b : mRNA-UTR, Stol : within coordinates of	C	C/T
rs13483926	ChrX:106088410	Atp7a : Coding-NonSynonymous, Stol : within coordinates of	T	C/T
rs33875836	ChrX:106125082	Atp7a : mRNA-UTR, Stol : within coordinates of	G	A/G
rs31411247	ChrX:106125962	Atp7a : mRNA-UTR, Stol : within coordinates of	G	G/T
rs31069423	ChrX:106127952	Atp7a : mRNA-UTR, Stol : within coordinates of	G	G/T
rs29095036	ChrX:106158262	Stol : within coordinates of, Tlr13 : Coding-NonSynonymous	C	C/T
rs13483928	ChrX:106421737	Fndc3c1 : Coding-Synonymous, Stol : within coordinates of	C	C/T
rs29096674	ChrX:106577602	Cysltr1 : mRNA-UTR, Stol : within coordinates of	A	A/C

rs29097039	ChrX:106671589	Gm5127 : mRNA-UTR, Stol : within coordinates of	C	A/C
rs29095079	ChrX:106709740	Gm5127 : Coding-NonSynonymous, Stol : within coordinates of	G	A/G
rs6164723	ChrX:108768973	Brwd3 : Coding-Synonymous, Stol : within coordinates of	C	C/T
rs29093046	ChrX:109005703	Hmgn5 : Coding-NonSynonymous, Stol : within coordinates of	T	C/T
rs29091513	ChrX:109013376	Hmgn5 : mRNA-UTR, Stol : within coordinates of	C	C/T
rs13474710	ChrX:109162091	Sh3bgrl : mRNA-UTR, Stol : within coordinates of	T	C/T
rs33877359	ChrX:111389226	Rps6ka6 : mRNA-UTR, Stol : within coordinates of	A	A/C
rs29103158	ChrX:111389482	Rps6ka6 : mRNA-UTR, Stol : within coordinates of	A	A/G
rs13460274	ChrX:111389589	Rps6ka6 : mRNA-UTR, Stol : within coordinates of	A	A/C
rs31308852	ChrX:111450299	Rps6ka6 : Coding-NonSynonymous, Stol : within coordinates of	C	C/T
rs29101893	ChrX:111941317	L1Md-Gf21 : Noncoding-Transcript-Variant, Stol : within coordinates of	T	A/T
rs29101043	ChrX:111963692	L1Md-Gf21 : Noncoding-Transcript-Variant, Stol : within coordinates of	T	C/T
rs29103050	ChrX:112119314	Stol : within coordinates of, Tex16 : Coding-Synonymous	C	C/T
rs29103049	ChrX:112119464	Stol : within coordinates of, Tex16 : Coding-Synonymous	T	C/T
rs29103048	ChrX:112119542	Stol : within coordinates of, Tex16 : Coding-NonSynonymous	T	A/T
rs29103047	ChrX:112119556	Stol : within coordinates of, Tex16 : Coding-NonSynonymous	G	C/G
rs29103045	ChrX:112120208	Stol : within coordinates of, Tex16 : Coding-NonSynonymous	A	A/T
rs29103044	ChrX:112120620	Stol : within coordinates of, Tex16 : Coding-NonSynonymous	A	A/C
rs31284292	ChrX:112120895	Stol : within coordinates of, Tex16 : Coding-Synonymous	A	A/C
rs29102573	ChrX:112120980	Stol : within coordinates of, Tex16 : Coding-NonSynonymous	C	C/G
rs29102148	ChrX:112127294	Stol : within coordinates of, Tex16 : mRNA-UTR	G	G/T
rs29099929	ChrX:112243795	4933403O08Rik : mRNA-UTR, Stol : within coordinates of	A	A/G

rs31069644	ChrX:112405761	Satl1 : Coding-NonSynonymous, Stol : within coordinates of	A	A/C
rs29100376	ChrX:112406204	Satl1 : Coding-NonSynonymous, Stol : within coordinates of	T	C/T
rs29100140	ChrX:112512563	2010106E10Rik : Coding-Synonymous, Stol : within coordinates of	G	A/G
rs29099968	ChrX:112557010	2010106E10Rik : Coding-NonSynonymous, Stol : within coordinates of	A	A/C
rs29099900	ChrX:112557928	2010106E10Rik : Coding-NonSynonymous, Stol : within coordinates of	A	A/C
rs29099899	ChrX:112557973	2010106E10Rik : mRNA-UTR, Stol : within coordinates of	A	A/C
rs29099898	ChrX:112558053	2010106E10Rik : mRNA-UTR, Stol : within coordinates of	T	A/T
rs29103145	ChrX:112633893	Stol : within coordinates of, Zfp711 : mRNA-UTR	G	G/T
rs29103144	ChrX:112634076	Stol : within coordinates of, Zfp711 : mRNA-UTR	G	A/G
rs29102813	ChrX:112634153	Stol : within coordinates of, Zfp711 : mRNA-UTR	A	A/G
rs29102812	ChrX:112634880	Stol : within coordinates of, Zfp711 : mRNA-UTR	G	G/T
rs29102806	ChrX:112638822	Pof1b : mRNA-UTR, Stol : within coordinates of	G	A/G
rs31388500	ChrX:112639989	Pof1b : mRNA-UTR, Stol : within coordinates of	G	A/G
rs31404111	ChrX:112640000	Pof1b : mRNA-UTR, Stol : within coordinates of	A	A/T
rs29102272	ChrX:112640243	Pof1b : mRNA-UTR, Stol : within coordinates of	G	G/T
rs13465297	ChrX:113041617	Chm : mRNA-UTR, Stol : within coordinates of	A	A/C
rs13465298	ChrX:113041657	Chm : mRNA-UTR, Stol : within coordinates of	A	A/C
rs29100859	ChrX:113041972	Chm : mRNA-UTR, Stol : within coordinates of	G	A/G
rs13465299	ChrX:113042214	Chm : mRNA-UTR, Stol : within coordinates of	A	A/C
rs29100856	ChrX:113043427	Chm : mRNA-UTR, Stol : within coordinates of	T	G/T
rs29100855	ChrX:113043628	Chm : Coding-NonSynonymous, Stol : within coordinates of	G	A/G
rs13483946	ChrX:113298447	Dach2 : Coding-Synonymous, Stol : within coordinates of	G	A/G
rs33869579	ChrX:113817702	Dach2 : Coding-Synonymous, Stol : within coordinates of	C	A/C

rs31047137	ChrX:113831214	Dach2 : mRNA-UTR, Dach2 : Intron, Stol : within coordinates of	A	A/G
rs31074907	ChrX:113831299	Dach2 : Intron, Dach2 : mRNA-UTR, Stol : within coordinates of	A	A/T
rs31138510	ChrX:113831317	Dach2 : mRNA-UTR, Dach2 : Intron, Stol : within coordinates of	C	C/T
rs33870778	ChrX:113831322	Dach2 : Intron, Dach2 : mRNA-UTR, Stol : within coordinates of	G	A/G
rs31298789	ChrX:113831601	Dach2 : Intron, Dach2 : mRNA-UTR, Stol : within coordinates of	G	A/G
rs33869644	ChrX:113831912	Dach2 : mRNA-UTR, Dach2 : Intron, Stol : within coordinates of	G	G/T
rs33876554	ChrX:113831999	Dach2 : mRNA-UTR, Dach2 : Intron, Stol : within coordinates of	G	A/G
rs31263427	ChrX:113833934	Dach2 : Intron, Dach2 : mRNA-UTR, Stol : within coordinates of	G	C/G
rs31223686	ChrX:113833943	Dach2 : Intron, Dach2 : mRNA-UTR, Stol : within coordinates of	T	C/T
rs31136157	ChrX:113834115	Dach2 : Intron, Dach2 : mRNA-UTR, Stol : within coordinates of	C	A/C
rs31211588	ChrX:113834743	Dach2 : Intron, Dach2 : mRNA-UTR, Stol : within coordinates of	A	A/G
rs31409360	ChrX:113834830	Dach2 : Intron, Dach2 : mRNA-UTR, Stol : within coordinates of	G	G/T
rs31150578	ChrX:113835058	Dach2 : mRNA-UTR, Dach2 : Intron, Stol : within coordinates of	G	G/T
rs29096876	ChrX:113836044	Dach2 : mRNA-UTR, Stol : within coordinates of	A	A/G
rs31234449	ChrX:113836125	Dach2 : mRNA-UTR, Stol : within coordinates of	A	A/G
rs29096875	ChrX:113836163	Dach2 : mRNA-UTR, Stol : within coordinates of	T	G/T
rs29096874	ChrX:113836363	Dach2 : mRNA-UTR, Stol : within coordinates of	T	C/T
rs29097204	ChrX:114474843	Klhl4 : Coding-Synonymous, Stol : within coordinates of	A	A/T
rs29098297	ChrX:114560093	Klhl4 : mRNA-UTR, Stol : within coordinates of	T	G/T
rs29098296	ChrX:114560097	Klhl4 : mRNA-UTR, Stol : within coordinates of	A	A/G
rs29104955	ChrX:116478111	Cpxcr1 : Coding-NonSynonymous, Stol : within coordinates of	G	A/G
rs29111911	ChrX:119928181	Pabpc5 : Coding-Synonymous, Stol : within coordinates of	T	C/T
rs29111907	ChrX:119929900	Pabpc5 : mRNA-UTR, Stol : within coordinates of	G	C/G

rs29111906	ChrX:119929937	Pabpc5 : mRNA-UTR, Stol : within coordinates of	G	A/G
rs29111905	ChrX:119929966	Pabpc5 : mRNA-UTR, Stol : within coordinates of	C	C/T
rs29111904	ChrX:119930075	Pabpc5 : mRNA-UTR, Stol : within coordinates of	T	G/T
rs31318064	ChrX:120313006	H2afb3 : Coding-Synonymous, Pcdh11x : Intron, Stol : within coordinates of	A	A/G
rs31141320	ChrX:120365327	Pcdh11x : Coding-Synonymous, Stol : within coordinates of	T	C/T
rs29110793	ChrX:120400943	Pcdh11x : Coding-Synonymous, Stol : within coordinates of	T	C/T
rs31302430	ChrX:122394810	Nap1l3 : mRNA-UTR, Stol : within coordinates of	A	A/G
rs33875855	ChrX:122394881	Nap1l3 : mRNA-UTR, Stol : within coordinates of	G	A/G
rs33872156	ChrX:122395015	Nap1l3 : mRNA-UTR, Stol : within coordinates of	A	A/G
rs29254526	ChrX:126814133	4932411N23Rik : Coding-NonSynonymous, Stol : within coordinates of	T	C/T
rs29254524	ChrX:126814988	4932411N23Rik : Coding-NonSynonymous, Stol : within coordinates of	C	A/C
rs29254112	ChrX:126815343	4932411N23Rik : Coding-Synonymous, Stol : within coordinates of	C	C/T
rs29254111	ChrX:126815554	4932411N23Rik : mRNA-UTR, Stol : within coordinates of	T	C/T
rs29254110	ChrX:126815587	4932411N23Rik : mRNA-UTR, Stol : within coordinates of	C	C/T
rs29254109	ChrX:126815686	4932411N23Rik : mRNA-UTR, Stol : within coordinates of	C	C/G
rs29253579	ChrX:127061303	Gm382 : Coding-Synonymous, Stol : within coordinates of	A	A/G
rs29253578	ChrX:127061391	Gm382 : Coding-NonSynonymous, Stol : within coordinates of	C	C/T
rs29253577	ChrX:127061444	Gm382 : Coding-Synonymous, Gm382 : Coding-NonSynonymous, Stol : within coordinates of	C	C/T
rs29253576	ChrX:127062260	Gm382 : Coding-Synonymous, Stol : within coordinates of	C	C/T
rs29253575	ChrX:127062271	Gm382 : Coding-NonSynonymous, Stol : within coordinates of	T	A/T
rs29253574	ChrX:127062312	Gm382 : Coding-NonSynonymous, Stol : within coordinates of	T	G/T
rs31253169	ChrX:127062938	Gm382 : Coding-Synonymous, Stol : within coordinates of	A	A/G

rs31117747	ChrX:127063394	Gm382 : Coding-Synonymous, Stol : within coordinates of	T	C/T
rs29257504	ChrX:127063759	Gm382 : Coding-NonSynonymous, Stol : within coordinates of	C	C/T
rs29251511	ChrX:127394655	4921511C20Rik : Noncoding-Transcript-Variant, Stol : within coordinates of	T	G/T
rs29251510	ChrX:127394810	4921511C20Rik : Noncoding-Transcript-Variant, Stol : within coordinates of	C	C/T
rs29251509	ChrX:127395091	4921511C20Rik : Noncoding-Transcript-Variant, Stol : within coordinates of	A	A/G
rs29117483	ChrX:127721263	Cldn34c4 : Coding-NonSynonymous, Stol : within coordinates of	A	A/T
rs13483989	ChrX:128376473	Gm4916 : Coding-NonSynonymous, Gm38390 : within coordinates of, Stol : within coordinates of	T	G/T
rs31304925	ChrX:130460947	Diaph2 : mRNA-UTR, Stol : within coordinates of	G	C/G
rs31061049	ChrX:130461310	Diaph2 : mRNA-UTR, Stol : within coordinates of	A	A/C
rs31317598	ChrX:130461674	Diaph2 : mRNA-UTR, Stol : within coordinates of	C	A/C
rs31120861	ChrX:130462282	Diaph2 : mRNA-UTR, Stol : within coordinates of	C	C/T
rs33875989	ChrX:130462491	Diaph2 : mRNA-UTR, Stol : within coordinates of	T	C/T
rs31202278	ChrX:130462781	Diaph2 : mRNA-UTR, Stol : within coordinates of	A	A/T
rs31204040	ChrX:130463279	Diaph2 : mRNA-UTR, Stol : within coordinates of	T	A/T
rs31349292	ChrX:130463418	Diaph2 : mRNA-UTR, Stol : within coordinates of	C	C/T
rs33873029	ChrX:130463571	Diaph2 : mRNA-UTR, Stol : within coordinates of	T	C/T
rs31152733	ChrX:130464347	Diaph2 : mRNA-UTR, Stol : within coordinates of	T	A/T
rs31109456	ChrX:130464491	Diaph2 : mRNA-UTR, Stol : within coordinates of	A	A/G
rs13484006	ChrX:134150612	Nox1 : within coordinates of, Xkrx : Coding-NonSynonymous	G	C/G
rs29269122	ChrX:134238207	Trmt2b : Coding-NonSynonymous	T	C/T
rs29269121	ChrX:134238510	Trmt2b : Coding-NonSynonymous	C	A/C
rs29269118	ChrX:134240163	Trmt2b : Coding-NonSynonymous	C	C/G
rs29269261	ChrX:134295265	Tmem35 : mRNA-UTR	T	C/T
rs4232625	ChrX:134305603	Tmem35 : mRNA-UTR	A	A/C
rs4232624	ChrX:134305691	Tmem35 : mRNA-UTR	G	A/G

rs4232623	ChrX:134305701	Tmem35 : mRNA-UTR	C	C/T
rs4232622	ChrX:134305742	Tmem35 : mRNA-UTR	A	A/T
rs4232621	ChrX:134305866	Tmem35 : mRNA-UTR	G	C/G
rs4232620	ChrX:134305932	Tmem35 : mRNA-UTR	A	A/G
rs4232626	ChrX:134453234	Drp2 : mRNA-UTR	C	C/T
rs4232627	ChrX:134453250	Drp2 : mRNA-UTR	A	A/C
rs4232628	ChrX:134453283	Drp2 : mRNA-UTR	G	G/T
rs4232629	ChrX:134453332	Drp2 : mRNA-UTR	A	A/G
rs4232630	ChrX:134453364	Drp2 : mRNA-UTR	G	C/G
rs4232631	ChrX:134453409	Drp2 : mRNA-UTR	T	C/T
rs4232632	ChrX:134453506	Drp2 : mRNA-UTR	G	A/G
rs29271257	ChrX:134556819	Btk : Coding-Synonymous	G	A/G
rs31095566	ChrX:134588390	Gla : mRNA-UTR, Rpl36a : Intron	A	A/G
rs8257377	ChrX:134588638	Gla : mRNA-UTR	C	C/T
rs8257371	ChrX:134588992	Gla : mRNA-UTR	T	C/T
rs8257370	ChrX:134589041	Gla : mRNA-UTR	A	A/T
rs8257369	ChrX:134589076	Gla : mRNA-UTR	C	C/T
rs8257368	ChrX:134589082	Gla : mRNA-UTR	C	C/T
rs8257367	ChrX:134589099	Gla : mRNA-UTR	G	A/G
rs8257366	ChrX:134589102	Gla : mRNA-UTR	A	A/G
rs33870276	ChrX:134589272	Gla : mRNA-UTR	G	A/G
rs8257365	ChrX:134589329	Gla : mRNA-UTR	G	A/G
rs16795290	ChrX:134589443	Gla : mRNA-UTR	G	G/T
rs8257408	ChrX:134591045	Gla : Coding-Synonymous	G	C/G
rs8257407	ChrX:134591051	Gla : Coding-NonSynonymous	C	C/T
rs8257441	ChrX:134592135	Gla : Coding-NonSynonymous	C	A/C
rs8257391	ChrX:134595196	Gla : Coding-Synonymous	A	A/G
rs8257340	ChrX:134596400	Gla : Coding-Synonymous	C	A/C
rs8257379	ChrX:134600916	Gla : Coding-NonSynonymous, Gla : Coding-Synonymous, Hnrnph2 : Locus-Region	A	A/G
rs29268827	ChrX:134606688	Hnrnph2 : mRNA-UTR	C	C/T
rs29267941	ChrX:134696873	Armcx4 : mRNA-UTR	T	A/T
rs29267275	ChrX:134718970	Armcx1 : mRNA-UTR	G	A/G
rs29266975	ChrX:134720089	Armcx1 : mRNA-UTR	A	A/G
rs13463897	ChrX:134721038	Armcx1 : Coding-Synonymous	C	C/T
rs6302291	ChrX:134745558	Gm10344 : Noncoding-Transcript-Variant	C	C/T
rs6315409	ChrX:134745685	Gm10344 : Noncoding-Transcript-Variant	G	G/T
rs6315504	ChrX:134745742	Gm10344 : Noncoding-Transcript-Variant	T	G/T
rs6333562	ChrX:134746761	Gm10344 : Noncoding-Transcript-Variant	G	A/G
rs6334022	ChrX:134746815	Gm10344 : Noncoding-Transcript-Variant	G	G/T
rs6334533	ChrX:134746905	Gm10344 : Noncoding-Transcript-Variant	C	C/T
rs29267646	ChrX:134806958	Armcx2 : mRNA-UTR	C	C/T

rs33876466	ChrX:135023999	3632454L22Rik : Noncoding-Transcript-Variant, Del(XNxf2-Nxf3)1Jw : within coordinates of	T	G/T
rs31168442	ChrX:135024204	3632454L22Rik : Noncoding-Transcript-Variant, Del(XNxf2-Nxf3)1Jw : within coordinates of	T	C/T
rs31349191	ChrX:135624141	Del(XNxf2-Nxf3)1Jw : within coordinates of, Prame : mRNA-UTR	C	C/T
rs29267668	ChrX:135624143	Del(XNxf2-Nxf3)1Jw : within coordinates of, Prame : mRNA-UTR	T	C/T
rs31308815	ChrX:135624144	Del(XNxf2-Nxf3)1Jw : within coordinates of, Prame : mRNA-UTR	G	A/G
rs29267667	ChrX:135624159	Del(XNxf2-Nxf3)1Jw : within coordinates of, Prame : mRNA-UTR	A	A/G
rs29266377	ChrX:136076263	Del(XNxf2-Nxf3)1Jw : within coordinates of, Nxf3 : Coding-Synonymous	C	C/T
rs4151952	ChrX:136214310	Bex1 : Coding-NonSynonymous	G	A/G
rs4151953	ChrX:136214366	Bex1 : Coding-NonSynonymous	G	G/T
rs4151954	ChrX:136214435	Bex1 : Coding-Synonymous	A	A/G
rs4151955	ChrX:136214460	Bex1 : Coding-NonSynonymous	C	C/T
rs31166916	ChrX:136526573	Kir3dl1 : Coding-NonSynonymous, Kir3dl2 : within coordinates of	C	C
rs31270241	ChrX:136526662	Kir3dl1 : Coding-NonSynonymous, Kir3dl2 : within coordinates of	A	A
rs31204132	ChrX:136533951	Kir3dl1 : Coding-NonSynonymous, Kir3dl2 : within coordinates of	C	C/T
rs8238551	ChrX:136836125	Dp(XTceal3-Plp1)1Gmh : within coordinates of, Plp1 : mRNA-UTR	C	C/T
rs8238558	ChrX:136836334	Dp(XTceal3-Plp1)1Gmh : within coordinates of, Plp1 : mRNA-UTR	G	A/G
rs8238559	ChrX:136836472	Dp(XTceal3-Plp1)1Gmh : within coordinates of, Plp1 : mRNA-UTR	G	A/G
rs8238560	ChrX:136836618	Dp(XTceal3-Plp1)1Gmh : within coordinates of, Plp1 : mRNA-UTR	A	A/G
rs8238561	ChrX:136836625	Dp(XTceal3-Plp1)1Gmh : within coordinates of, Plp1 : mRNA-UTR	T	-/T
rs8238562	ChrX:136836626	Dp(XTceal3-Plp1)1Gmh : within coordinates of, Plp1 : mRNA-UTR	A	-/A
rs8238584	ChrX:136836757	Dp(XTceal3-Plp1)1Gmh : within coordinates of, Plp1 : mRNA-UTR	C	C/T
rs8238585	ChrX:136836947	Dp(XTceal3-Plp1)1Gmh : within coordinates of, Plp1 : mRNA-UTR	T	A/T
rs8238586	ChrX:136837075	Dp(XTceal3-Plp1)1Gmh : within coordinates of, Plp1 : mRNA-UTR	C	C/T

rs8238587	ChrX:136837135	Dp(XTceal3-Plp1)1Gmh : within coordinates of, Plp1 : mRNA-UTR	T	A/T
rs8238588	ChrX:136837151	Dp(XTceal3-Plp1)1Gmh : within coordinates of, Plp1 : mRNA-UTR	A	A/C
rs13484019	ChrX:136975409	Tmsb15b1 : Coding-NonSynonymous, Tmsb15l : Coding-NonSynonymous	G	A/G
rs29260514	ChrX:136976801	Tmsb15b1 : mRNA-UTR, Tmsb15l : mRNA-UTR	C	C/T
rs29263210	ChrX:136981301	Slc25a53 : mRNA-UTR	G	A/G
rs29263207	ChrX:136981487	Slc25a53 : mRNA-UTR	G	C/G
rs31365941	ChrX:136982425	Slc25a53 : mRNA-UTR	T	A/T
rs29263191	ChrX:136982547	Slc25a53 : mRNA-UTR	C	C/T
rs31404752	ChrX:136983522	Slc25a53 : Coding-Synonymous	C	C/T
rs29263125	ChrX:136993463	Slc25a53 : Locus-Region, Slc25a53 : Intron, Zcchc18 : mRNA-UTR, Zcchc18 : Intron	A	A/G
rs31066946	ChrX:136995254	Slc25a53 : Intron, Zcchc18 : Coding-NonSynonymous	A	A/G
rs29263092	ChrX:136996437	Slc25a53 : Intron, Zcchc18 : mRNA-UTR	T	C/T
rs31237155	ChrX:137049668	Cpgi22958 : within coordinates of, Fam199x : mRNA-UTR	T	C/T
rs31142707	ChrX:137049722	Cpgi22958 : within coordinates of, Fam199x : mRNA-UTR	A	A/G
rs31123183	ChrX:137062724	Fam199x : Coding-Synonymous	C	C/G
rs31162746	ChrX:137062841	Fam199x : Coding-Synonymous	G	A/G
rs29263163	ChrX:137077416	Fam199x : mRNA-UTR	A	A/G
rs31311003	ChrX:137080432	Fam199x : mRNA-UTR	G	A/G
rs29263157	ChrX:137080501	Fam199x : mRNA-UTR	C	C/T
rs29263122	ChrX:137081771	Fam199x : mRNA-UTR	G	A/G
rs29263121	ChrX:137081905	Fam199x : mRNA-UTR	A	A/G
rs29254709	ChrX:139275263	Platr21 : Noncoding-Transcript-Variant	G	A/G
rs29258580	ChrX:139347542	Platr21 : Intron, Platr21 : Noncoding-Transcript-Variant	C	C/T
rs29258579	ChrX:139347569	Platr21 : Intron, Platr21 : Noncoding-Transcript-Variant	T	C/T
rs33876724	ChrX:139543150	D330045A20Rik : Coding-Synonymous	A	A/G
rs33876014	ChrX:139610652	Cpgi22965 : within coordinates of, Rnf128 : Intron, Rnf128 : mRNA-UTR	T	C/T
rs6306761	ChrX:139665457	Rnf128 : Coding-NonSynonymous	G	A/G
rs13459159	ChrX:139810681	Cldn2 : mRNA-UTR	C	C/T
rs13468202	ChrX:139810899	Cldn2 : mRNA-UTR	C	C/G
rs13484031	ChrX:139840710	Morc4 : Coding-NonSynonymous	A	A/G
rs29257197	ChrX:139942639	Rbm41 : Intron, Rbm41 : mRNA-UTR	T	C/T
rs29257194	ChrX:139944137	Rbm41 : mRNA-UTR, Rbm41 : Intron	T	C/T

rs29257288	ChrX:140106482	Pih1h3b : mRNA-UTR	T	A/T
rs29257287	ChrX:140106509	Pih1h3b : mRNA-UTR	T	A/T
rs31253124	ChrX:140106702	Pih1h3b : mRNA-UTR	T	C/T
rs33868821	ChrX:140539714	Tsc22d3 : mRNA-UTR	T	C/T
rs13484034	ChrX:140678432	Mid2 : Coding-Synonymous	G	A/G
rs31323140	ChrX:140678678	Mid2 : Coding-Synonymous	C	C/T
rs29289362	ChrX:141165407	Col4a6 : mRNA-UTR	T	C/T
rs13484036	ChrX:141165858	Col4a6 : mRNA-UTR	T	C/T
rs31308246	ChrX:141192771	Col4a6 : Coding-NonSynonymous	C	C/T
rs29289403	ChrX:141192785	Col4a6 : Coding-Synonymous	T	C/T
rs33869865	ChrX:141192797	Col4a6 : Coding-Synonymous	C	C/T
rs8255333	ChrX:141721884	Cpgi22974 : within coordinates of, Irs4 : Coding-Synonymous	C	C/T
rs8255331	ChrX:141722133	Irs4 : Coding-Synonymous	C	C/T
rs8255330	ChrX:141722584	Irs4 : Coding-NonSynonymous	C	C/G
rs8255329	ChrX:141722670	Irs4 : Coding-Synonymous	A	A/G
rs8255328	ChrX:141722694	Irs4 : Coding-Synonymous	G	A/G
rs8255327	ChrX:141722763	Irs4 : Coding-Synonymous	C	C/T
rs8255326	ChrX:141723248	Irs4 : Coding-NonSynonymous	T	C/T
rs8255325	ChrX:141723264	Irs4 : Coding-Synonymous	A	A/G
rs8255324	ChrX:141723330	Irs4 : Coding-Synonymous	A	A/G
rs8255323	ChrX:141723838	Irs4 : Coding-NonSynonymous	C	A/C
rs8255322	ChrX:141723879	Irs4 : Coding-Synonymous	G	A/G
rs8255321	ChrX:141724206	Cpgi22975 : within coordinates of, Irs4 : Coding-Synonymous	C	C/T
rs29288973	ChrX:141724440	Cpgi22975 : within coordinates of, Irs4 : Coding-Synonymous	G	A/G
rs8255320	ChrX:141724611	Cpgi22975 : within coordinates of, Irs4 : Coding-Synonymous	C	C/T
rs29284231	ChrX:142853665	Ammecr1 : mRNA-UTR	G	A/G
rs29284230	ChrX:142854741	Ammecr1 : mRNA-UTR	A	A/G
rs29284224	ChrX:142855548	Ammecr1 : Coding-Synonymous	G	A/G
rs29283938	ChrX:143287145	Chrdl1 : mRNA-UTR	G	G/T
rs29284347	ChrX:143297830	Chrdl1 : mRNA-UTR, Chrdl1 : Intron	A	A/G
rs29283398	ChrX:143303477	Chrdl1 : Coding-NonSynonymous	T	C/T
rs29286933	ChrX:143518719	Pak3 : mRNA-UTR	C	C/T
rs16793515	ChrX:143802524	Capn6 : mRNA-UTR	G	G/T
rs16793516	ChrX:143802705	Capn6 : mRNA-UTR	T	A/T
rs6321562	ChrX:143802886	Capn6 : mRNA-UTR	T	C/T
rs6324456	ChrX:143803446	Capn6 : mRNA-UTR	G	A/G
rs16793519	ChrX:143805454	Capn6 : Coding-Synonymous	T	C/T
rs29283011	ChrX:143855928	Dcx : mRNA-UTR	G	G/T

rs6241586	ChrX:143857208	Dcx : mRNA-UTR	A	A/G
rs6255078	ChrX:143857440	Dcx : mRNA-UTR	T	G/T
rs29284112	ChrX:143875864	Dcx : Coding-Synonymous	A	A/T
rs31149333	ChrX:144456585	Trpc5 : Intron, Trpc5os : mRNA-UTR	C	C/T
rs29296085	ChrX:150585242	Apex2 : Coding-Synonymous, Stpy : within coordinates of	G	A/G
rs29293675	ChrX:150643586	Pfkfb1 : Coding-Synonymous, Stpy : within coordinates of	C	C/T
rs29297517	ChrX:150646933	Stpy : within coordinates of, Tro : Coding-NonSynonymous, Tro : Intron	G	G/T
rs29296632	ChrX:150655043	Stpy : within coordinates of, Tro : Coding-NonSynonymous	A	A/G
rs31156594	ChrX:150850001	Itih5l-ps : Coding-NonSynonymous, Stpy : within coordinates of	T	G/T
rs29297117	ChrX:151295915	A230072E10Rik : within coordinates of, Stpy : within coordinates of, Wnk3 : Noncoding-Transcript-Variant	C	C/T
rs31397803	ChrX:151313875	A230072E10Rik : Noncoding-Transcript-Variant, Stpy : within coordinates of, Wnk3 : Intron	C	C/T
rs31298718	ChrX:151313880	A230072E10Rik : Noncoding-Transcript-Variant, Stpy : within coordinates of, Wnk3 : Intron	G	C/G
rs29294869	ChrX:151317505	A230072E10Rik : Intron, Stpy : within coordinates of, Wnk3 : Noncoding-Transcript-Variant	G	G/T
rs33866804	ChrX:151340836	A230072E10Rik : Noncoding-Transcript-Variant, Stpy : within coordinates of	G	A/G
rs31266096	ChrX:151470124	Fam120c : mRNA-UTR, Stpy : within coordinates of	T	C/T
rs6158018	ChrX:151472161	Fam120c : mRNA-UTR, Stpy : within coordinates of	T	C/T
rs6158556	ChrX:151472243	Fam120c : mRNA-UTR, Stpy : within coordinates of	T	C/T
rs6160268	ChrX:151472573	Fam120c : mRNA-UTR, Stpy : within coordinates of	C	C/T
rs31140939	ChrX:151472932	Fam120c : mRNA-UTR, Stpy : within coordinates of	C	C/T
rs29294514	ChrX:151473162	Fam120c : mRNA-UTR, Stpy : within coordinates of	C	C/T
rs31200552	ChrX:151521357	Cpgi22993 : within coordinates of, Phf8 : mRNA-UTR, Stpy : within coordinates of	A	A/C
rs31119632	ChrX:151625745	Phf8 : Intron, Phf8 : mRNA-UTR, Stpy : within coordinates of	C	C/T
rs29296351	ChrX:151874619	Huwe1 : Coding-Synonymous, Stpy : within coordinates of	G	C/G

rs6207314	ChrX:151890062	Huwe1 : Coding-Synonymous, Stpy : within coordinates of	A	A/G
rs29295859	ChrX:151918438	Huwe1 : Coding-Synonymous, Stpy : within coordinates of	C	C/T
rs31172003	ChrX:151918486	Huwe1 : Coding-Synonymous, Stpy : within coordinates of	T	C/T
rs31204598	ChrX:153113039	Gm5396 : Coding-NonSynonymous, Stpy : within coordinates of	G	A/G
rs13484077	ChrX:153113053	Gm5396 : Coding-NonSynonymous, Stpy : within coordinates of	G	A/G
rs3703205	ChrX:153119069	Foxr2 : mRNA-UTR, Stpy : within coordinates of	A	A/G
rs29289138	ChrX:153131997	Foxr2 : mRNA-UTR, Stpy : within coordinates of	T	C/T
rs33875338	ChrX:153499425	Stpy : within coordinates of, Ubqln2 : Coding-Synonymous	A	A/C
rs33872487	ChrX:155125038	Cldn34b2 : mRNA-UTR, Stpy : within coordinates of	G	A/G
rs29312082	ChrX:157364293	Phex : Coding-Synonymous, Stpy : within coordinates of	A	A/G
rs13484095	ChrX:157444417	Sms : mRNA-UTR, Stpy : within coordinates of	A	A/G
rs29312615	ChrX:157445557	Sms : mRNA-UTR, Stpy : within coordinates of	G	C/G
rs31315682	ChrX:157445876	Sms : mRNA-UTR, Stpy : within coordinates of	A	A/T
rs8255395	ChrX:157461266	Sms : Coding-Synonymous, Stpy : within coordinates of	T	C/T
rs8255393	ChrX:157462118	Sms : Coding-Synonymous, Stpy : within coordinates of	C	C/G
rs16783933	ChrX:157550759	Mbtps2 : mRNA-UTR, Stpy : within coordinates of	A	A/G
rs16783932	ChrX:157550914	Mbtps2 : mRNA-UTR, Stpy : within coordinates of	C	C/T
rs16783929	ChrX:157554941	Mbtps2 : Coding-Synonymous, Stpy : within coordinates of	A	A/G
rs16783928	ChrX:157555006	Mbtps2 : Coding-NonSynonymous, Stpy : within coordinates of	G	A/G
rs16783927	ChrX:157555010	Mbtps2 : Coding-Synonymous, Stpy : within coordinates of	A	A/G
rs16802211	ChrX:157555609	Mbtps2 : Coding-Synonymous, Stpy : within coordinates of	C	C/T
rs16802210	ChrX:157555675	Mbtps2 : Coding-Synonymous, Stpy : within coordinates of	T	C/T

rs16802174	ChrX:157572952	Mbtps2 : Coding-NonSynonymous, Stpy : within coordinates of	C	C/T
rs16802152	ChrX:157576128	Mbtps2 : Coding-Synonymous, Stpy : within coordinates of	T	C/T
rs16802151	ChrX:157576143	Mbtps2 : Coding-Synonymous, Stpy : within coordinates of	C	C/T
rs6225166	ChrX:157822062	Cnksr2 : mRNA-UTR, Stpy : within coordinates of	C	C/G
rs6157681	ChrX:162927503	Ctps2 : Noncoding-Transcript-Variant, Ctps2 : Coding-Synonymous, Li : within coordinates of	C	C/T
rs31454705	ChrX:164433462	Piga : mRNA-UTR	C	C/T
rs13484100	ChrX:164533536	Asb9 : Coding-Synonymous	G	A/G
rs31078373	ChrX:165130094	Glra2 : mRNA-UTR	A	A/C
rs31277961	ChrX:165257182	Glra2 : Coding-Synonymous	G	A/G
rs8255433	ChrX:165281831	Glra2 : Coding-Synonymous	C	C/T
rs13460315	ChrX:166180639	Gemin8 : Coding-NonSynonymous	G	A/G
rs13459161	ChrX:166387243	Gpm6b : mRNA-UTR	G	A/G
rs31297907	ChrX:166527447	Egfl6 : Coding-NonSynonymous	G	G/T
rs8252436	ChrX:167209185	Cpgi23052 : within coordinates of, Tmsb4x : mRNA-UTR	C	C/G

Appendix III. GSEA for biological processes and pathway overrepresentation analysis on the DEGs

GSEA positive enrichment

BIOLOGICAL PROCESS	NES	ES	FDR
GO_CELL_CELL_SIGNALING	3.35	0.23	0.00E+00
GO_RESPONSE_TO_TYPE_I_INTERFERON	3.25	0.58	0.00E+00
GO_CYTOKINE_MEDIATED_SIGNALING_PATHWAY	3.25	0.29	0.00E+00
GO_IMMUNE_RESPONSE	3.09	0.19	1.74E-04
GO_DEFENSE_RESPONSE	3.06	0.18	2.10E-04
GO_INNATE_IMMUNE_RESPONSE	2.99	0.25	3.74E-04
GO_HUMORAL_IMMUNE_RESPONSE	2.94	0.51	5.05E-04
GO_NEUROPEPTIDE_SIGNALING_PATHWAY	2.87	0.62	5.97E-04
GO_CENTRAL_NERVOUS_SYSTEM_NEURON_DIFFERENTIATION	2.89	0.38	6.26E-04
GO_ORGAN_MORPHOGENESIS	2.84	0.18	6.34E-04
GO_RESPONSE_TO_BIOTIC_STIMULUS	2.87	0.19	6.56E-04
GO_CELLULAR_RESPONSE_TO_CYTOKINE_STIMULUS	2.90	0.23	6.75E-04
GO_POSITIVE_REGULATION_OF_MULTICELLULAR_ORGANISMAL_PROCESS	2.80	0.14	8.90E-04
GO_IMMUNE_SYSTEM_PROCESS	2.74	0.13	1.39E-03
GO_RESPONSE_TO_CYTOKINE	2.70	0.19	2.01E-03
GO_G_PROTEIN_COUPLED_RECEPTOR_SIGNALING_PATHWAY	2.67	0.21	2.21E-03
GO_RESPONSE_TO_EXTERNAL_STIMULUS	2.68	0.12	2.26E-03
GO_RESPONSE_TO_BACTERIUM	2.66	0.24	2.29E-03
GO_REGULATION_OF_HORMONE_LEVELS	2.63	0.23	2.69E-03
GO_RESPONSE_TO_PURINE_CONTAINING_COMPOUND	2.51	0.35	6.96E-03
GO_REGULATION_OF_INTERFERON_GAMMA_PRODUCTION	2.50	0.45	7.08E-03
GO_REGULATION_OF_SYNAPSE_ORGANIZATION	2.47	0.39	7.10E-03
GO_TELENCEPHALON_DEVELOPMENT	2.46	0.28	7.11E-03
GO_MESENCHYMAL_CELL_DIFFERENTIATION	2.47	0.32	7.28E-03
GO_RESPONSE_TO_CAMP	2.47	0.42	7.35E-03
GO_NEGATIVE_REGULATION_OF_PEPTIDASE_ACTIVITY	2.49	0.27	7.43E-03
GO_SENSORY_PERCEPTION_OF_PAIN	2.47	0.45	7.59E-03
GO_TISSUE_DEVELOPMENT	2.47	0.11	7.87E-03
GO_REGULATION_OF_PATHWAY_RESTRICTED_SMAD_PROTEIN_PHOSPHORYLATION	2.44	0.47	8.29E-03
GO_REGIONALIZATION	2.41	0.23	9.11E-03
GO_NEGATIVE_REGULATION_OF_MULTICELLULAR_ORGANISMAL_PROCESS	2.42	0.14	9.19E-03
GO_RESPONSE_TO_ORGANOPHOSPHORUS	2.41	0.35	9.37E-03
GO_EYE_DEVELOPMENT	2.41	0.23	9.42E-03
GO_TAXIS	2.39	0.18	9.76E-03
GO_SKIN_DEVELOPMENT	2.39	0.26	1.00E-02
GO_FOREBRAIN_DEVELOPMENT	2.37	0.22	1.10E-02
GO_CELL_FATE_COMMITMENT	2.35	0.28	1.18E-02
GO_NEGATIVE_REGULATION_OF_HORMONE_SECRETION	2.34	0.43	1.22E-02

GO_STEM_CELL_DIFFERENTIATION	2.34	0.27	1.24E-02
GO_POSITIVE_REGULATION_OF_NERVOUS_SYSTEM_DEVELOPMENT	2.34	0.19	1.25E-02
GO_POSITIVE_REGULATION_OF_IMMUNE_SYSTEM_PROCESS	2.31	0.16	1.44E-02
GO_RESPONSE_TO_VIRUS	2.30	0.25	1.52E-02
GO_HORMONE_METABOLIC_PROCESS	2.29	0.31	1.59E-02
GO_DIENCEPHALON_DEVELOPMENT	2.26	0.40	1.86E-02
GO_ADULT_BEHAVIOR	2.26	0.35	1.88E-02
GO_INFLAMMATORY_RESPONSE	2.25	0.20	1.92E-02
GO_REGULATION_OF_CELL_DIFFERENTIATION	2.25	0.11	1.98E-02
GO_TISSUE_MORPHOGENESIS	2.23	0.17	2.21E-02
GO_PROTEIN_ACTIVATION_CASCADE	2.22	0.48	2.22E-02
GO_POSITIVE_REGULATION_OF_CELL_DIFFERENTIATION	2.21	0.13	2.32E-02
GO_NEURON_MIGRATION	2.20	0.33	2.43E-02
GO_RESPONSE_TO_INTERFERON_GAMMA	2.19	0.30	2.57E-02
GO_NEGATIVE_REGULATION_OF_AXONOGENESIS	2.19	0.39	2.64E-02
GO_EYE_MORPHOGENESIS	2.18	0.31	2.66E-02
GO_REGULATION_OF_NERVOUS_SYSTEM_DEVELOPMENT	2.18	0.14	2.73E-02
GO_CENTRAL_NERVOUS_SYSTEM_NEURON_DEVELOPMENT	2.17	0.43	2.82E-02
GO_SODIUM_ION_TRANSMEMBRANE_TRANSPORT	2.16	0.43	2.89E-02
GO_FEEDING_BEHAVIOR	2.16	0.39	2.91E-02
GO_MORPHOGENESIS_OF_A_BRANCHING_STRUCTURE	2.14	0.25	3.14E-02
GO_POSITIVE_REGULATION_OF_LEUKOCYTE_CHEMOTAXIS	2.13	0.43	3.36E-02
GO_DEFENSE_RESPONSE_TO_OTHER_ORGANISM	2.12	0.19	3.43E-02
GO_REGULATION_OF_SMOOTH_MUSCLE_CELL_MIGRATION	2.12	0.42	3.47E-02
GO_UROGENITAL_SYSTEM_DEVELOPMENT	2.12	0.20	3.50E-02
GO_POSITIVE_REGULATION_OF_IMMUNE_RESPONSE	2.10	0.18	3.70E-02
GO_REGULATION_OF_TRANSMEMBRANE_RECEPTOR_PROTEIN_SERINE_THREONINE_KINASE_SIGNALING_PATHWAY	2.10	0.26	3.72E-02
GO_CELL_CHEMOTAXIS	2.10	0.27	3.74E-02
GO_REGULATION_OF_MULTICELLULAR_ORGANISMAL_DEVELOPMENT	2.09	0.10	3.74E-02
GO_REGULATION_OF_DEVELOPMENTAL_GROWTH	2.09	0.21	3.75E-02
GO_OSTEOBLAST_DIFFERENTIATION	2.10	0.31	3.76E-02
GO_METAL_ION_TRANSPORT	2.09	0.17	3.79E-02
GO_RESPONSE_TO_WOUNDING	2.08	0.15	3.99E-02
GO_NEGATIVE_REGULATION_OF_HYDROLASE_ACTIVITY	2.07	0.20	4.01E-02
GO_REGULATION_OF_SYNAPSE_ASSEMBLY	2.07	0.40	4.02E-02
GO_POTASSIUM_ION_TRANSPORT	2.07	0.32	4.02E-02
GO_NEGATIVE_REGULATION_OF_PROTEOLYSIS	2.06	0.20	4.09E-02
GO_POSITIVE_REGULATION_OF_DEVELOPMENTAL_PROCESS	2.06	0.11	4.14E-02
GO_REGULATION_OF_LYMPHOCYTE_DIFFERENTIATION	2.05	0.32	4.19E-02
GO_REGULATION_OF_NEURON_DIFFERENTIATION	2.05	0.14	4.19E-02
GO_REGULATION_OF_AXONOGENESIS	2.04	0.25	4.36E-02
GO_RETINA_DEVELOPMENT_IN_CAMERA_TYPE_EYE	2.04	0.31	4.38E-02
GO_RESPONSE_TO_RETINOIC_ACID	2.04	0.41	4.46E-02

GO_REGULATION_OF_CELL_DEVELOPMENT	2.03	0.12	4.61E-02
GO_CIRCULATORY_SYSTEM_PROCESS	2.02	0.19	4.65E-02
GO_POSITIVE_REGULATION_OF_NEURON_DIFFERENTIATION	2.02	0.19	4.68E-02
GO_EPITHELIAL_CELL_DIFFERENTIATION	2.02	0.16	4.71E-02
GO_SIGNAL_RELEASE	2.00	0.26	4.94E-02
GO_INORGANIC_ION_TRANSMEMBRANE_TRANSPORT	2.01	0.17	4.95E-02
GO_POSITIVE_REGULATION_OF_RESPONSE_TO_STIMULUS	2.01	0.09	4.96E-02
GO_DEVELOPMENTAL_MATURATION	2.00	0.27	4.96E-02
GO_SENSORY_ORGAN_DEVELOPMENT	2.00	0.16	4.98E-02

GSEA negative enrichment

BIOLOGICAL PROCESS	NES	ES	FDR
GO_CELL_CYCLE	-5.83	-0.31	0.00E+00
GO_CELL_CYCLE_PROCESS	-5.24	-0.30	0.00E+00
GO_CELL_DIVISION	-4.71	-0.39	0.00E+00
GO_MITOTIC_CELL_CYCLE	-4.66	-0.31	0.00E+00
GO_RNA_PROCESSING	-4.62	-0.38	0.00E+00
GO_MITOTIC_NUCLEAR_DIVISION	-4.58	-0.42	0.00E+00
GO_ORGANELLE_FISSION	-4.48	-0.38	0.00E+00
GO_TRANSLATIONAL_INITIATION	-4.04	-0.53	0.00E+00
GO_MRNA_METABOLIC_PROCESS	-4.02	-0.38	0.00E+00
GO_PROTEIN_LOCALIZATION_TO_ENDOPLASMIC_RETICULUM	-3.97	-0.54	0.00E+00
GO_CHROMOSOME_SEGREGATION	-3.90	-0.42	0.00E+00
GO_DNA_METABOLIC_PROCESS	-3.89	-0.29	0.00E+00
GO_NCRNA_PROCESSING	-3.86	-0.38	0.00E+00
GO_ESTABLISHMENT_OF_PROTEIN_LOCALIZATION_TO_ENDOPLASMIC_RETICULUM	-3.86	-0.54	0.00E+00
GO_CHROMOSOME_ORGANIZATION	-3.77	-0.25	0.00E+00
GO_NUCLEAR_TRANSCRIBED_MRNA_CATABOLIC_PROCESS_NONSENSE_MEDIATED_DECAY	-3.76	-0.52	0.00E+00
GO_MACROMOLECULE_CATABOLIC_PROCESS	-3.75	-0.25	0.00E+00
GO_PROTEIN_LOCALIZATION_TO_ORGANELLE	-3.70	-0.31	0.00E+00
GO_NCRNA_METABOLIC_PROCESS	-3.69	-0.34	0.00E+00
GO_RIBONUCLEOPROTEIN_COMPLEX_BIOGENESIS	-3.68	-0.37	0.00E+00
GO_MULTI_ORGANISM_METABOLIC_PROCESS	-3.66	-0.50	0.00E+00
GO_NUCLEAR_CHROMOSOME_SEGREGATION	-3.60	-0.42	0.00E+00
GO_ESTABLISHMENT_OF_PROTEIN_LOCALIZATION_TO_ORGANELLE	-3.53	-0.36	0.00E+00
GO_SISTER_CHROMATID_SEGREGATION	-3.53	-0.44	0.00E+00
GO_PROTEIN_TARGETING_TO_MEMBRANE	-3.52	-0.43	0.00E+00
GO_CELLULAR_RESPONSE_TO_STRESS	-3.48	-0.18	0.00E+00
GO_RIBOSOME_BIOGENESIS	-3.48	-0.38	0.00E+00
GO_ESTABLISHMENT_OF_LOCALIZATION_IN_CELL	-3.48	-0.17	0.00E+00
GO_RNA_CATABOLIC_PROCESS	-3.40	-0.40	2.83E-06
GO_PEPTIDE_METABOLIC_PROCESS	-3.39	-0.29	2.74E-06
GO_RRNA_METABOLIC_PROCESS	-3.38	-0.39	2.65E-06
GO_PROTEIN_TARGETING	-3.36	-0.31	2.56E-06
GO_DNA_REPLICATION	-3.34	-0.40	2.49E-06

GO_CELLULAR_RESPONSE_TO_DNA_DAMAGE_STIMULUS	-3.18	-0.24	7.25E-06
GO_VIRAL_LIFE_CYCLE	-3.18	-0.31	7.04E-06
GO_DNA_RECOMBINATION	-3.16	-0.42	9.14E-06
GO_DNA_REPAIR	-3.15	-0.31	8.90E-06
GO_INTRACELLULAR_PROTEIN_TRANSPORT	-3.10	-0.22	1.75E-05
GO_CELLULAR_CATABOLIC_PROCESS	-3.08	-0.17	2.14E-05
GO_ESTABLISHMENT_OF_PROTEIN_LOCALIZATION_TO_MEMBRANE	-3.07	-0.31	2.09E-05
GO_AMIDE_BIOSYNTHETIC_PROCESS	-3.07	-0.27	2.23E-05
GO_CELLULAR_AMIDE_METABOLIC_PROCESS	-3.04	-0.22	2.77E-05
GO_ORGANONITROGEN_COMPOUND_METABOLIC_PROCESS	-3.01	-0.14	3.65E-05
GO_CELL_CYCLE_PHASE_TRANSITION	-2.98	-0.33	3.95E-05
GO_REGULATION_OF_ORGANELLE_ORGANIZATION	-2.96	-0.16	4.97E-05
GO_ESTABLISHMENT_OF_PROTEIN_LOCALIZATION	-2.91	-0.16	8.28E-05
GO_SMALL_MOLECULE_METABOLIC_PROCESS	-2.83	-0.13	1.70E-04
GO_CELLULAR_MACROMOLECULE_LOCALIZATION	-2.81	-0.16	1.93E-04
GO_CATABOLIC_PROCESS	-2.79	-0.14	2.21E-04
GO_MITOCHONDRION_ORGANIZATION	-2.79	-0.25	2.18E-04
GO_DNA_BIOSYNTHETIC_PROCESS	-2.77	-0.49	2.48E-04
GO_REGULATION_OF_PROTEIN_COMPLEX_DISASSEMBLY	-2.76	-0.32	2.67E-04
GO_PROTEIN_LOCALIZATION	-2.76	-0.13	2.62E-04
GO_DNA_DEPENDENT_DNA_REPLICATION	-2.73	-0.47	3.26E-04
GO_REGULATION_OF_CYTOSKELETON_ORGANIZATION	-2.71	-0.22	3.57E-04
GO_MITOTIC_SISTER_CHROMATID_SEGREGATION	-2.68	-0.45	4.37E-04
GO_SISTER_CHROMATID_COHESION	-2.68	-0.45	4.30E-04
GO_MRNA_PROCESSING	-2.67	-0.37	4.75E-04
GO_INTERSPECIES_INTERACTION_BETWEEN_ORGANISMS	-2.66	-0.20	5.08E-04
GO_NEGATIVE_REGULATION_OF_PROTEIN_COMPLEX_DISASSEMBLY	-2.65	-0.34	5.21E-04
GO_ORGANONITROGEN_COMPOUND_BIOSYNTHETIC_PROCESS	-2.64	-0.16	5.64E-04
GO_NEGATIVE_REGULATION_OF_CELL_CYCLE_PROCESS	-2.63	-0.33	5.86E-04
GO_GLYCOSYL_COMPOUND_METABOLIC_PROCESS	-2.52	-0.29	1.50E-03
GO_MEIOTIC_CELL_CYCLE	-2.50	-0.34	1.62E-03
GO_SINGLE_ORGANISM_BIOSYNTHETIC_PROCESS	-2.49	-0.14	1.66E-03
GO_CELL_CYCLE_CHECKPOINT	-2.47	-0.30	1.88E-03
GO_PROTEIN_MODIFICATION_BY_SMALL_PROTEIN_CONJUGATION_OR_REMOVAL	-2.44	-0.18	2.37E-03
GO_CHROMATIN_MODIFICATION	-2.43	-0.23	2.46E-03
GO_CARBOHYDRATE_CATABOLIC_PROCESS	-2.43	-0.44	2.47E-03
GO_NEGATIVE_REGULATION_OF_CELL_CYCLE	-2.42	-0.21	2.67E-03
GO_RNA_SPLICING	-2.41	-0.36	2.80E-03
GO_SINGLE_ORGANISM_CELLULAR_LOCALIZATION	-2.41	-0.15	2.78E-03
GO_NEGATIVE_REGULATION_OF_ORGANELLE_ORGANIZATION	-2.41	-0.22	2.85E-03
GO_CELL_CYCLE_G2_M_PHASE_TRANSITION	-2.39	-0.39	3.10E-03
GO_REGULATION_OF_CELL_CYCLE	-2.39	-0.15	3.23E-03
GO_DNA_SYNTHESIS_INVOLVED_IN_DNA_REPAIR	-2.38	-0.47	3.28E-03
GO_NEGATIVE_REGULATION_OF_CELL_CYCLE_PHASE_TRANSITION	-2.37	-0.35	3.59E-03
GO_REGULATION_OF_MICROTUBULE_BASED_PROCESS	-2.35	-0.27	3.99E-03
GO_OXIDATION_REDUCTION_PROCESS	-2.35	-0.15	4.08E-03

GO_ORGANELLE_ASSEMBLY	-2.34	-0.20	4.33E-03
GO_COFACTOR_METABOLIC_PROCESS	-2.32	-0.24	4.71E-03
GO_COENZYME_METABOLIC_PROCESS	-2.32	-0.27	4.86E-03
GO_CYTOKINESIS	-2.31	-0.45	4.92E-03
GO_DNA_GEOMETRIC_CHANGE	-2.31	-0.47	4.96E-03
GO_NEGATIVE_REGULATION_OF_MITOTIC_CELL_CYCLE	-2.31	-0.28	5.04E-03
GO_REGULATION_OF_MITOTIC_CELL_CYCLE	-2.30	-0.20	5.07E-03
GO_SULFUR_COMPOUND_BIOSYNTHETIC_PROCESS	-2.30	-0.25	5.07E-03
GO_MEMBRANE_ORGANIZATION	-2.29	-0.14	5.60E-03
GO_MACROMOLECULAR_COMPLEX_ASSEMBLY	-2.28	-0.13	5.67E-03
GO_MICROTUBULE_BASED_PROCESS	-2.26	-0.19	6.43E-03
GO_NEGATIVE_REGULATION_OF_CYTOSKELETON_ORGANIZATION	-2.26	-0.26	6.51E-03
GO_CELL_CYCLE_G1_S_PHASE_TRANSITION	-2.25	-0.33	6.67E-03
GO_ORGANIC_ACID_METABOLIC_PROCESS	-2.24	-0.13	7.21E-03
GO_MICROTUBULE_CYTOSKELETON_ORGANIZATION	-2.23	-0.21	7.95E-03
GO_MICROTUBULE_ORGANIZING_CENTER_ORGANIZATION	-2.23	-0.46	7.86E-03
GO_DNA_INTEGRITY_CHECKPOINT	-2.23	-0.32	7.79E-03
GO_REGULATION_OF_CELL_CYCLE_PROCESS	-2.23	-0.19	7.72E-03
GO_PROTEIN_LOCALIZATION_TO_MEMBRANE	-2.23	-0.20	7.68E-03
GO_REGULATION_OF_DNA_REPAIR	-2.22	-0.47	7.74E-03
GO_REGULATION_OF_CELL_CYCLE_PHASE_TRANSITION	-2.22	-0.25	7.74E-03
GO_SULFUR_COMPOUND_METABOLIC_PROCESS	-2.18	-0.19	1.02E-02
GO_CELLULAR_MACROMOLECULAR_COMPLEX_ASSEMBLY	-2.14	-0.17	1.28E-02
GO_ORGANIC_CYCLIC_COMPOUND_CATABOLIC_PROCESS	-2.14	-0.19	1.32E-02
GO_PROTEASOMAL_PROTEIN_CATABOLIC_PROCESS	-2.14	-0.30	1.32E-02
GO_REGULATION_OF_CELLULAR_PROTEIN_LOCALIZATION	-2.10	-0.17	1.65E-02
GO_PROTEIN_CATABOLIC_PROCESS	-2.08	-0.19	1.78E-02
GO_POSTTRANSCRIPTIONAL_REGULATION_OF_GENE_EXPRESSION	-2.08	-0.22	1.78E-02
GO_PEPTIDYL_LYSINE_MODIFICATION	-2.07	-0.28	1.92E-02
GO_RNA_SPLICING_VIA_TRANSESTERIFICATION_REACTIONS	-2.07	-0.39	1.92E-02
GO_REGULATION_OF_RESPONSE_TO_EXTRACELLULAR_STIMULUS	-2.06	-0.30	1.95E-02
GO_REGULATION_OF_MICROTUBULE_POLYMERIZATION_OR_DEPOLYMERIZATION	-2.06	-0.27	1.97E-02
GO_CYTOSKELETON_ORGANIZATION	-2.05	-0.12	2.03E-02
GO_MEIOTIC_CELL_CYCLE_PROCESS	-2.05	-0.31	2.02E-02
GO_COFACTOR_BIOSYNTHETIC_PROCESS	-2.02	-0.28	2.52E-02
GO_SPINDLE_ASSEMBLY	-2.01	-0.41	2.55E-02
GO_CELLULAR_ALDEHYDE_METABOLIC_PROCESS	-2.01	-0.41	2.55E-02
GO_ENERGY_DERIVATION_BY_OXIDATION_OF_ORGANIC_COMPOUNDS	-2.00	-0.27	2.66E-02
GO_CELLULAR_PROTEIN_COMPLEX_ASSEMBLY	-1.99	-0.24	2.85E-02
GO_CARBOHYDRATE_DERIVATIVE_METABOLIC_PROCESS	-1.99	-0.12	2.86E-02
GO_TRNA_PROCESSING	-1.97	-0.36	3.11E-02
GO_PROTEIN_POLYMERIZATION	-1.95	-0.38	3.60E-02
GO_MITOTIC_CELL_CYCLE_CHECKPOINT	-1.93	-0.28	3.90E-02
GO_STEROL_METABOLIC_PROCESS	-1.93	-0.29	3.94E-02
GO_GLYCOSYL_COMPOUND_BIOSYNTHETIC_PROCESS	-1.93	-0.38	3.95E-02
GO_PEPTIDYL_AMINO_ACID_MODIFICATION	-1.91	-0.14	4.19E-02
GO_POSITIVE_REGULATION_OF_ORGANELLE_ORGANIZATION	-1.91	-0.16	4.30E-02

GO_COVALENT_CHROMATIN_MODIFICATION	-1.91	-0.25	4.30E-02
GO_NUCLEIC_ACID_PHOSPHODIESTER_BOND_HYDROLYSIS	-1.89	-0.26	4.61E-02
GO_DNA_CONFORMATION_CHANGE	-1.89	-0.20	4.66E-02
GO_NUCLEOSIDE_MONOPHOSPHATE_METABOLIC_PROCESS	-1.88	-0.30	4.93E-02

Reactome pathways overrepresented among upregulated genes

Pathway	***p* value**	***q* value**	**external_id**
Interferon alpha/beta signaling	7.95E-09	5.66E-06	R-HSA-909733
Cytokine Signaling in Immune system	8.89E-07	2.12E-04	R-HSA-1280215
Collagen formation	8.91E-07	2.12E-04	R-HSA-1474290
Extracellular matrix organization	1.69E-06	3.01E-04	R-HSA-1474244
Interferon Signaling	2.42E-06	3.44E-04	R-HSA-913531
Axon guidance	8.50E-06	1.01E-03	R-HSA-422475
Neuronal System	1.86E-05	1.89E-03	R-HSA-112316
Developmental Biology	4.25E-05	3.78E-03	R-HSA-1266738
Collagen biosynthesis and modifying enzymes	6.95E-05	5.50E-03	R-HSA-1650814
Endosomal/Vacuolar pathway	7.80E-05	5.55E-03	R-HSA-1236977
O-glycosylation of TSR domain-containing proteins	1.25E-04	8.12E-03	R-HSA-5173214
MAPK family signaling cascades	4.63E-04	2.75E-02	R-HSA-5683057
Antigen Presentation: Folding, assembly and peptide loading of class I MHC	5.63E-04	2.91E-02	R-HSA-983170
Potassium Channels	5.73E-04	2.91E-02	R-HSA-1296071
RAF/MAP kinase cascade	6.27E-04	2.98E-02	R-HSA-5673001
Assembly of collagen fibrils and other multimeric structures	7.65E-04	3.40E-02	R-HSA-2022090
Interferon gamma signaling	8.66E-04	3.63E-02	R-HSA-877300
Class B/2 (Secretin family receptors)	9.69E-04	3.68E-02	R-HSA-373080
MAPK1/MAPK3 signaling	9.81E-04	3.68E-02	R-HSA-5684996
Peptide hormone biosynthesis	1.30E-03	4.63E-02	R-HSA-209952

Reactome pathways overrepresented among downregulated genes

Pathway	***p* value**	***q* value**	**external_id**
Eukaryotic Translation Elongation	1.29E-08	7.74E-06	R-HSA-156842
Peptide chain elongation	3.40E-08	1.02E-05	R-HSA-156902
Eukaryotic Translation Termination	3.03E-07	4.03E-05	R-HSA-72764
Selenocysteine synthesis	3.03E-07	4.03E-05	R-HSA-2408557
Selenoamino acid metabolism	3.44E-07	4.03E-05	R-HSA-2408522
Nonsense Mediated Decay (NMD) independent of the Exon Junction Complex (EJC)	4.03E-07	4.03E-05	R-HSA-975956
Formation of a pool of free 40S subunits	1.18E-06	1.01E-04	R-HSA-72689
Nonsense Mediated Decay (NMD) enhanced by the Exon Junction Complex (EJC)	1.93E-06	1.29E-04	R-HSA-975957
Nonsense-Mediated Decay (NMD)	1.93E-06	1.29E-04	R-HSA-927802
SRP-dependent cotranslational protein targeting to membrane	3.90E-06	2.12E-04	R-HSA-1799339
L13a-mediated translational silencing of Ceruloplasmin expression	3.90E-06	2.12E-04	R-HSA-156827
GTP hydrolysis and joining of the 60S ribosomal subunit	4.36E-06	2.18E-04	R-HSA-72706
Cap-dependent Translation Initiation	9.25E-06	3.96E-04	R-HSA-72737

Eukaryotic Translation Initiation	9.25E-06	3.96E-04	R-HSA-72613
Formation of the ternary complex, and subsequently, the 43S complex	6.72E-05	2.68E-03	R-HSA-72695
Metabolism of amino acids and derivatives	7.92E-05	2.96E-03	R-HSA-71291
G alpha (12/13) signalling events	1.05E-04	3.69E-03	R-HSA-416482
Semaphorin interactions	1.37E-04	4.55E-03	R-HSA-373755
Translation initiation complex formation	1.78E-04	5.15E-03	R-HSA-72649
Ribosomal scanning and start codon recognition	1.78E-04	5.15E-03	R-HSA-72702
Beta-catenin independent WNT signaling	1.81E-04	5.15E-03	R-HSA-3858494
Activation of the mRNA upon binding of the cap-binding complex and eIFs, and subsequent binding to 43S	2.02E-04	5.51E-03	R-HSA-72662
Signaling by Rho GTPases	2.13E-04	5.54E-03	R-HSA-194315
Axon guidance	5.07E-04	1.24E-02	R-HSA-422475
Cell death signalling via NRAGE, NRIF and NADE	5.74E-04	1.24E-02	R-HSA-204998
Prostacyclin signalling through prostacyclin receptor	5.79E-04	1.24E-02	R-HSA-392851
Other semaphorin interactions	5.79E-04	1.24E-02	R-HSA-416700
Glycosaminoglycan metabolism	5.80E-04	1.24E-02	R-HSA-1630316
G beta:gamma signalling through PLC beta	7.48E-04	1.55E-02	R-HSA-418217
Activation of kainate receptors upon glutamate binding	7.95E-04	1.56E-02	R-HSA-451326
Ca2+ pathway	8.06E-04	1.56E-02	R-HSA-4086398
Presynaptic function of Kainate receptors	9.52E-04	1.78E-02	R-HSA-500657
TFAP2A acts as a transcriptional repressor during retinoic acid induced cell differentiation	1.70E-03	3.09E-02	R-HSA-8869496
Glucagon-like Peptide-1 (GLP1) regulates insulin secretion	1.85E-03	3.25E-02	R-HSA-381676
Metabolism	1.90E-03	3.25E-02	R-HSA-1430728
G0 and Early G1	2.19E-03	3.55E-02	R-HSA-1538133
NOTCH3 Activation and Transmission of Signal to the Nucleus	2.19E-03	3.55E-02	R-HSA-9013507
Death Receptor Signalling	2.28E-03	3.60E-02	R-HSA-73887
Rho GTPase cycle	2.59E-03	3.86E-02	R-HSA-194840
VEGF binds to VEGFR leading to receptor dimerization	2.65E-03	3.86E-02	R-HSA-195399
VEGF ligand-receptor interactions	2.65E-03	3.86E-02	R-HSA-194313
p75 NTR receptor-mediated signalling	3.46E-03	4.93E-02	R-HSA-193704

Appendix IV. Identified BORIS-interacting proteins

Criteria: a minimum detection of 2 peptides, 99% protein threshold, 99% peptide threshold, log2FC>0.6 and p<0.05

Protein ID	Log2FC	*p* value	Ranking score	Known BORIS partner	Shared with CTCF
CTCFL	3.7	1.00E-04	49.16		
HCFC1	2.67	2.00E-04	32.81	X	
SUPT16H	2.6	1.00E-04	34.55		
PDCD11	2.52	2.80E-02	13.00		
PKM	2.36	8.00E-03	16.44		
CDC73	2.06	2.80E-02	10.63		
EIF4A1	1.9	8.00E-03	13.23		
SND1	1.89	1.00E-04	25.11		
IPO7	1.78	1.00E-03	17.74		
BAZ1B	1.76	2.80E-02	9.08		
SMARCA5	1.74	1.00E-04	23.12		
RPS11	1.72	8.00E-03	11.98		
RPL10A	1.68	8.00E-03	11.70		
EIF3C	1.67	2.80E-02	8.61		
CHD4	1.62	2.80E-02	8.36		
KRT18	1.56	1.00E-03	15.55		
NONO	1.53	1.00E-03	15.25		
SSRP1	1.51	1.00E-04	20.06		
MSH6	1.5	1.00E-03	14.95		
XPO1	1.49	1.00E-03	14.85		
SFXN1	1.47	8.00E-03	10.24		
TRAP1	1.47	8.00E-03	10.24		
HNRNPH2	1.4	2.80E-02	7.22		
KRT8	1.4	2.80E-02	7.22		
PRPF6	1.39	2.80E-02	7.17		
EIF3A	1.38	1.00E-03	13.75		
RPL30	1.37	2.80E-02	7.07		
SLC25A11	1.37	2.80E-02	7.07		
HSPD1	1.37	1.00E-04	18.20		
NOP14	1.34	8.00E-03	9.33		
RPN1	1.32	8.40E-04	13.49		
SF3A3	1.3	2.80E-02	6.71		
SENP3	1.29	2.80E-02	6.65		
RPL9	1.27	2.80E-02	6.55		
PRKDC	1.26	1.00E-04	16.74		
HSP90AA1	1.26	1.00E-04	16.74		
RPL7A	1.25	1.00E-04	16.61		

DDX3X	1.24	1.00E-03	12.36		
PTBP1	1.22	8.00E-03	8.50		
RPL15	1.22	8.00E-03	8.50		
PARP1	1.21	1.00E-04	16.08		X
GATAD2A	1.18	2.80E-02	6.09		
FASN	1.17	9.80E-04	11.69		
LRRC59	1.17	2.80E-02	6.04		
MYBBP1A	1.17	1.00E-04	15.55		
VDAC2	1.15	8.00E-03	8.01		
EPRS	1.13	1.00E-04	15.02		
DDX52	1.07	2.80E-02	5.52		
CLTC	1.06	3.30E-04	12.26		
DDX46	1.06	1.00E-03	10.56		
UBTF	1.06	1.00E-03	10.56	X	X
CAD	1.05	2.80E-02	5.42		
SF3B1	1.04	6.00E-03	7.68		
NOP2	1.03	1.00E-03	10.26		
COPA	1.03	2.30E-02	5.61		
XRCC6	1.02	1.00E-04	13.55		
EIF4A3	1.01	2.80E-02	5.21		
RBM14	0.99	8.00E-03	6.90		
EIF3I	0.99	2.80E-02	5.11		
PCBP1	0.98	2.80E-02	5.06		
TMPO	0.97	2.80E-02	5.00		
PDIA6	0.97	2.80E-02	5.00		
MSH2	0.97	2.80E-02	5.00		
LRPPRC	0.96	1.00E-04	12.76		
DDOST	0.95	8.00E-03	6.62		
TCP1	0.86	1.00E-03	8.57		
MATR3	0.86	5.00E-03	6.57		
CKAP4	0.86	2.80E-02	4.44		
PRPF8	0.83	3.10E-04	9.67		
G3BP1	0.83	2.80E-02	4.28		
RAD50	0.82	2.80E-02	4.23		
XRCC5	0.81	1.00E-03	8.07		
ALDH18A1	0.81	2.80E-02	4.18		
RPL5	0.8	1.00E-03	7.97		
VIM	0.79	1.00E-04	10.50		
RPLP0	0.77	8.00E-03	5.36		
RARS	0.76	1.00E-03	7.57		
IARS	0.76	2.80E-02	3.92		

RPF2	0.74	2.80E-02	3.82		
PHB2	0.7	6.00E-03	5.17		
SMC1A	0.68	4.20E-03	5.37		X
DNTTIP2	0.63	1.00E-03	6.28		
SNRNP200	0.6	4.00E-03	4.78		
NOP56	0.6	8.00E-03	4.18		

Appendix V. Protein complexes among BORIS-interacting proteins identified by MIST

Complex ID	Complex Name	Entrez Gene ID	Gene In Search Results
HC1026	eIF3	8663	EIF3C
HC1026	eIF3	8668	EIF3I
HC1026	eIF3	8661	EIF3A
HC1045	MLL1-WDR5	26168	SENP3
HC1045	MLL1-WDR5	3054	HCFC1
HC1079	BRCA1/BARD1	2956	MSH6
HC1079	BRCA1/BARD1	8243	SMC1A
HC108	SNF2h-cohesin-NuRD complex	8243	SMC1A
HC108	SNF2h-cohesin-NuRD complex	8467	SMARCA5
HC1090	generation of neurons	6175	RPLP0
HC1090	generation of neurons	790	CAD
HC1100	ICEN	11198	SUPT16H
HC1100	ICEN	8467	SMARCA5
HC1100	ICEN	6749	SSRP1
HC1145	MMS22L-NFKBIL2	11198	SUPT16H
HC1145	MMS22L-NFKBIL2	6749	SSRP1
HC1153	DHX9-ADAR-vigilin-DNA-PK-Ku antigen complex	5591	PRKDC
HC1153	DHX9-ADAR-vigilin-DNA-PK-Ku antigen complex	2547	XRCC6
HC1153	DHX9-ADAR-vigilin-DNA-PK-Ku antigen complex	7520	XRCC5
HC1170	eIF3 complex (EIF3S6, EIF3S5, EIF3S4, EIF3S3, EIF3S6IP, EIF3S2, EIF3S9, EIF3S12, EIF3S10, EIF3S8, EIF3S1, EIF3S7)	8661	EIF3A
HC1170	eIF3 complex (EIF3S6, EIF3S5, EIF3S4, EIF3S3, EIF3S6IP, EIF3S2, EIF3S9, EIF3S12, EIF3S10, EIF3S8, EIF3S1, EIF3S7)	8668	EIF3I
HC1170	eIF3 complex (EIF3S6, EIF3S5, EIF3S4, EIF3S3, EIF3S6IP, EIF3S2, EIF3S9, EIF3S12, EIF3S10, EIF3S8, EIF3S1, EIF3S7)	8663	EIF3C
HC1267	nucleosome assembly	9031	BAZ1B
HC1267	nucleosome assembly	8467	SMARCA5
HC1278	MLL1 complex	26168	SENP3
HC1278	MLL1 complex	3054	HCFC1
HC129	transcription elongation from RNA polymerase II promoter	6749	SSRP1
HC129	transcription elongation from RNA polymerase II promoter	11198	SUPT16H
HC1300	C complex spliceosome	10946	SF3A3
HC1300	C complex spliceosome	10594	PRPF8
HC1300	C complex spliceosome	24148	PRPF6
HC1300	C complex spliceosome	9775	EIF4A3
HC1300	C complex spliceosome	23451	SF3B1
HC1300	C complex spliceosome	23020	SNRNP200
HC1322	MutS-alpha complex	4436	MSH2
HC1322	MutS-alpha complex	2956	MSH6
HC1329	nuclear-transcribed mRNA poly(A) tail shortening	1973	EIF4A1
HC1329	nuclear-transcribed mRNA poly(A) tail shortening	9775	EIF4A3
HC1357	Rag1-Rag2-Ku70-Ku80 protein-DNA complex	2547	XRCC6
HC1357	Rag1-Rag2-Ku70-Ku80 protein-DNA complex	7520	XRCC5

HC137	translational initiation	8661	EIF3A
HC137	translational initiation	8663	EIF3C
HC1410	53BP1-containing complex	7520	XRCC5
HC1410	53BP1-containing complex	2547	XRCC6
HC1442	tRNA aminoacylation for protein translation	3376	IARS
HC1442	tRNA aminoacylation for protein translation	5917	RARS
HC1442	tRNA aminoacylation for protein translation	2058	EPRS
HC1513	CEN complex	11198	SUPT16H
HC1513	CEN complex	6749	SSRP1
HC1513	CEN complex	8467	SMARCA5
HC1574	U2 snRNP complex	10946	SF3A3
HC1574	U2 snRNP complex	23451	SF3B1
HC1608	WICH complex	9031	BAZ1B
HC1608	WICH complex	8467	SMARCA5
HC161	Vigilin-DNA-PK-Ku antigen complex	5591	PRKDC
HC161	Vigilin-DNA-PK-Ku antigen complex	2547	XRCC6
HC161	Vigilin-DNA-PK-Ku antigen complex	7520	XRCC5
HC1620	BRCA1-associated genome surveillance complex (BASC)	4436	MSH2
HC1620	BRCA1-associated genome surveillance complex (BASC)	2956	MSH6
HC1639	heterogeneous nuclear ribonucleoprotein complex	3188	HNRNPH2
HC1639	heterogeneous nuclear ribonucleoprotein complex	5725	PTBP1
HC1645	EIF3 core complex (EIF3A, EIF3B, EIF3G, EIF3I)	8668	EIF3I
HC1645	EIF3 core complex (EIF3A, EIF3B, EIF3G, EIF3I)	8661	EIF3A
HC1682	RNA metabolic process	6205	RPS11
HC1682	RNA metabolic process	10528	NOP56
HC1759	HCF-1N	3054	HCFC1
HC1759	HCF-1N	3320	HSP90AA1
HC176	DNA-PK-Ku-eIF2-NF90-NF45 complex	7520	XRCC5
HC176	DNA-PK-Ku-eIF2-NF90-NF45 complex	2547	XRCC6
HC176	DNA-PK-Ku-eIF2-NF90-NF45 complex	5591	PRKDC
HC1767	DNA-dependent protein kinase-DNA ligase 4 complex	5591	PRKDC
HC1767	DNA-dependent protein kinase-DNA ligase 4 complex	2547	XRCC6
HC1767	DNA-dependent protein kinase-DNA ligase 4 complex	7520	XRCC5
HC1815	EIF3 complex (EIF3A, EIF3B, EIF3G, EIF3I, EIF3C)	8661	EIF3A
HC1815	EIF3 complex (EIF3A, EIF3B, EIF3G, EIF3I, EIF3C)	8663	EIF3C
HC1815	EIF3 complex (EIF3A, EIF3B, EIF3G, EIF3I, EIF3C)	8668	EIF3I
HC1832	Spliceosome, U2-snRNP	10946	SF3A3
HC1832	Spliceosome, U2-snRNP	23451	SF3B1
HC1907	CDC5L complex	5591	PRKDC
HC1907	CDC5L complex	23451	SF3B1
HC1922	Ku antigen-NARG1 complex	2547	XRCC6
HC1922	Ku antigen-NARG1 complex	7520	XRCC5
HC1926	regulation of transcription, DNA-dependent	7520	XRCC5
HC1926	regulation of transcription, DNA-dependent	2547	XRCC6
HC1926	regulation of transcription, DNA-dependent	142	PARP1
HC1929	NuRD complex	1108	CHD4
HC1929	NuRD complex	54815	GATAD2A

HC197	Ribosome, cytoplasmic	6205	RPS11
HC197	Ribosome, cytoplasmic	6125	RPL5
HC197	Ribosome, cytoplasmic	6138	RPL15
HC197	Ribosome, cytoplasmic	6130	RPL7A
HC197	Ribosome, cytoplasmic	6175	RPLP0
HC197	Ribosome, cytoplasmic	6133	RPL9
HC197	Ribosome, cytoplasmic	6156	RPL30
HC197	Ribosome, cytoplasmic	4736	RPL10A
HC1974	eIF4F	1973	EIF4A1
HC1974	eIF4F	9775	EIF4A3
HC1997	multi-eIF complex	8668	EIF3I
HC1997	multi-eIF complex	8663	EIF3C
HC1998	TERF2-RAP1 complex	7520	XRCC5
HC1998	TERF2-RAP1 complex	2547	XRCC6
HC2009	SNF2h/cohesin	1108	CHD4
HC2009	SNF2h/cohesin	8467	SMARCA5
HC2009	SNF2h/cohesin	8243	SMC1A
HC2014	H3.3 com	142	PARP1
HC2014	H3.3 com	7520	XRCC5
HC2014	H3.3 com	5591	PRKDC
HC2023	rRNA processing	10514	MYBBP1A
HC2023	rRNA processing	10528	NOP56
HC206	BASC (Ab 80) complex (BRCA1-associated genome surveillance complex)	4436	MSH2
HC206	BASC (Ab 80) complex (BRCA1-associated genome surveillance complex)	2956	MSH6
HC2067	Ribosome, archaea	6130	RPL7A
HC2067	Ribosome, archaea	6156	RPL30
HC2067	Ribosome, archaea	6138	RPL15
HC2069	transcription, DNA-dependent	1108	CHD4
HC2069	transcription, DNA-dependent	6749	SSRP1
HC2069	transcription, DNA-dependent	79577	CDC73
HC2069	transcription, DNA-dependent	11198	SUPT16H
HC209	TLE1 corepressor complex (MASH1 promoter-corepressor complex)	142	PARP1
HC209	TLE1 corepressor complex (MASH1 promoter-corepressor complex)	4841	NONO
HC213	MMS22L-TONSL	6749	SSRP1
HC213	MMS22L-TONSL	11198	SUPT16H
HC2134	BRCA1-TRRAP/hGCN5	2956	MSH6
HC2134	BRCA1-TRRAP/hGCN5	4436	MSH2
HC2175	nonhomologous end joining complex	2547	XRCC6
HC2175	nonhomologous end joining complex	7520	XRCC5
HC2175	nonhomologous end joining complex	5591	PRKDC
HC2329	protein transport	3320	HSP90AA1
HC2329	protein transport	7417	VDAC2
HC2329	protein transport	7514	XPO1

HC240	Aminoacyl-tRNA biosynthesis, eukaryotes	5917	RARS
HC240	Aminoacyl-tRNA biosynthesis, eukaryotes	2058	EPRS
HC240	Aminoacyl-tRNA biosynthesis, eukaryotes	3376	IARS
HC2429	TRF2-Ku complex	7520	XRCC5
HC2429	TRF2-Ku complex	2547	XRCC6
HC2461	FACT complex, UV-activated	6749	SSRP1
HC2461	FACT complex, UV-activated	11198	SUPT16H
HC2497	18S U11/U12 snRNP	10594	PRPF8
HC2497	18S U11/U12 snRNP	23451	SF3B1
HC250	Rap1 complex	2547	XRCC6
HC250	Rap1 complex	142	PARP1
HC250	Rap1 complex	7520	XRCC5
HC2500	Oligosaccharyltransferase complex (Stt3B variant)	1650	DDOST
HC2500	Oligosaccharyltransferase complex (Stt3B variant)	6184	RPN1
HC2502	CAMK2-delta-MASH1 promoter-coactivator complex	4841	NONO
HC2502	CAMK2-delta-MASH1 promoter-coactivator complex	142	PARP1
HC2527	oligosaccharyltransferase complex	1650	DDOST
HC2527	oligosaccharyltransferase complex	6184	RPN1
HC2529	TLE1 corepressor complex (MASH1 promoter-corepressor complex)	142	PARP1
HC2529	TLE1 corepressor complex (MASH1 promoter-corepressor complex)	4841	NONO
HC2578	Ku70/Ku86/Werner complex	2547	XRCC6
HC2578	Ku70/Ku86/Werner complex	7520	XRCC5
HC2579	p54(nrb)-PSF-matrin3 complex	9782	MATR3
HC2579	p54(nrb)-PSF-matrin3 complex	4841	NONO
HC2657	SNW1 complex	4436	MSH2
HC2657	SNW1 complex	9782	MATR3
HC2657	SNW1 complex	23020	SNRNP200
HC2657	SNW1 complex	10594	PRPF8
HC2662	MutS-alpha-PK-zeta complex	2956	MSH6
HC2662	MutS-alpha-PK-zeta complex	4436	MSH2
HC2708	60S ribosomal subunit, cytoplasmic	4736	RPL10A
HC2708	60S ribosomal subunit, cytoplasmic	6175	RPLP0
HC2708	60S ribosomal subunit, cytoplasmic	6130	RPL7A
HC2708	60S ribosomal subunit, cytoplasmic	6138	RPL15
HC2708	60S ribosomal subunit, cytoplasmic	6156	RPL30
HC2708	60S ribosomal subunit, cytoplasmic	6125	RPL5
HC2708	60S ribosomal subunit, cytoplasmic	6133	RPL9
HC2764	Parvulin-associated pre-rRNP complex	6138	RPL15
HC2764	Parvulin-associated pre-rRNP complex	6175	RPLP0
HC2764	Parvulin-associated pre-rRNP complex	10514	MYBBP1A
HC2764	Parvulin-associated pre-rRNP complex	6130	RPL7A
HC2764	Parvulin-associated pre-rRNP complex	4839	NOP2
HC2764	Parvulin-associated pre-rRNP complex	10528	NOP56
HC2764	Parvulin-associated pre-rRNP complex	4736	RPL10A
HC2782	Oligosaccharyltransferase	1650	DDOST
HC2782	Oligosaccharyltransferase	6184	RPN1
HC2785	B-WICH	9031	BAZ1B

HC2785	B-WICH	10514	MYBBP1A
HC2785	B-WICH	23451	SF3B1
HC2792	Spliceosome	10594	PRPF8
HC2792	Spliceosome	23451	SF3B1
HC2792	Spliceosome	10946	SF3A3
HC2792	Spliceosome	23020	SNRNP200
HC2792	Spliceosome	1654	DDX3X
HC2792	Spliceosome	9879	DDX46
HC2792	Spliceosome	9775	EIF4A3
HC2792	Spliceosome	24148	PRPF6
HC2796	PCNA-MutS-alpha-MutL-alpha-DNA complex	2956	MSH6
HC2796	PCNA-MutS-alpha-MutL-alpha-DNA complex	4436	MSH2
HC280	FACT-Nek9	11198	SUPT16H
HC280	FACT-Nek9	6749	SSRP1
HC2808	WRN-Ku70-Ku80-PARP1 complex	142	PARP1
HC2808	WRN-Ku70-Ku80-PARP1 complex	2547	XRCC6
HC2808	WRN-Ku70-Ku80-PARP1 complex	7520	XRCC5
HC290	DNA-PK-Ku complex	5591	PRKDC
HC290	DNA-PK-Ku complex	7520	XRCC5
HC290	DNA-PK-Ku complex	2547	XRCC6
HC291	cytoplasmic ribosomal large subunit	6175	RPLP0
HC291	cytoplasmic ribosomal large subunit	6133	RPL9
HC291	cytoplasmic ribosomal large subunit	6130	RPL7A
HC291	cytoplasmic ribosomal large subunit	4736	RPL10A
HC291	cytoplasmic ribosomal large subunit	6156	RPL30
HC291	cytoplasmic ribosomal large subunit	6138	RPL15
HC291	cytoplasmic ribosomal large subunit	6125	RPL5
HC2914	MutS-alpha-histone H4 complex	4436	MSH2
HC2914	MutS-alpha-histone H4 complex	2956	MSH6
HC2923	Yph1-L	6125	RPL5
HC2923	Yph1-L	6130	RPL7A
HC2930	EIF3 complex (EIF3A, EIF3B, EIF3G, EIF3I, EIF3J)	8668	EIF3I
HC2930	EIF3 complex (EIF3A, EIF3B, EIF3G, EIF3I, EIF3J)	8661	EIF3A
HC2935	HCF-1 complex	3320	HSP90AA1
HC2935	HCF-1 complex	3054	HCFC1
HC2992	primary metabolic process	2194	FASN
HC2992	primary metabolic process	6184	RPN1
HC301	Spliceosome, 35S U5-snRNP	23020	SNRNP200
HC301	Spliceosome, 35S U5-snRNP	10594	PRPF8
HC3021	H2AX complex, isolated from cells without IR exposure	6749	SSRP1
HC3021	H2AX complex, isolated from cells without IR exposure	11198	SUPT16H
HC304	eukaryotic translation initiation factor 3 complex	8663	EIF3C
HC304	eukaryotic translation initiation factor 3 complex	8661	EIF3A
HC304	eukaryotic translation initiation factor 3 complex	8668	EIF3I
HC3041	nuclear mRNA splicing, via spliceosome	10946	SF3A3
HC3041	nuclear mRNA splicing, via spliceosome	10594	PRPF8
HC3041	nuclear mRNA splicing, via spliceosome	24148	PRPF6
HC3041	nuclear mRNA splicing, via spliceosome	23451	SF3B1
HC3041	nuclear mRNA splicing, via spliceosome	23020	SNRNP200

HC31	TNF-alpha/Nf-kappa B signaling complex (RPL6, RPL30, RPS13, CHUK, DDX3X, NFKB2, NFKBIB, REL, IKBKG, NFKB1, MAP3K8, RELB, GLG1, NFKBIA, RELA, TNIP2, GTF2I)	1654	DDX3X
HC31	TNF-alpha/Nf-kappa B signaling complex (RPL6, RPL30, RPS13, CHUK, DDX3X, NFKB2, NFKBIB, REL, IKBKG, NFKB1, MAP3K8, RELB, GLG1, NFKBIA, RELA, TNIP2, GTF2I)	6156	RPL30
HC3120	U5 snRNP complex	10594	PRPF8
HC3120	U5 snRNP complex	23020	SNRNP200
HC3157	TRF2.RAP1	7520	XRCC5
HC3157	TRF2.RAP1	2547	XRCC6
HC3164	ALL-1	1108	CHD4
HC3164	ALL-1	8467	SMARCA5
HC3185	TRBP containing complex (DICER, RPL7A, EIF6, MOV10 and subunits of the 60S ribosomal particle)	6130	RPL7A
HC3185	TRBP containing complex (DICER, RPL7A, EIF6, MOV10 and subunits of the 60S ribosomal particle)	4736	RPL10A
HC3185	TRBP containing complex (DICER, RPL7A, EIF6, MOV10 and subunits of the 60S ribosomal particle)	6138	RPL15
HC3185	TRBP containing complex (DICER, RPL7A, EIF6, MOV10 and subunits of the 60S ribosomal particle)	6156	RPL30
HC3198	ISW2	9031	BAZ1B
HC3198	ISW2	8467	SMARCA5
HC320	spliceosomal complex	10946	SF3A3
HC320	spliceosomal complex	24148	PRPF6
HC320	spliceosomal complex	23451	SF3B1
HC320	spliceosomal complex	23020	SNRNP200
HC3215	nuclear telomere cap complex	2547	XRCC6
HC3215	nuclear telomere cap complex	7520	XRCC5
HC3287	Oligosaccharyltransferase complex (Stt3A variant)	1650	DDOST
HC3287	Oligosaccharyltransferase complex (Stt3A variant)	6184	RPN1
HC3289	MSH2-MSH6-PMS1-MLH1 complex	2956	MSH6
HC3289	MSH2-MSH6-PMS1-MLH1 complex	4436	MSH2
HC3298	SPN	6749	SSRP1
HC3298	SPN	11198	SUPT16H
HC3319	Aminoacyl-tRNA biosynthesis, prokaryotes	5917	RARS
HC3319	Aminoacyl-tRNA biosynthesis, prokaryotes	3376	IARS
HC3326	isotype switching	2956	MSH6
HC3326	isotype switching	4436	MSH2
HC3331	DNA double-strand break end-joining complex	7520	XRCC5
HC3331	DNA double-strand break end-joining complex	2547	XRCC6
HC3335	MASH1 promoter-coactivator complex	142	PARP1
HC3335	MASH1 promoter-coactivator complex	4841	NONO
HC3347	17S U2 snRNP	9879	DDX46
HC3347	17S U2 snRNP	23451	SF3B1
HC3347	17S U2 snRNP	10946	SF3A3
HC3388	translational initiation	6205	RPS11
HC3388	translational initiation	8668	EIF3I

HC3388	translational initiation	1654	DDX3X
HC3388	translational initiation	8661	EIF3A
HC3388	translational initiation	8663	EIF3C
HC3420	Ribosome, eukaryotes	6205	RPS11
HC3420	Ribosome, eukaryotes	6125	RPL5
HC3420	Ribosome, eukaryotes	6138	RPL15
HC3420	Ribosome, eukaryotes	6175	RPLP0
HC3420	Ribosome, eukaryotes	6130	RPL7A
HC3420	Ribosome, eukaryotes	6133	RPL9
HC3420	Ribosome, eukaryotes	6156	RPL30
HC3420	Ribosome, eukaryotes	4736	RPL10A
HC3437	transcription factor complex	2547	XRCC6
HC3437	transcription factor complex	142	PARP1
HC3437	transcription factor complex	5591	PRKDC
HC3437	transcription factor complex	10432	RBM14
HC3470	rRNA metabolic process	1108	CHD4
HC3470	rRNA metabolic process	10528	NOP56
HC3520	mRNA metabolic process	6156	RPL30
HC3520	mRNA metabolic process	10528	NOP56
HC3520	mRNA metabolic process	6175	RPLP0
HC3520	mRNA metabolic process	4839	NOP2
HC3520	mRNA metabolic process	6133	RPL9
HC3520	mRNA metabolic process	4736	RPL10A
HC3520	mRNA metabolic process	1654	DDX3X
HC3520	mRNA metabolic process	6130	RPL7A
HC3520	mRNA metabolic process	6138	RPL15
HC3520	mRNA metabolic process	6205	RPS11
HC3520	mRNA metabolic process	6125	RPL5
HC3521	BASC complex (BRCA1-associated genome surveillance complex)	2956	MSH6
HC3521	BASC complex (BRCA1-associated genome surveillance complex)	4436	MSH2
HC3559	TNF-alpha/NF-kappa B signaling complex 6	3320	HSP90AA1
HC3559	TNF-alpha/NF-kappa B signaling complex 6	6156	RPL30
HC3559	TNF-alpha/NF-kappa B signaling complex 6	6205	RPS11
HC359	oligosaccharyl transferase complex	6184	RPN1
HC359	oligosaccharyl transferase complex	1650	DDOST
HC3599	U4/U6 x U5 tri-snRNP complex	24148	PRPF6
HC3599	U4/U6 x U5 tri-snRNP complex	23020	SNRNP200
HC3599	U4/U6 x U5 tri-snRNP complex	10594	PRPF8
HC369	PCNA-KU antigen complex	2547	XRCC6
HC369	PCNA-KU antigen complex	7520	XRCC5
HC393	ribonucleoprotein complex	3188	HNRNPH2
HC393	ribonucleoprotein complex	6175	RPLP0
HC393	ribonucleoprotein complex	6125	RPL5
HC393	ribonucleoprotein complex	5093	PCBP1
HC393	ribonucleoprotein complex	10432	RBM14
HC393	ribonucleoprotein complex	7514	XPO1
HC414	NCOA6-DNA-PK-Ku-PARP1 complex	142	PARP1
HC414	NCOA6-DNA-PK-Ku-PARP1 complex	5591	PRKDC

HC414	NCOA6-DNA-PK-Ku-PARP1 complex	2547	XRCC6
HC414	NCOA6-DNA-PK-Ku-PARP1 complex	7520	XRCC5
HC426	carboxylic acid metabolic process	6749	SSRP1
HC426	carboxylic acid metabolic process	27044	SND1
HC429	Ku antigen complex	7520	XRCC5
HC429	Ku antigen complex	2547	XRCC6
HC444	PCNA-MSH2-MSH6 complex	2956	MSH6
HC444	PCNA-MSH2-MSH6 complex	4436	MSH2
HC464	Ku antigen-YY1-alphaMyHC promoter complex	2547	XRCC6
HC464	Ku antigen-YY1-alphaMyHC promoter complex	7520	XRCC5
HC485	FACT complex	6749	SSRP1
HC485	FACT complex	11198	SUPT16H
HC572	Nop56p-associated pre-rRNA complex	6125	RPL5
HC572	Nop56p-associated pre-rRNA complex	6205	RPS11
HC572	Nop56p-associated pre-rRNA complex	10514	MYBBP1A
HC572	Nop56p-associated pre-rRNA complex	6130	RPL7A
HC572	Nop56p-associated pre-rRNA complex	6175	RPLP0
HC572	Nop56p-associated pre-rRNA complex	6133	RPL9
HC572	Nop56p-associated pre-rRNA complex	6156	RPL30
HC572	Nop56p-associated pre-rRNA complex	10528	NOP56
HC572	Nop56p-associated pre-rRNA complex	6138	RPL15
HC572	Nop56p-associated pre-rRNA complex	4736	RPL10A
HC575	Toposome	10594	PRPF8
HC575	Toposome	6749	SSRP1
HC59	MSH2-MSH6-PMS2-MLH1 complex	2956	MSH6
HC59	MSH2-MSH6-PMS2-MLH1 complex	4436	MSH2
HC71	MGC1-DNA-PKcs-Ku complex	7520	XRCC5
HC71	MGC1-DNA-PKcs-Ku complex	5591	PRKDC
HC71	MGC1-DNA-PKcs-Ku complex	2547	XRCC6
HC745	MLL1-WDR5 complex	26168	SENP3
HC745	MLL1-WDR5 complex	3054	HCFC1
HC750	eIF3 complex (EIF3S6, EIF3S5, EIF3S4, EIF3S3, EIF3S6IP, EIF3S2, EIF3S9, EIF3S12, EIF3S10, EIF3S8, EIF3S1, EIF3S7, PCID1)	8668	EIF3I
HC750	eIF3 complex (EIF3S6, EIF3S5, EIF3S4, EIF3S3, EIF3S6IP, EIF3S2, EIF3S9, EIF3S12, EIF3S10, EIF3S8, EIF3S1, EIF3S7, PCID1)	8661	EIF3A
HC750	eIF3 complex (EIF3S6, EIF3S5, EIF3S4, EIF3S3, EIF3S6IP, EIF3S2, EIF3S9, EIF3S12, EIF3S10, EIF3S8, EIF3S1, EIF3S7, PCID1)	8663	EIF3C
HC812	Spliceosome, U4/U6.U5 tri-snRNP	23020	SNRNP200
HC812	Spliceosome, U4/U6.U5 tri-snRNP	10594	PRPF8
HC812	Spliceosome, U4/U6.U5 tri-snRNP	24148	PRPF6
HC864	H3.1 com	5591	PRKDC
HC864	H3.1 com	142	PARP1
HC864	H3.1 com	7520	XRCC5
HC883	ISW2/yCHRAC	8467	SMARCA5
HC883	ISW2/yCHRAC	9031	BAZ1B
HC916	MSH2/6-BLM-p53-RAD51 complex	2956	MSH6

HC916	MSH2/6-BLM-p53-RAD51 complex	4436	MSH2
HC942	WINAC complex	11198	SUPT16H
HC942	WINAC complex	9031	BAZ1B
HC943	WICH	9031	BAZ1B
HC943	WICH	8467	SMARCA5
HC953	Multisynthetase complex	2058	EPRS
HC953	Multisynthetase complex	3376	IARS
HC953	Multisynthetase complex	5917	RARS
HC963	Ku-ORC complex	7520	XRCC5
HC963	Ku-ORC complex	2547	XRCC6

Appendix VI. Known CTCF-interacting partners

Known CTCF partner
UBTF
Oct4
LS Pol II (POLR2A)
HDLBP (Vigilin)
Kaiso (ZBTB33)
TFII-I
NPM (B23)
YBX1 (YB1)
YY1
TopoII
PARP1
CIITA
SIN3A
CHD8
Suz12
SMC1 (SMC1A)
SMC3
STAG1 (SA1)*
STAG2 (SA2)*
RAD21 (SCC1)
CSB
NCL
SMC2
SMC4
BPTF
LMNA (Lamin A/C)**
KPNA2**
KPNA4**
H2AZ**
TAF1 (TFIID)**
DDX5***
RFX5***

* Direct interacting protein within Cohesin
** Identified by MS
*** Indirect interacting protein

Appendix VII. Overrepresentation analysis for biological processes and pathways on the identified BORIS-interacting proteins

PANTHER GO-Slim Biological processes
#1 number of genes in the reference (Human)
#2 number of genes from the client textbox input
#3 number of expected genes
#4 fold enrichment

Biological Process	#1	#2	#3	#4	+/-	Raw *p* value	FDR	GO ID
tRNA aminoacylation for protein translation	44	3	0.19	16.12	+	1.02E-03	1.32E-02	GO:0006418
translation	198	8	0.84	9.55	+	2.56E-06	6.93E-05	GO:0007126
protein metabolic process	1583	16	6.7	2.39	+	9.63E-04	1.31E-02	GO:0006310
primary metabolic process	4753	45	20.1	2.24	+	1.13E-08	2.75E-06	GO:0000375
metabolic process	5878	48	24.86	1.93	+	3.17E-07	1.29E-05	GO:0000398
meiosis	62	4	0.26	15.25	+	1.74E-04	2.65E-03	GO:0016072
cellular process	8247	50	34.88	1.43	+	1.47E-03	1.80E-02	GO:0006457
DNA recombination	52	3	0.22	13.64	+	1.62E-03	1.79E-02	GO:0006397
DNA metabolic process	391	10	1.65	6.05	+	7.25E-06	1.77E-04	GO:0006412
nucleobase-containing compound metabolic process	2797	32	11.83	2.7	+	6.08E-08	4.95E-06	GO:0006281
RNA splicing, via transesterification reactions	156	8	0.66	12.12	+	4.60E-07	1.60E-05	GO:0051169
RNA metabolic process	1570	24	6.64	3.61	+	2.60E-08	3.18E-06	GO:0006259
mRNA splicing, via spliceosome	178	9	0.75	11.95	+	9.49E-08	5.79E-06	GO:0006260
mRNA processing	244	10	1.03	9.69	+	1.17E-07	5.69E-06	GO:0006520
rRNA metabolic process	113	5	0.48	10.46	+	1.42E-04	2.32E-03	GO:0044085
protein folding	94	4	0.4	10.06	+	7.85E-04	1.13E-02	GO:0016070
DNA repair	167	6	0.71	8.49	+	9.15E-05	1.72E-03	GO:0006950
nuclear transport	123	4	0.52	7.69	+	2.05E-03	2.08E-02	GO:0006139
DNA replication	157	4	0.66	6.02	+	4.80E-03	4.34E-02	GO:0006807
cellular amino acid metabolic process	230	5	0.97	5.14	+	3.17E-03	3.10E-02	GO:0071840
cellular component biogenesis	776	13	3.28	3.96	+	2.54E-05	5.17E-04	GO:0009058
cellular component organization or biogenesis	2099	23	8.88	2.59	+	1.63E-05	3.61E-04	GO:0019538
response to stress	653	9	2.76	3.26	+	1.88E-03	2.00E-02	GO:0044238
nitrogen compound metabolic process	2524	28	10.68	2.62	+	1.07E-06	3.27E-05	GO:0016043
biosynthetic process	1745	19	7.38	2.57	+	1.14E-04	1.99E-03	GO:0008152
cellular component organization	1964	18	8.31	2.17	+	1.52E-03	1.77E-02	GO:0009987
cell communication	2686	3	11.36	0.26	-	3.89E-03	3.65E-02	GO:0007154

+ is overrepresented, - is underrepresented

REViGO on identified biological processes

GO ID	Description	Frequency	log10 *p* value	Uniqueness	Dispensability	Eliminated
GO:0006397	mRNA processing	2.78%	-5.24	0.36	0.00	0
GO:0007126	meiotic nuclear division	1.09%	-2.58	0.48	0.23	0
GO:0006457	protein folding	1.40%	-1.95	0.48	0.23	0
GO:0051169	nuclear transport	2.75%	-1.68	0.46	0.24	0
GO:0016072	rRNA metabolic process	1.52%	-2.63	0.42	0.31	0
GO:0006281	DNA repair	2.98%	-2.76	0.36	0.33	0
GO:0006412	translation	3.69%	-4.16	0.37	0.34	0
GO:0006520	cellular amino acid metabolic process	2.15%	-1.51	0.41	0.43	0
GO:0044238	primary metabolic process	60.31%	-5.56	0.20	0.44	0
GO:0016070	RNA metabolic process	26.92%	-5.50	0.22	0.71	1
GO:0019538	protein metabolic process	31.99%	-1.88	0.24	0.75	1
GO:0009987	cellular process	90.33%	-1.74	0.21	0.81	1
GO:0009058	biosynthetic process	36.92%	-2.70	0.24	0.78	1
GO:0006139	nucleobase-containing compound metabolic process	33.60%	-5.31	0.20	0.89	1
GO:0008152	metabolic process	64.84%	-4.89	0.24	0.71	1
GO:0006807	nitrogen compound metabolic process	40.08%	-4.49	0.24	0.81	1
GO:0006418	tRNA aminoacylation for protein translation	0.28%	-1.88	0.47	0.47	0
GO:0006259	DNA metabolic process	5.66%	-3.75	0.35	0.48	0
GO:0044085	cellular component biogenesis	17.15%	-3.29	0.33	0.50	0
GO:0016043	cellular component organization	35.68%	-1.75	0.27	0.75	1
GO:0006950	response to stress	21.31%	-1.70	0.33	0.53	0
GO:0006310	DNA recombination	1.57%	-1.75	0.40	0.55	0
GO:0006260	DNA replication	1.71%	-1.36	0.40	0.56	0
GO:0071840	cellular component organization or biogenesis	36.53%	-3.44	0.29	0.60	0
GO:0007154	cell communication	36.71%	-1.44	0.28	0.60	0
GO:0000398	mRNA splicing, via spliceosome	1.84%	-5.24	0.37	0.68	0
GO:0000375	RNA splicing, via transesterification reactions	1.86%	-4.80	0.38	0.81	1

Reactome pathways

pathway	external_id	*p* value	*q* value
Metabolism of RNA	R-HSA-8953854	2.62E-16	1.41E-14
Translation	R-HSA-72766	2.84E-12	7.66E-11
L13a-mediated translational silencing of Ceruloplasmin expression	R-HSA-156827	8.57E-12	1.27E-10
GTP hydrolysis and joining of the 60S ribosomal subunit	R-HSA-72706	9.44E-12	1.27E-10
Cap-dependent Translation Initiation	R-HSA-72737	1.82E-11	1.64E-10
Eukaryotic Translation Initiation	R-HSA-72613	1.82E-11	1.64E-10
Formation of a pool of free 40S subunits	R-HSA-72689	6.92E-11	5.33E-10

Selenoamino acid metabolism	R-HSA-2408522	2.69E-10	1.82E-09
SRP-dependent cotranslational protein targeting to membrane	R-HSA-1799339	3.19E-09	1.91E-08
Nonsense Mediated Decay (NMD) enhanced by the Exon Junction Complex (EJC)	R-HSA-975957	3.35E-08	1.64E-07
Nonsense-Mediated Decay (NMD)	R-HSA-927802	3.35E-08	1.64E-07
mRNA Splicing - Major Pathway	R-HSA-72163	1.11E-07	4.99E-07
Peptide chain elongation	R-HSA-156902	1.50E-07	6.22E-07
mRNA Splicing	R-HSA-72172	1.67E-07	6.36E-07
Eukaryotic Translation Termination	R-HSA-72764	1.89E-07	6.36E-07
Selenocysteine synthesis	R-HSA-2408557	1.89E-07	6.36E-07
Eukaryotic Translation Elongation	R-HSA-156842	2.19E-07	6.57E-07
Nonsense Mediated Decay (NMD) independent of the Exon Junction Complex (EJC)	R-HSA-975956	2.19E-07	6.57E-07
Metabolism of amino acids and derivatives	R-HSA-71291	7.37E-07	2.10E-06
Processing of Capped Intron-Containing Pre-mRNA	R-HSA-72203	1.74E-06	4.71E-06
Metabolism of proteins	R-HSA-392499	2.46E-06	6.33E-06
rRNA modification in the nucleus and cytosol	R-HSA-6790901	2.56E-05	6.01E-05
rRNA processing in the nucleus and cytosol	R-HSA-8868773	2.56E-05	6.01E-05
Epigenetic regulation of gene expression	R-HSA-212165	3.35E-05	7.55E-05
rRNA processing	R-HSA-72312	4.14E-05	8.94E-05
Positive epigenetic regulation of rRNA expression	R-HSA-5250913	4.77E-05	9.27E-05
Translation initiation complex formation	R-HSA-72649	4.81E-05	9.27E-05
Ribosomal scanning and start codon recognition	R-HSA-72702	4.81E-05	9.27E-05
Activation of the mRNA upon binding of the cap-binding complex and eIFs, and subsequent binding to 43S	R-HSA-72662	5.17E-05	9.62E-05
DNA Double-Strand Break Repair	R-HSA-5693532	6.65E-05	1.20E-04
DNA Repair	R-HSA-73894	1.22E-04	2.12E-04
mRNA Splicing - Minor Pathway	R-HSA-72165	2.71E-04	4.57E-04
Formation of the ternary complex, and subsequently, the 43S complex	R-HSA-72695	4.41E-04	7.22E-04
Nonhomologous End-Joining (NHEJ)	R-HSA-5693571	8.00E-04	1.24E-03
Infectious disease	R-HSA-5663205	8.02E-04	1.24E-03
Gene expression (Transcription)	R-HSA-74160	1.09E-03	1.63E-03
HIV Life Cycle	R-HSA-162587	2.40E-03	3.44E-03
B-WICH complex positively regulates rRNA expression	R-HSA-5250924	2.42E-03	3.44E-03
Regulation of TP53 Activity	R-HSA-5633007	2.91E-03	4.03E-03
HIV Infection	R-HSA-162906	6.55E-03	8.84E-03
Transcriptional Regulation by TP53	R-HSA-3700989	7.03E-03	9.27E-03

REFERENCES

1. Kanitakis J: Anatomy, histology and immunohistochemistry of normal human skin. *European journal of dermatology : EJD* 2002, 12(4):390-399; quiz 400-391.
2. Lin JY, Fisher DE: Melanocyte biology and skin pigmentation. *Nature* 2007, 445(7130):843-850.
3. Jimbow K, Roth SI, Fitzpatrick TB, Szabo G: Mitotic activity in non-neoplastic melanocytes in vivo as determined by histochemical, autoradiographic, and electron microscope studies. *The Journal of cell biology* 1975, 66(3):663-670.
4. Shain AH, Bastian BC: From melanocytes to melanomas. *Nat Rev Cancer* 2016, 16(6):345-358.
5. Schadendorf D, van Akkooi ACJ, Berking C, Griewank KG, Gutzmer R, Hauschild A, Stang A, Roesch A, Ugurel S: Melanoma. *The Lancet* 2018, 392(10151):971-984.
6. Committee CCSA: Canadian Cancer Statistics 2018. *Toronto, ON: Canadian Cancer Society* 2018.
7. Boniol M, Autier P, Boyle P, Gandini S: Cutaneous melanoma attributable to sunbed use: systematic review and meta-analysis. *BMJ (Clinical research ed)* 2012, 345:e4757.
8. FitzGerald MG, Harkin DP, Silva-Arrieta S, MacDonald DJ, Lucchina LC, Unsal H, O'Neill E, Koh J, Finkelstein DM, Isselbacher KJ *et al*: Prevalence of germ-line mutations in p16, p19ARF, and CDK4 in familial melanoma: analysis of a clinic-based population. *Proc Natl Acad Sci U S A* 1996, 93(16):8541-8545.
9. Gandini S, Sera F, Cattaruzza MS, Pasquini P, Zanetti R, Masini C, Boyle P, Melchi CF: Meta-analysis of risk factors for cutaneous melanoma: III. Family history, actinic damage and phenotypic factors. *Eur J Cancer* 2005, 41(14):2040-2059.
10. Brash DE: UV signature mutations. *Photochem Photobiol* 2015, 91(1):15-26.
11. Cleaver JE, Lam ET, Revet I: Disorders of nucleotide excision repair: the genetic and molecular basis of heterogeneity. *Nat Rev Genet* 2009, 10(11):756-768.
12. Hodis E, Watson IR, Kryukov GV, Arold ST, Imielinski M, Theurillat JP, Nickerson E, Auclair D, Li L, Place C *et al*: A landscape of driver mutations in melanoma. *Cell* 2012, 150(2):251-263.
13. Krauthammer M, Kong Y, Bacchiocchi A, Evans P, Pornputtapong N, Wu C, McCusker JP, Ma S, Cheng E, Straub R *et al*: Exome sequencing identifies recurrent mutations in NF1 and RASopathy genes in sun-exposed melanomas. *Nat Genet* 2015, 47(9):996-1002.
14. CancerGenomeAtlasNetwork: Genomic Classification of Cutaneous Melanoma. *Cell* 2015, 161(7):1681-1696.
15. Eggermont AMM, Spatz A, Robert C: Cutaneous melanoma. *The Lancet* 2014, 383(9919):816-827.
16. Horn S, Figl A, Rachakonda PS, Fischer C, Sucker A, Gast A, Kadel S, Moll I, Nagore E, Hemminki K *et al*: TERT promoter mutations in familial and sporadic melanoma. *Science (New York, NY)* 2013, 339(6122):959-961.
17. Huang FW, Hodis E, Xu MJ, Kryukov GV, Chin L, Garraway LA: Highly recurrent TERT promoter mutations in human melanoma. *Science (New York, NY)* 2013, 339(6122):957-959.
18. Hussussian CJ, Struewing JP, Goldstein AM, Higgins PA, Ally DS, Sheahan MD, Clark WH, Jr., Tucker MA, Dracopoli NC: Germline p16 mutations in familial melanoma. *Nat Genet* 1994, 8(1):15-21.
19. Robles-Espinoza CD, Harland M, Ramsay AJ, Aoude LG, Quesada V, Ding Z, Pooley KA, Pritchard AL, Tiffen JC, Petljak M *et al*: POT1 loss-of-function variants predispose to familial melanoma. *Nat Genet* 2014, 46(5):478-481.
20. Shi J, Yang XR, Ballew B, Rotunno M, Calista D, Fargnoli MC, Ghiorzo P, Bressac-de Paillerets B, Nagore E, Avril MF *et al*: Rare missense variants in POT1 predispose to familial cutaneous malignant melanoma. *Nat Genet* 2014, 46(5):482-486.
21. Zuo L, Weger J, Yang Q, Goldstein AM, Tucker MA, Walker GJ, Hayward N, Dracopoli NC: Germline mutations in the p16INK4a binding domain of CDK4 in familial melanoma. *Nat Genet* 1996, 12(1):97-99.
22. Tsao H, Olazagasti JM, Cordoro KM, Brewer JD, Taylor SC, Bordeaux JS, Chren MM, Sober AJ, Tegeler C, Bhushan R *et al*: Early detection of melanoma: reviewing the ABCDEs. *Journal of the American Academy of Dermatology* 2015, 72(4):717-723.
23. Chin L: The genetics of malignant melanoma: lessons from mouse and man. *Nat Rev Cancer* 2003, 3(8):559-570.
24. Gershenwald JE, Scolyer RA: Melanoma Staging: American Joint Committee on Cancer (AJCC) 8th Edition and Beyond. *Ann Surg Oncol* 2018, 25(8):2105-2110.
25. Pitcovski J, Shahar E, Aizenshtein E, Gorodetsky R: Melanoma antigens and related immunological markers. *Crit Rev Oncol Hematol* 2017, 115:36-49.

26. Guarino M, Rubino B, Ballabio G: The role of epithelial-mesenchymal transition in cancer pathology. *Pathology* 2007, 39(3):305-318.
27. Kalluri R, Weinberg RA: The basics of epithelial-mesenchymal transition. *The Journal of clinical investigation* 2009, 119(6):1420-1428.
28. Lamouille S, Xu J, Derynck R: Molecular mechanisms of epithelial-mesenchymal transition. *Nat Rev Mol Cell Biol* 2014, 15(3):178-196.
29. Nieto MA, Huang RY, Jackson RA, Thiery JP: Emt: 2016. *Cell* 2016, 166(1):21-45.
30. Roche J: The Epithelial-to-Mesenchymal Transition in Cancer. *Cancers (Basel)* 2018, 10(2).
31. Alonso SR, Tracey L, Ortiz P, Perez-Gomez B, Palacios J, Pollan M, Linares J, Serrano S, Saez-Castillo AI, Sanchez L *et al*: A high-throughput study in melanoma identifies epithelial-mesenchymal transition as a major determinant of metastasis. *Cancer Res* 2007, 67(7):3450-3460.
32. Hoek KS, Eichhoff OM, Schlegel NC, Dobbeling U, Kobert N, Schaerer L, Hemmi S, Dummer R: In vivo switching of human melanoma cells between proliferative and invasive states. *Cancer Res* 2008, 68(3):650-656.
33. Hoek KS, Goding CR: Cancer stem cells versus phenotype-switching in melanoma. *Pigment Cell Melanoma Res* 2010, 23(6):746-759.
34. Caramel J, Papadogeorgakis E, Hill L, Browne GJ, Richard G, Wierinckx A, Saldanha G, Osborne J, Hutchinson P, Tse G *et al*: A switch in the expression of embryonic EMT-inducers drives the development of malignant melanoma. *Cancer Cell* 2013, 24(4):466-480.
35. Denecker G, Vandamme N, Akay O, Koludrovic D, Taminau J, Lemeire K, Gheldof A, De Craene B, Van Gele M, Brochez L *et al*: Identification of a ZEB2-MITF-ZEB1 transcriptional network that controls melanogenesis and melanoma progression. *Cell death and differentiation* 2014, 21(8):1250-1261.
36. Fenouille N, Tichet M, Dufies M, Pottier A, Mogha A, Soo JK, Rocchi S, Mallavialle A, Galibert MD, Khammari A *et al*: The epithelial-mesenchymal transition (EMT) regulatory factor SLUG (SNAI2) is a downstream target of SPARC and AKT in promoting melanoma cell invasion. *PLoS One* 2012, 7(7):e40378.
37. Wels C, Joshi S, Koefinger P, Bergler H, Schaider H: Transcriptional activation of ZEB1 by Slug leads to cooperative regulation of the epithelial-mesenchymal transition-like phenotype in melanoma. *The Journal of investigative dermatology* 2011, 131(9):1877-1885.
38. Hartman ML, Czyz M: MITF in melanoma: mechanisms behind its expression and activity. *Cell Mol Life Sci* 2015, 72(7):1249-1260.
39. Seberg HE, Van Otterloo E, Cornell RA: Beyond MITF: Multiple transcription factors directly regulate the cellular phenotype in melanocytes and melanoma. *Pigment Cell Melanoma Res* 2017, 30(5):454-466.
40. Opdecamp K, Nakayama A, Nguyen MT, Hodgkinson CA, Pavan WJ, Arnheiter H: Melanocyte development in vivo and in neural crest cell cultures: crucial dependence on the Mitf basic-helix-loop-helix-zipper transcription factor. *Development (Cambridge, England)* 1997, 124(12):2377-2386.
41. Levy C, Khaled M, Fisher DE: MITF: master regulator of melanocyte development and melanoma oncogene. *Trends in molecular medicine* 2006, 12(9):406-414.
42. Carreira S, Goodall J, Denat L, Rodriguez M, Nuciforo P, Hoek KS, Testori A, Larue L, Goding CR: Mitf regulation of Dia1 controls melanoma proliferation and invasiveness. *Genes Dev* 2006, 20(24):3426-3439.
43. Arozarena I, Wellbrock C: Targeting invasive properties of melanoma cells. *FEBS J* 2017, 284(14):2148-2162.
44. Lin K, Baritaki S, Militello L, Malaponte G, Bevelacqua Y, Bonavida B: The Role of B-RAF Mutations in Melanoma and the Induction of EMT via Dysregulation of the NF-kappaB/Snail/RKIP/PTEN Circuit. *Genes Cancer* 2010, 1(5):409-420.
45. Whipple CA, Brinckerhoff CE: BRAF(V600E) melanoma cells secrete factors that activate stromal fibroblasts and enhance tumourigenicity. *Br J Cancer* 2014, 111(8):1625-1633.
46. Huntington JT, Shields JM, Der CJ, Wyatt CA, Benbow U, Slingluff CL, Jr., Brinckerhoff CE: Overexpression of collagenase 1 (MMP-1) is mediated by the ERK pathway in invasive melanoma cells: role of BRAF mutation and fibroblast growth factor signaling. *J Biol Chem* 2004, 279(32):33168-33176.
47. De Craene B, van Roy F, Berx G: Unraveling signalling cascades for the Snail family of transcription factors. *Cellular signalling* 2005, 17(5):535-547.
48. Vandamme N, Berx G: Melanoma cells revive an embryonic transcriptional network to dictate phenotypic heterogeneity. *Front Oncol* 2014, 4:352.

49. Gregory PA, Bracken CP, Smith E, Bert AG, Wright JA, Roslan S, Morris M, Wyatt L, Farshid G, Lim YY *et al*: An autocrine TGF-beta/ZEB/miR-200 signaling network regulates establishment and maintenance of epithelial-mesenchymal transition. *Molecular biology of the cell* 2011, 22(10):1686-1698.
50. Perrot CY, Gilbert C, Marsaud V, Postigo A, Javelaud D, Mauviel A: GLI2 cooperates with ZEB1 for transcriptional repression of CDH1 expression in human melanoma cells. *Pigment Cell Melanoma Res* 2013, 26(6):861-873.
51. Hoek KS, Schlegel NC, Brafford P, Sucker A, Ugurel S, Kumar R, Weber BL, Nathanson KL, Phillips DJ, Herlyn M *et al*: Metastatic potential of melanomas defined by specific gene expression profiles with no BRAF signature. *Pigment Cell Res* 2006, 19(4):290-302.
52. Kaur A, Webster MR, Weeraratna AT: In the Wnt-er of life: Wnt signalling in melanoma and ageing. *Br J Cancer* 2016, 115(11):1273-1279.
53. Demunter A, Libbrecht L, Degreef H, De Wolf-Peeters C, van den Oord JJ: Loss of membranous expression of beta-catenin is associated with tumor progression in cutaneous melanoma and rarely caused by exon 3 mutations. *Mod Pathol* 2002, 15(4):454-461.
54. Zhou BP, Deng J, Xia W, Xu J, Li YM, Gunduz M, Hung MC: Dual regulation of Snail by GSK-3beta-mediated phosphorylation in control of epithelial-mesenchymal transition. *Nature cell biology* 2004, 6(10):931-940.
55. Arozarena I, Bischof H, Gilby D, Belloni B, Dummer R, Wellbrock C: In melanoma, beta-catenin is a suppressor of invasion. *Oncogene* 2011, 30(45):4531-4543.
56. Pearlman RL, Montes de Oca MK, Pal HC, Afaq F: Potential therapeutic targets of epithelial-mesenchymal transition in melanoma. *Cancer Lett* 2017, 391:125-140.
57. Webster MR, Kugel CH, 3rd, Weeraratna AT: The Wnts of change: How Wnts regulate phenotype switching in melanoma. *Biochim Biophys Acta* 2015, 1856(2):244-251.
58. Dissanayake SK, Olkhanud PB, O'Connell MP, Carter A, French AD, Camilli TC, Emeche CD, Hewitt KJ, Rosenthal DT, Leotlela PD *et al*: Wnt5A regulates expression of tumor-associated antigens in melanoma via changes in signal transducers and activators of transcription 3 phosphorylation. *Cancer Res* 2008, 68(24):10205-10214.
59. Verfaillie A, Imrichova H, Atak ZK, Dewaele M, Rambow F, Hulselmans G, Christiaens V, Svetlichnyy D, Luciani F, Van den Mooter L *et al*: Decoding the regulatory landscape of melanoma reveals TEADS as regulators of the invasive cell state. *Nat Commun* 2015, 6:6683.
60. Widmer DS, Cheng PF, Eichhoff OM, Belloni BC, Zipser MC, Schlegel NC, Javelaud D, Mauviel A, Dummer R, Hoek KS: Systematic classification of melanoma cells by phenotype-specific gene expression mapping. *Pigment Cell Melanoma Res* 2012, 25(3):343-353.
61. Aibar S, Gonzalez-Blas CB, Moerman T, Huynh-Thu VA, Imrichova H, Hulselmans G, Rambow F, Marine JC, Geurts P, Aerts J *et al*: SCENIC: single-cell regulatory network inference and clustering. *Nat Methods* 2017, 14(11):1083-1086.
62. Afshar N, English DR, Thursfield V, Mitchell PL, Te Marvelde L, Farrugia H, Giles GG, Milne RL: Differences in cancer survival by sex: a population-based study using cancer registry data. *Cancer causes & control : CCC* 2018, 29(11):1059-1069.
63. de Vries E, Nijsten TE, Visser O, Bastiaannet E, van Hattem S, Janssen-Heijnen ML, Coebergh JW: Superior survival of females among 10,538 Dutch melanoma patients is independent of Breslow thickness, histologic type and tumor site. *Annals of oncology : official journal of the European Society for Medical Oncology* 2008, 19(3):583-589.
64. Gamba CS, Clarke CA, Keegan TH, Tao L, Swetter SM: Melanoma survival disadvantage in young, non-Hispanic white males compared with females. *JAMA dermatology* 2013, 149(8):912-920.
65. Kemeny MM, Busch E, Stewart AK, Menck HR: Superior survival of young women with malignant melanoma. *American journal of surgery* 1998, 175(6):437-444; discussion 444-435.
66. Joosse A, Collette S, Suciu S, Nijsten T, Lejeune F, Kleeberg UR, Coebergh JW, Eggermont AM, de Vries E: Superior outcome of women with stage I/II cutaneous melanoma: pooled analysis of four European Organisation for Research and Treatment of Cancer phase III trials. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2012, 30(18):2240-2247.
67. Joosse A, Collette S, Suciu S, Nijsten T, Patel PM, Keilholz U, Eggermont AM, Coebergh JW, de Vries E: Sex is an independent prognostic indicator for survival and relapse/progression-free survival in metastasized stage III to IV melanoma: a pooled analysis of five European organisation for research and treatment of cancer randomized controlled trials. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2013, 31(18):2337-2346.

68. Fisher DE, Geller AC: Disproportionate burden of melanoma mortality in young U.S. men: the possible role of biology and behavior. *JAMA dermatology* 2013, 149(8):903-904.
69. Sondak VK, Swetter SM, Berwick MA: Gender disparities in patients with melanoma: breaking the glass ceiling. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2012, 30(18):2177-2178.
70. Chapman PB, Hauschild A, Robert C, Haanen JB, Ascierto P, Larkin J, Dummer R, Garbe C, Testori A, Maio M *et al*: Improved survival with vemurafenib in melanoma with BRAF V600E mutation. *The New England journal of medicine* 2011, 364(26):2507-2516.
71. Kim A, Cohen MS: The discovery of vemurafenib for the treatment of BRAF-mutated metastatic melanoma. *Expert opinion on drug discovery* 2016, 11(9):907-916.
72. Puzanov I, Amaravadi RK, McArthur GA, Flaherty KT, Chapman PB, Sosman JA, Ribas A, Shackleton M, Hwu P, Chmielowski B *et al*: Long-term outcome in BRAF(V600E) melanoma patients treated with vemurafenib: Patterns of disease progression and clinical management of limited progression. *Eur J Cancer* 2015, 51(11):1435-1443.
73. Hodi FS, O'Day SJ, McDermott DF, Weber RW, Sosman JA, Haanen JB, Gonzalez R, Robert C, Schadendorf D, Hassel JC *et al*: Improved survival with ipilimumab in patients with metastatic melanoma. *The New England journal of medicine* 2010, 363(8):711-723.
74. Robert C, Ribas A, Wolchok JD, Hodi FS, Hamid O, Kefford R, Weber JS, Joshua AM, Hwu WJ, Gangadhar TC *et al*: Anti-programmed-death-receptor-1 treatment with pembrolizumab in ipilimumab-refractory advanced melanoma: a randomised dose-comparison cohort of a phase 1 trial. *Lancet (London, England)* 2014, 384(9948):1109-1117.
75. Robert C, Schachter J, Long GV, Arance A, Grob JJ, Mortier L, Daud A, Carlino MS, McNeil C, Lotem M *et al*: Pembrolizumab versus Ipilimumab in Advanced Melanoma. *The New England journal of medicine* 2015, 372(26):2521-2532.
76. Herin M, Lemoine C, Weynants P, Vessiere F, Van Pel A, Knuth A, Devos R, Boon T: Production of stable cytolytic T-cell clones directed against autologous human melanoma. *Int J Cancer* 1987, 39(3):390-396.
77. Muul LM, Spiess PJ, Director EP, Rosenberg SA: Identification of specific cytolytic immune responses against autologous tumor in humans bearing malignant melanoma. *J Immunol* 1987, 138(3):989-995.
78. Coulie PG, Van den Eynde BJ, van der Bruggen P, Boon T: Tumour antigens recognized by T lymphocytes: at the core of cancer immunotherapy. *Nat Rev Cancer* 2014, 14(2):135-146.
79. Ilyas S, Yang JC: Landscape of Tumor Antigens in T Cell Immunotherapy. *J Immunol* 2015, 195(11):5117-5122.
80. Whiteside TL: Immune responses to malignancies. *J Allergy Clin Immunol* 2010, 125(2 Suppl 2):S272-283.
81. Knuth A, Wolfel T, Klehmann E, Boon T, Meyer zum Buschenfelde KH: Cytolytic T-cell clones against an autologous human melanoma: specificity study and definition of three antigens by immunoselection. *Proc Natl Acad Sci U S A* 1989, 86(8):2804-2808.
82. Traversari C, van der Bruggen P, Luescher IF, Lurquin C, Chomez P, Van Pel A, De Plaen E, Amar-Costesec A, Boon T: A nonapeptide encoded by human gene MAGE-1 is recognized on HLA-A1 by cytolytic T lymphocytes directed against tumor antigen MZ2-E. *The Journal of experimental medicine* 1992, 176(5):1453-1457.
83. van der Bruggen P, Traversari C, Chomez P, Lurquin C, De Plaen E, Van den Eynde B, Knuth A, Boon T: A gene encoding an antigen recognized by cytolytic T lymphocytes on a human melanoma. *Science (New York, NY)* 1991, 254(5038):1643-1647.
84. Barrow C, Browning J, MacGregor D, Davis ID, Sturrock S, Jungbluth AA, Cebon J: Tumor antigen expression in melanoma varies according to antigen and stage. *Clin Cancer Res* 2006, 12(3 Pt 1):764-771.
85. Brasseur F, Rimoldi D, Lienard D, Lethe B, Carrel S, Arienti F, Suter L, Vanwijck R, Bourlond A, Humblet Y *et al*: Expression of MAGE genes in primary and metastatic cutaneous melanoma. *Int J Cancer* 1995, 63(3):375-380.
86. Hodi FS: Well-defined melanoma antigens as progression markers for melanoma: insights into differential expression and host response based on stage. *Clin Cancer Res* 2006, 12(3 Pt 1):673-678.
87. Houghton AN, Gold JS, Blachere NE: Immunity against cancer: lessons learned from melanoma. *Current opinion in immunology* 2001, 13(2):134-140.
88. Bakker AB, Schreurs MW, de Boer AJ, Kawakami Y, Rosenberg SA, Adema GJ, Figdor CG: Melanocyte lineage-specific antigen gp100 is recognized by melanoma-derived tumor-infiltrating lymphocytes. *The Journal of experimental medicine* 1994, 179(3):1005-1009.

89. Kawakami Y, Eliyahu S, Delgado CH, Robbins PF, Rivoltini L, Topalian SL, Miki T, Rosenberg SA: Cloning of the gene coding for a shared human melanoma antigen recognized by autologous T cells infiltrating into tumor. *Proc Natl Acad Sci U S A* 1994, 91(9):3515-3519.
90. Brichard V, Van Pel A, Wolfel T, Wolfel C, De Plaen E, Lethe B, Coulie P, Boon T: The tyrosinase gene codes for an antigen recognized by autologous cytolytic T lymphocytes on HLA-A2 melanomas. *The Journal of experimental medicine* 1993, 178(2):489-495.
91. Wang RF, Robbins PF, Kawakami Y, Kang XQ, Rosenberg SA: Identification of a gene encoding a melanoma tumor antigen recognized by HLA-A31-restricted tumor-infiltrating lymphocytes. *The Journal of experimental medicine* 1995, 181(2):799-804.
92. Wang RF, Appella E, Kawakami Y, Kang X, Rosenberg SA: Identification of TRP-2 as a human tumor antigen recognized by cytotoxic T lymphocytes. *The Journal of experimental medicine* 1996, 184(6):2207-2216.
93. Vonderheide RH, Hahn WC, Schultze JL, Nadler LM: The telomerase catalytic subunit is a widely expressed tumor-associated antigen recognized by cytotoxic T lymphocytes. *Immunity* 1999, 10(6):673-679.
94. Andersen MH, Pedersen LO, Becker JC, Straten PT: Identification of a cytotoxic T lymphocyte response to the apoptosis inhibitor protein survivin in cancer patients. *Cancer Res* 2001, 61(3):869-872.
95. Robbins PF, El-Gamil M, Li YF, Kawakami Y, Loftus D, Appella E, Rosenberg SA: A mutated beta-catenin gene encodes a melanoma-specific antigen recognized by tumor infiltrating lymphocytes. *The Journal of experimental medicine* 1996, 183(3):1185-1192.
96. Wolfel T, Hauer M, Schneider J, Serrano M, Wolfel C, Klehmann-Hieb E, De Plaen E, Hankeln T, Meyer zum Buschenfelde KH, Beach D: A p16INK4a-insensitive CDK4 mutant targeted by cytolytic T lymphocytes in a human melanoma. *Science (New York, NY)* 1995, 269(5228):1281-1284.
97. Veatch JR, Lee SM, Fitzgibbon M, Chow IT, Jesernig B, Schmitt T, Kong YY, Kargl J, Houghton AM, Thompson JA *et al*: Tumor-infiltrating BRAFV600E-specific CD4+ T cells correlated with complete clinical response in melanoma. *The Journal of clinical investigation* 2018, 128(4):1563-1568.
98. Lee AK, Potts PR: A Comprehensive Guide to the MAGE Family of Ubiquitin Ligases. *J Mol Biol* 2017, 429(8):1114-1142.
99. Caballero OL, Chen YT: Cancer/testis (CT) antigens: potential targets for immunotherapy. *Cancer Sci* 2009, 100(11):2014-2021.
100. Fratta E, Coral S, Covre A, Parisi G, Colizzi F, Danielli R, Nicolay HJ, Sigalotti L, Maio M: The biology of cancer testis antigens: putative function, regulation and therapeutic potential. *Mol Oncol* 2011, 5(2):164-182.
101. Simpson AJ, Caballero OL, Jungbluth A, Chen YT, Old LJ: Cancer/testis antigens, gametogenesis and cancer. *Nat Rev Cancer* 2005, 5(8):615-625.
102. Grigoriadis A, Caballero OL, Hoek KS, da Silva L, Chen YT, Shin SJ, Jungbluth AA, Miller LD, Clouston D, Cebon J *et al*: CT-X antigen expression in human breast cancer. *Proc Natl Acad Sci U S A* 2009, 106(32):13493-13498.
103. Ross MT, Grafham DV, Coffey AJ, Scherer S, McLay K, Muzny D, Platzer M, Howell GR, Burrows C, Bird CP *et al*: The DNA sequence of the human X chromosome. *Nature* 2005, 434(7031):325-337.
104. Salmaninejad A, Zamani MR, Pourvahedi M, Golchehre Z, Hosseini Bereshneh A, Rezaei N: Cancer/Testis Antigens: Expression, Regulation, Tumor Invasion, and Use in Immunotherapy of Cancers. *Immunol Invest* 2016, 45(7):619-640.
105. Cheng YH, Wong EW, Cheng CY: Cancer/testis (CT) antigens, carcinogenesis and spermatogenesis. *Spermatogenesis* 2011, 1(3):209-220.
106. Sahin U, Tureci O, Chen YT, Seitz G, Villena-Heinsen C, Old LJ, Pfreundschuh M: Expression of multiple cancer/testis (CT) antigens in breast cancer and melanoma: basis for polyvalent CT vaccine strategies. *Int J Cancer* 1998, 78(3):387-389.
107. Jungbluth AA, Stockert E, Chen YT, Kolb D, Iversen K, Coplan K, Williamson B, Altorki N, Busam KJ, Old LJ: Monoclonal antibody MA454 reveals a heterogeneous expression pattern of MAGE-1 antigen in formalin-fixed paraffin embedded lung tumours. *Br J Cancer* 2000, 83(4):493-497.
108. De Backer O, Arden KC, Boretti M, Vantomme V, De Smet C, Czekay S, Viars CS, De Plaen E, Brasseur F, Chomez P *et al*: Characterization of the GAGE genes that are expressed in various human cancers and in normal testis. *Cancer Res* 1999, 59(13):3157-3165.
109. Lethe B, van der Bruggen P, Brasseur F, Boon T: MAGE-1 expression threshold for the lysis of melanoma cell lines by a specific cytotoxic T lymphocyte. *Melanoma research* 1997, 7 Suppl 2:S83-88.

110. De Smet C, Courtois SJ, Faraoni I, Lurquin C, Szikora JP, De Backer O, Boon T: Involvement of two Ets binding sites in the transcriptional activation of the MAGE1 gene. *Immunogenetics* 1995, 42(4):282-290.
111. De Smet C, Lurquin C, Lethe B, Martelange V, Boon T: DNA methylation is the primary silencing mechanism for a set of germ line- and tumor-specific genes with a CpG-rich promoter. *Mol Cell Biol* 1999, 19(11):7327-7335.
112. Lian Y, Meng L, Ding P, Sang M: Epigenetic regulation of MAGE family in human cancer progression-DNA methylation, histone modification, and non-coding RNAs. *Clin Epigenetics* 2018, 10(1):115.
113. Weber J, Salgaller M, Samid D, Johnson B, Herlyn M, Lassam N, Treisman J, Rosenberg SA: Expression of the MAGE-1 tumor antigen is up-regulated by the demethylating agent 5-aza-2'-deoxycytidine. *Cancer Res* 1994, 54(7):1766-1771.
114. Sigalotti L, Coral S, Altomonte M, Natali L, Gaudino G, Cacciotti P, Libener R, Colizzi F, Vianale G, Martini F *et al*: Cancer testis antigens expression in mesothelioma: role of DNA methylation and bioimmunotherapeutic implications. *Br J Cancer* 2002, 86(6):979-982.
115. Sigalotti L, Coral S, Nardi G, Spessotto A, Cortini E, Cattarossi I, Colizzi F, Altomonte M, Maio M: Promoter methylation controls the expression of MAGE2, 3 and 4 genes in human cutaneous melanoma. *Journal of immunotherapy (Hagerstown, Md : 1997)* 2002, 25(1):16-26.
116. Weiser TS, Ohnmacht GA, Guo ZS, Fischette MR, Chen GA, Hong JA, Nguyen DM, Schrump DS: Induction of MAGE-3 expression in lung and esophageal cancer cells. *The Annals of thoracic surgery* 2001, 71(1):295-301; discussion 301-292.
117. Heninger E, Krueger TE, Thiede SM, Sperger JM, Byers BL, Kircher MR, Kosoff D, Yang B, Jarrard DF, McNeel DG *et al*: Inducible expression of cancer-testis antigens in human prostate cancer. *Oncotarget* 2016, 7(51):84359-84374.
118. Dubovsky JA, McNeel DG: Inducible expression of a prostate cancer-testis antigen, SSX-2, following treatment with a DNA methylation inhibitor. *Prostate* 2007, 67(16):1781-1790.
119. Bao L, Dunham K, Lucas K: MAGE-A1, MAGE-A3, and NY-ESO-1 can be upregulated on neuroblastoma cells to facilitate cytotoxic T lymphocyte-mediated tumor cell killing. *Cancer immunology, immunotherapy : CII* 2011, 60(9):1299-1307.
120. Grunwald C, Koslowski M, Arsiray T, Dhaene K, Praet M, Victor A, Morresi-Hauf A, Lindner M, Passlick B, Lehr HA *et al*: Expression of multiple epigenetically regulated cancer/germline genes in nonsmall cell lung cancer. *Int J Cancer* 2006, 118(10):2522-2528.
121. Natsume A, Wakabayashi T, Tsujimura K, Shimato S, Ito M, Kuzushima K, Kondo Y, Sekido Y, Kawatsura H, Narita Y *et al*: The DNA demethylating agent 5-aza-2'-deoxycytidine activates NY-ESO-1 antigenicity in orthotopic human glioma. *Int J Cancer* 2008, 122(11):2542-2553.
122. James SR, Link PA, Karpf AR: Epigenetic regulation of X-linked cancer/germline antigen genes by DNMT1 and DNMT3b. *Oncogene* 2006, 25(52):6975-6985.
123. Akers SN, Odunsi K, Karpf AR: Regulation of cancer germline antigen gene expression: implications for cancer immunotherapy. *Future Oncol* 2010, 6(5):717-732.
124. Wischnewski F, Pantel K, Schwarzenbach H: Promoter demethylation and histone acetylation mediate gene expression of MAGE-A1, -A2, -A3, and -A12 in human cancer cells. *Mol Cancer Res* 2006, 4(5):339-349.
125. Tachibana M, Matsumura Y, Fukuda M, Kimura H, Shinkai Y: G9a/GLP complexes independently mediate H3K9 and DNA methylation to silence transcription. *Embo j* 2008, 27(20):2681-2690.
126. Tachibana M, Ueda J, Fukuda M, Takeda N, Ohta T, Iwanari H, Sakihama T, Kodama T, Hamakubo T, Shinkai Y: Histone methyltransferases G9a and GLP form heteromeric complexes and are both crucial for methylation of euchromatin at H3-K9. *Genes Dev* 2005, 19(7):815-826.
127. Link PA, Gangisetty O, James SR, Woloszynska-Read A, Tachibana M, Shinkai Y, Karpf AR: Distinct roles for histone methyltransferases G9a and GLP in cancer germ-line antigen gene regulation in human cancer cells and murine embryonic stem cells. *Mol Cancer Res* 2009, 7(6):851-862.
128. Sun F, Chan E, Wu Z, Yang X, Marquez VE, Yu Q: Combinatorial pharmacologic approaches target EZH2-mediated gene repression in breast cancer cells. *Molecular cancer therapeutics* 2009, 8(12):3191-3202.
129. Sang M, Wang L, Ding C, Zhou X, Wang B, Wang L, Lian Y, Shan B: Melanoma-associated antigen genes - an update. *Cancer Lett* 2011, 302(2):85-90.
130. Laduron S, Deplus R, Zhou S, Kholmanskikh O, Godelaine D, De Smet C, Hayward SD, Fuks F, Boon T, De Plaen E: MAGE-A1 interacts with adaptor SKIP and the deacetylase HDAC1 to repress transcription. *Nucleic Acids Res* 2004, 32(14):4340-4350.

131. Wang D, Wang J, Ding N, Li Y, Yang Y, Fang X, Zhao H: MAGE-A1 promotes melanoma proliferation and migration through C-JUN activation. *Biochem Biophys Res Commun* 2016, 473(4):959-965.
132. Nagao T, Higashitsuji H, Nonoguchi K, Sakurai T, Dawson S, Mayer RJ, Itoh K, Fujita J: MAGE-A4 interacts with the liver oncoprotein gankyrin and suppresses its tumorigenic activity. *J Biol Chem* 2003, 278(12):10668-10674.
133. Nardiello T, Jungbluth AA, Mei A, Diliberto M, Huang X, Dabrowski A, Andrade VC, Wasserstrum R, Ely S, Niesvizky R *et al*: MAGE-A inhibits apoptosis in proliferating myeloma cells through repression of Bax and maintenance of survivin. *Clin Cancer Res* 2011, 17(13):4309-4319.
134. Feng Y, Gao J, Yang M: When MAGE meets RING: insights into biological functions of MAGE proteins. *Protein & cell* 2011, 2(1):7-12.
135. Cilensek ZM, Yehiely F, Kular RK, Deiss LP: A member of the GAGE family of tumor antigens is an anti-apoptotic gene that confers resistance to Fas/CD95/APO-1, Interferon-gamma, taxol and gamma-irradiation. *Cancer biology & therapy* 2002, 1(4):380-387.
136. Cronwright G, Le Blanc K, Gotherstrom C, Darcy P, Ehnman M, Brodin B: Cancer/testis antigen expression in human mesenchymal stem cells: down-regulation of SSX impairs cell migration and matrix metalloproteinase 2 expression. *Cancer Res* 2005, 65(6):2207-2215.
137. Velazquez EF, Jungbluth AA, Yancovitz M, Gnjatic S, Adams S, O'Neill D, Zavilevich K, Albukh T, Christos P, Mazumdar M *et al*: Expression of the cancer/testis antigen NY-ESO-1 in primary and metastatic malignant melanoma (MM)--correlation with prognostic factors. *Cancer Immun* 2007, 7:11.
138. Kholmanskikh O, Loriot A, Brasseur F, De Plaen E, De Smet C: Expression of BORIS in melanoma: lack of association with MAGE-A1 activation. *Int J Cancer* 2008, 122(4):777-784.
139. Loukinov DI, Pugacheva E, Vatolin S, Pack SD, Moon H, Chernukhin I, Mannan P, Larsson E, Kanduri C, Vostrov AA *et al*: BORIS, a novel male germ-line-specific protein associated with epigenetic reprogramming events, shares the same 11-zinc-finger domain with CTCF, the insulator protein involved in reading imprinting marks in the soma. *Proc Natl Acad Sci U S A* 2002, 99(10):6806-6811.
140. Phillips JE, Corces VG: CTCF: master weaver of the genome. *Cell* 2009, 137(7):1194-1211.
141. Klenova EM, Morse HC, 3rd, Ohlsson R, Lobanenkov VV: The novel BORIS + CTCF gene family is uniquely involved in the epigenetics of normal biology and cancer. *Semin Cancer Biol* 2002, 12(5):399-414.
142. Marshall AD, Bailey CG, Rasko JE: CTCF and BORIS in genome regulation and cancer. *Curr Opin Genet Dev* 2014, 24:8-15.
143. Hore TA, Deakin JE, Marshall Graves JA: The evolution of epigenetic regulators CTCF and BORIS/CTCFL in amniotes. *PLoS Genet* 2008, 4(8):e1000169.
144. Pugacheva EM, Rivero-Hinojosa S, Espinoza CA, Mendez-Catala CF, Kang S, Suzuki T, Kosaka-Suzuki N, Robinson S, Nagarajan V, Ye Z *et al*: Comparative analyses of CTCF and BORIS occupancies uncover two distinct classes of CTCF binding genomic regions. *Genome Biol* 2015, 16:161.
145. Renaud S, Pugacheva EM, Delgado MD, Braunschweig R, Abdullaev Z, Loukinov D, Benhattar J, Lobanenkov V: Expression of the CTCF-paralogous cancer-testis gene, brother of the regulator of imprinted sites (BORIS), is regulated by three alternative promoters modulated by CpG methylation and by CTCF and p53 transcription factors. *Nucleic Acids Res* 2007, 35(21):7372-7388.
146. Pugacheva EM, Suzuki T, Pack SD, Kosaka-Suzuki N, Yoon J, Vostrov AA, Barsov E, Strunnikov AV, Morse HC, 3rd, Loukinov D *et al*: The structural complexity of the human BORIS gene in gametogenesis and cancer. *PLoS One* 2010, 5(11):e13872.
147. Hoffmann MJ, Muller M, Engers R, Schulz WA: Epigenetic control of CTCFL/BORIS and OCT4 expression in urogenital malignancies. *Biochem Pharmacol* 2006, 72(11):1577-1588.
148. Martin-Kleiner I: BORIS in human cancers -- a review. *Eur J Cancer* 2012, 48(6):929-935.
149. Soltanian S, Dehghani H: BORIS: a key regulator of cancer stemness. *Cancer Cell Int* 2018, 18:154.
150. Woloszynska-Read A, James SR, Link PA, Yu J, Odunsi K, Karpf AR: DNA methylation-dependent regulation of BORIS/CTCFL expression in ovarian cancer. *Cancer Immun* 2007, 7:21.
151. Hong JA, Kang Y, Abdullaev Z, Flanagan PT, Pack SD, Fischette MR, Adnani MT, Loukinov DI, Vatolin S, Risinger JI *et al*: Reciprocal binding of CTCF and BORIS to the NY-ESO-1 promoter coincides with derepression of this cancer-testis gene in lung cancer cells. *Cancer Res* 2005, 65(17):7763-7774.
152. Link PA, Zhang W, Odunsi K, Karpf AR: BORIS/CTCFL mRNA isoform expression and epigenetic regulation in epithelial ovarian cancer. *Cancer Immun* 2013, 13:6.
153. Sleutels F, Soochit W, Bartkuhn M, Heath H, Dienstbach S, Bergmaier P, Franke V, Rosa-Garrido M, van de Nobelen S, Caesar L *et al*: The male germ cell gene regulator CTCFL is functionally different from

CTCF and binds CTCF-like consensus sites in a nucleosome composition-dependent manner. *Epigenetics Chromatin* 2012, 5(1):8.
154. Suzuki T, Kosaka-Suzuki N, Pack S, Shin DM, Yoon J, Abdullaev Z, Pugacheva E, Morse HC, 3rd, Loukinov D, Lobanenkov V: Expression of a testis-specific form of Gal3st1 (CST), a gene essential for spermatogenesis, is regulated by the CTCF paralogous gene BORIS. *Mol Cell Biol* 2010, 30(10):2473-2484.
155. Monk M, Hitchins M, Hawes S: Differential expression of the embryo/cancer gene ECSA(DPPA2), the cancer/testis gene BORIS and the pluripotency structural gene OCT4, in human preimplantation development. *Mol Hum Reprod* 2008, 14(6):347-355.
156. Jones TA, Ogunkolade BW, Szary J, Aarum J, Mumin MA, Patel S, Pieri CA, Sheer D: Widespread expression of BORIS/CTCFL in normal and cancer cells. *PLoS One* 2011, 6(7):e22399.
157. Rosa-Garrido M, Ceballos L, Alonso-Lecue P, Abraira C, Delgado MD, Gandarillas A: A cell cycle role for the epigenetic factor CTCF-L/BORIS. *PLoS One* 2012, 7(6):e39371.
158. Cheema Z, Hari-Gupta Y, Kita GX, Farrar D, Seddon I, Corr J, Klenova E: Expression of the cancer-testis antigen BORIS correlates with prostate cancer. *Prostate* 2014, 74(2):164-176.
159. Vatolin S, Abdullaev Z, Pack SD, Flanagan PT, Custer M, Loukinov DI, Pugacheva E, Hong JA, Morse H, 3rd, Schrump DS *et al*: Conditional expression of the CTCF-paralogous transcriptional factor BORIS in normal cells results in demethylation and derepression of MAGE-A1 and reactivation of other cancer-testis genes. *Cancer Res* 2005, 65(17):7751-7762.
160. Ghochikyan A, Mkrtichyan M, Loukinov D, Mamikonyan G, Pack SD, Movsesyan N, Ichim TE, Cribbs DH, Lobanenkov VV, Agadjanyan MG: Elicitation of T cell responses to histologically unrelated tumors by immunization with the novel cancer-testis antigen, brother of the regulator of imprinted sites. *J Immunol* 2007, 178(1):566-573.
161. Hoivik EA, Kusonmano K, Halle MK, Berg A, Wik E, Werner HM, Petersen K, Oyan AM, Kalland KH, Krakstad C *et al*: Hypomethylation of the CTCFL/BORIS promoter and aberrant expression during endometrial cancer progression suggests a role as an Epi-driver gene. *Oncotarget* 2014, 5(4):1052-1061.
162. Okabayashi K, Fujita T, Miyazaki J, Okada T, Iwata T, Hirao N, Noji S, Tsukamoto N, Goshima N, Hasegawa H *et al*: Cancer-testis antigen BORIS is a novel prognostic marker for patients with esophageal cancer. *Cancer Sci* 2012, 103(9):1617-1624.
163. Risinger JI, Chandramouli GV, Maxwell GL, Custer M, Pack S, Loukinov D, Aprelikova O, Litzi T, Schrump DS, Murphy SK *et al*: Global expression analysis of cancer/testis genes in uterine cancers reveals a high incidence of BORIS expression. *Clin Cancer Res* 2007, 13(6):1713-1719.
164. Soltanian S, Dehghani H, Matin MM, Bahrami AR: Expression analysis of BORIS during pluripotent, differentiated, cancerous, and non-cancerous cell states. *Acta Biochim Biophys Sin (Shanghai)* 2014, 46(8):647-658.
165. Novak Kujundzic R, Grbesa I, Ivkic M, Kruslin B, Konjevoda P, Gall Troselj K: Possible prognostic value of BORIS transcript variants ratio in laryngeal squamous cell carcinomas - a pilot study. *Pathol Oncol Res* 2014, 20(3):687-695.
166. Dougherty CJ, Ichim TE, Liu L, Reznik G, Min WP, Ghochikyan A, Agadjanyan MG, Reznik BN: Selective apoptosis of breast cancer cells by siRNA targeting of BORIS. *Biochem Biophys Res Commun* 2008, 370(1):109-112.
167. Nguyen P, Bar-Sela G, Sun L, Bisht KS, Cui H, Kohn E, Feinberg AP, Gius D: BAT3 and SET1A form a complex with CTCFL/BORIS to modulate H3K4 histone dimethylation and gene expression. *Mol Cell Biol* 2008, 28(21):6720-6729.
168. Zhao R, Chen K, Zhou J, He J, Liu J, Guan P, Li B, Qin Y: The prognostic role of BORIS and SOCS3 in human hepatocellular carcinoma. *Medicine (Baltimore)* 2017, 96(12):e6420.
169. Ogunkolade BW, Jones TA, Aarum J, Szary J, Owen N, Ottaviani D, Mumin MA, Patel S, Pieri CA, Silver AR *et al*: BORIS/CTCFL is an RNA-binding protein that associates with polysomes. *BMC Cell Biol* 2013, 14:52.
170. D'Arcy V, Abdullaev ZK, Pore N, Docquier F, Torrano V, Chernukhin I, Smart M, Farrar D, Metodiev M, Fernandez N *et al*: The potential of BORIS detected in the leukocytes of breast cancer patients as an early marker of tumorigenesis. *Clin Cancer Res* 2006, 12(20 Pt 1):5978-5986.
171. D'Arcy V, Pore N, Docquier F, Abdullaev ZK, Chernukhin I, Kita GX, Rai S, Smart M, Farrar D, Pack S *et al*: BORIS, a paralogue of the transcription factor, CTCF, is aberrantly expressed in breast tumours. *Br J Cancer* 2008, 98(3):571-579.

172. Cuffel C, Rivals JP, Zaugg Y, Salvi S, Seelentag W, Speiser DE, Lienard D, Monnier P, Romero P, Bron L *et al*: Pattern and clinical significance of cancer-testis gene expression in head and neck squamous cell carcinoma. *Int J Cancer* 2011, 128(11):2625-2634.
173. Chen K, Huang W, Huang B, Wei Y, Li B, Ge Y, Qin Y: BORIS, brother of the regulator of imprinted sites, is aberrantly expressed in hepatocellular carcinoma. *Genet Test Mol Biomarkers* 2013, 17(2):160-165.
174. Ulaner GA, Vu TH, Li T, Hu JF, Yao XM, Yang Y, Gorlick R, Meyers P, Healey J, Ladanyi M *et al*: Loss of imprinting of IGF2 and H19 in osteosarcoma is accompanied by reciprocal methylation changes of a CTCF-binding site. *Human molecular genetics* 2003, 12(5):535-549.
175. de Necochea-Campion R, Ghochikyan A, Josephs SF, Zacharias S, Woods E, Karimi-Busheri F, Alexandrescu DT, Chen CS, Agadjanyan MG, Carrier E: Expression of the epigenetic factor BORIS (CTCFL) in the human genome. *J Transl Med* 2011, 9:213.
176. Campbell AE, Martinez SR, Miranda JJ: Molecular architecture of CTCFL. *Biochem Biophys Res Commun* 2010, 396(3):648-650.
177. Martinez SR, Miranda JL: CTCF terminal segments are unstructured. *Protein science : a publication of the Protein Society* 2010, 19(5):1110-1116.
178. Tompa P, Fuxreiter M: Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions. *Trends in biochemical sciences* 2008, 33(1):2-8.
179. Uversky VN: Intrinsic Disorder, Protein-Protein Interactions, and Disease. *Advances in protein chemistry and structural biology* 2018, 110:85-121.
180. UniProt Consortium T: UniProt: the universal protein knowledgebase. *Nucleic Acids Res* 2018, 46(5):2699.
181. Rivero-Hinojosa S, Kang S, Lobanenkov VV, Zentner GE: Testis-specific transcriptional regulators selectively occupy BORIS-bound CTCF target regions in mouse male germ cells. *Sci Rep* 2017, 7:41279.
182. Kang Y, Hong JA, Chen GA, Nguyen DM, Schrump DS: Dynamic transcriptional regulatory complexes including BORIS, CTCF and Sp1 modulate NY-ESO-1 expression in lung cancer cells. *Oncogene* 2007, 26(30):4394-4403.
183. van de Nobelen S, Rosa-Garrido M, Leers J, Heath H, Soochit W, Joosen L, Jonkers I, Demmers J, van der Reijden M, Torrano V *et al*: CTCF regulates the local epigenetic state of ribosomal DNA repeats. *Epigenetics Chromatin* 2010, 3(1):19.
184. Jelinic P, Stehle JC, Shaw P: The testis-specific factor CTCFL cooperates with the protein methyltransferase PRMT7 in H19 imprinting control region methylation. *PLoS Biol* 2006, 4(11):e355.
185. Jabbari K, Heger P, Sharma R, Wiehe T: The Diverging Routes of BORIS and CTCF: An Interactomic and Phylogenomic Analysis. *Life (Basel)* 2018, 8(1).
186. Lobanenkov VV, Zentner GE: Discovering a binary CTCF code with a little help from BORIS. *Nucleus* 2018, 9(1):33-41.
187. Kosaka-Suzuki N, Suzuki T, Pugacheva EM, Vostrov AA, Morse HC, 3rd, Loukinov D, Lobanenkov V: Transcription factor BORIS (Brother of the Regulator of Imprinted Sites) directly induces expression of a cancer-testis antigen, TSP50, through regulated binding of BORIS to the promoter. *J Biol Chem* 2011, 286(31):27378-27388.
188. Bergmaier P, Weth O, Dienstbach S, Boettger T, Galjart N, Mernberger M, Bartkuhn M, Renkawitz R: Choice of binding sites for CTCFL compared to CTCF is driven by chromatin and by sequence preference. *Nucleic Acids Res* 2018, 46(14):7097-7107.
189. Gaykalova D, Vatapalli R, Glazer CA, Bhan S, Shao C, Sidransky D, Ha PK, Califano JA: Dose-dependent activation of putative oncogene SBSN by BORIS. *PLoS One* 2012, 7(7):e40389.
190. Liu Q, Chen K, Liu Z, Huang Y, Zhao R, Wei L, Yu X, He J, Liu J, Qi J *et al*: BORIS up-regulates OCT4 via histone methylation to promote cancer stem cell-like properties in human liver cancer cells. *Cancer Lett* 2017, 403:165-174.
191. Zampieri M, Ciccarone F, Palermo R, Cialfi S, Passananti C, Chiaretti S, Nocchia D, Talora C, Screpanti I, Caiafa P: The epigenetic factor BORIS/CTCFL regulates the NOTCH3 gene expression in cancer cells. *Biochim Biophys Acta* 2014, 1839(9):813-825.
192. Bhan S, Negi SS, Shao C, Glazer CA, Chuang A, Gaykalova DA, Sun W, Sidransky D, Ha PK, Califano JA: BORIS binding to the promoters of cancer testis antigens, MAGEA2, MAGEA3, and MAGEA4, is associated with their transcriptional activation in lung cancer. *Clin Cancer Res* 2011, 17(13):4267-4276.
193. Sun L, Huang L, Nguyen P, Bisht KS, Bar-Sela G, Ho AS, Bradbury CM, Yu W, Cui H, Lee S *et al*: DNA methyltransferase 1 and 3B activate BAG-1 expression via recruitment of CTCFL/BORIS and modulation of promoter histone methylation. *Cancer Res* 2008, 68(8):2726-2735.
194. Holwerda S, de Laat W: Chromatin loops, gene positioning, and gene expression. *Front Genet* 2012, 3:217.

195. Teplyakov E, Wu Q, Liu J, Pugacheva EM, Loukinov D, Boukaba A, Lobanenkov V, Strunnikov A: The downregulation of putative anticancer target BORIS/CTCFL in an addicted myeloid cancer cell line modulates the expression of multiple protein coding and ncRNA genes. *Oncotarget* 2017, 8(43):73448-73468.
196. Pugacheva EM, Teplyakov E, Wu Q, Li J, Chen C, Meng C, Liu J, Robinson S, Loukinov D, Boukaba A *et al*: The cancer-associated CTCFL/BORIS protein targets multiple classes of genomic repeats, with a distinct binding and functional preference for humanoid-specific SVA transposable elements. *Epigenetics Chromatin* 2016, 9(1):35.
197. Smith IM, Glazer CA, Mithani SK, Ochs MF, Sun W, Bhan S, Vostrov A, Abdullaev Z, Lobanenkov V, Gray A *et al*: Coordinated activation of candidate proto-oncogenes and cancer testes antigens via promoter demethylation in head and neck cancer and lung cancer. *PLoS One* 2009, 4(3):e4961.
198. Kim S, Yu NK, Kaang BK: CTCF as a multifunctional protein in genome regulation and gene expression. *Exp Mol Med* 2015, 47:e166.
199. Lewis A, Murrell A: Genomic imprinting: CTCF protects the boundaries. *Curr Biol* 2004, 14(7):R284-286.
200. Bell AC, Felsenfeld G: Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. *Nature* 2000, 405(6785):482-485.
201. Hark AT, Schoenherr CJ, Katz DJ, Ingram RS, Levorse JM, Tilghman SM: CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature* 2000, 405(6785):486-489.
202. Nguyen P, Cui H, Bisht KS, Sun L, Patel K, Lee RS, Kugoh H, Oshimura M, Feinberg AP, Gius D: CTCFL/BORIS is a methylation-independent DNA-binding protein that preferentially binds to the paternal H19 differentially methylated region. *Cancer Res* 2008, 68(14):5546-5551.
203. Chaligne R, Heard E: X-chromosome inactivation in development and cancer. *FEBS Lett* 2014, 588(15):2514-2522.
204. Boumil RM, Ogawa Y, Sun BK, Huynh KD, Lee JT: Differential methylation of Xite and CTCF sites in Tsix mirrors the pattern of X-inactivation choice in mice. *Mol Cell Biol* 2006, 26(6):2109-2117.
205. Yang F, Deng X, Ma W, Berletch JB, Rabaia N, Wei G, Moore JM, Filippova GN, Xu J, Liu Y *et al*: The lncRNA Firre anchors the inactive X chromosome to the nucleolus by binding CTCF and maintains H3K27me3 methylation. *Genome Biol* 2015, 16:52.
206. Chao W, Huynh KD, Spencer RJ, Davidow LS, Lee JT: CTCF, a candidate trans-acting factor for X-inactivation choice. *Science (New York, NY)* 2002, 295(5553):345-347.
207. Kung JT, Kesner B, An JY, Ahn JY, Cifuentes-Rojas C, Colognori D, Jeon Y, Szanto A, del Rosario BC, Pinter SF *et al*: Locus-specific targeting to the X chromosome revealed by the RNA interactome of CTCF. *Mol Cell* 2015, 57(2):361-375.
208. Spencer RJ, del Rosario BC, Pinter SF, Lessing D, Sadreyev RI, Lee JT: A boundary element between Tsix and Xist binds the chromatin insulator Ctcf and contributes to initiation of X-chromosome inactivation. *Genetics* 2011, 189(2):441-454.
209. Xu N, Donohoe ME, Silva SS, Lee JT: Evidence that homologous X-chromosome pairing requires transcription and Ctcf protein. *Nat Genet* 2007, 39(11):1390-1396.
210. Sun S, Del Rosario BC, Szanto A, Ogawa Y, Jeon Y, Lee JT: Jpx RNA activates Xist by evicting CTCF. *Cell* 2013, 153(7):1537-1551.
211. Berletch JB, Ma W, Yang F, Shendure J, Noble WS, Disteche CM, Deng X: Escape from X inactivation varies in mouse tissues. *PLoS Genet* 2015, 11(3):e1005079.
212. Kemp CJ, Moore JM, Moser R, Bernard B, Teater M, Smith LE, Rabaia NA, Gurley KE, Guinney J, Busch SE *et al*: CTCF haploinsufficiency destabilizes DNA methylation and predisposes to cancer. *Cell Rep* 2014, 7(4):1020-1029.
213. Tiffen JC, Bailey CG, Marshall AD, Metierre C, Feng Y, Wang Q, Watson SL, Holst J, Rasko JE: The cancer-testis antigen BORIS phenocopies the tumor suppressor CTCF in normal and neoplastic cells. *Int J Cancer* 2013, 133(7):1603-1613.
214. Alberti L, Losi L, Leyvraz S, Benhattar J: Different Effects of BORIS/CTCFL on Stemness Gene Expression, Sphere Formation and Cell Survival in Epithelial Cancer Stem Cells. *PLoS One* 2015, 10(7):e0132977.
215. Garikapati KR, Patel N, Makani VKK, Cilamkoti P, Bhadra U, Bhadra MP: Down-regulation of BORIS/CTCFL efficiently regulates cancer stemness and metastasis in MYCN amplified neuroblastoma cell line by modulating Wnt/beta-catenin signaling pathway. *Biochem Biophys Res Commun* 2017, 484(1):93-99.

216. Loukinov D, Ghochikyan A, Mkrtichyan M, Ichim TE, Lobanenkov VV, Cribbs DH, Agadjanyan MG: Antitumor efficacy of DNA vaccination to the epigenetically acting tumor promoting transcription factor BORIS and CD80 molecular adjuvant. *J Cell Biochem* 2006, 98(5):1037-1043.
217. Mkrtichyan M, Ghochikyan A, Loukinov D, Davtyan H, Ichim TE, Cribbs DH, Lobanenkov VV, Agadjanyan MG: DNA, but not protein vaccine based on mutated BORIS antigen significantly inhibits tumor growth and prolongs the survival of mice. *Gene Ther* 2008, 15(1):61-64.
218. Mkrtichyan M, Ghochikyan A, Davtyan H, Movsesyan N, Loukinov D, Lobanenkov V, Cribbs DH, Laust AK, Nelson EL, Agadjanyan MG: Cancer-testis antigen, BORIS based vaccine delivered by dendritic cells is extremely effective against a very aggressive and highly metastatic mouse mammary carcinoma. *Cell Immunol* 2011, 270(2):188-197.
219. AmericanCancerSociety: Cancer Facts & Figures. In.; 2018.
220. Ohayon S: BORIS/CTCFL is an epigenetic modifier of melanoma tumorigenicity. McGill University; 2015.
221. Spatz A, Borg C, Feunteun J: X-chromosome genetics and human cancer. *Nat Rev Cancer* 2004, 4(8):617-629.
222. van Kempen LC, Redpath M, Elchebly M, Klein KO, Papadakis AI, Wilmott JS, Scolyer RA, Edqvist PH, Ponten F, Schadendorf D *et al*: The protein phosphatase 2A regulatory subunit PR70 is a gonosomal melanoma tumor suppressor gene. *Science translational medicine* 2016, 8(369):369ra177.
223. Berletch JB, Yang F, Disteche CM: Escape from X inactivation in mice and humans. *Genome Biol* 2010, 11(6):213.
224. Meerbrey KL, Hu G, Kessler JD, Roarty K, Li MZ, Fang JE, Herschkowitz JI, Burrows AE, Ciccia A, Sun T *et al*: The pINDUCER lentiviral toolkit for inducible RNA interference in vitro and in vivo. *Proc Natl Acad Sci U S A* 2011, 108(9):3665-3670.
225. Smith CL, Blake JA, Kadin JA, Richardson JE, Bult CJ: Mouse Genome Database (MGD)-2018: knowledgebase for the laboratory mouse. *Nucleic Acids Res* 2018, 46(D1):D836-d842.
226. Rao X, Huang X, Zhou Z, Lin X: An improvement of the 2^(-delta delta CT) method for quantitative real-time polymerase chain reaction data analysis. *Biostatistics, bioinformatics and biomathematics* 2013, 3(3):71-85.
227. Nesvizhskii AI, Keller A, Kolker E, Aebersold R: A statistical model for identifying proteins by tandem mass spectrometry. *Analytical chemistry* 2003, 75(17):4646-4658.
228. Benjamini Y HY: Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B (Methodological)* 1995, 57(1):289-300.
229. Hu Y, Vinayagam A, Nand A, Comjean A, Chung V, Hao T, Mohr SE, Perrimon N: Molecular Interaction Search Tool (MIST): an integrated resource for mining gene and protein interaction data. *Nucleic Acids Res* 2018, 46(D1):D567-d574.
230. Andrews S: FastQC: a quality control tool for high throughput sequence data. 2010.
231. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR: STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 2013, 29(1):15-21.
232. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R: The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 2009, 25(16):2078-2079.
233. Anders S, McCarthy DJ, Chen Y, Okoniewski M, Smyth GK, Huber W, Robinson MD: Count-based differential expression analysis of RNA sequencing data using R and Bioconductor. *Nature protocols* 2013, 8(9):1765-1786.
234. Anders S, Pyl PT, Huber W: HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* 2015, 31(2):166-169.
235. Robinson MD, McCarthy DJ, Smyth GK: edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2010, 26(1):139-140.
236. Bushnell B: BBMap short-read aligner, and other bioinformatics tools. 2014.
237. Parker S: A toolkit for ATAC-seq data -- atactk 0.1.5. 2015.
238. Martin M: Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnetjournal* 2011, 17:10-12.
239. Li H: Aligning sequence reads, clone sequences and asembly contigs with BWA-MEM. *arXiv:13033997v2* 2013:1-3.
240. Wysoker AT, K. Fennell, T: Picard tools version 2.17.10. 2009.
241. Feng J, Liu T, Qin B, Zhang Y, Liu XS: Identifying ChIP-seq enrichment using MACS. *Nature protocols* 2012, 7(9):1728-1740.

242. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W *et al*: Model-based analysis of ChIP-Seq (MACS). *Genome Biol* 2008, 9(9):R137.
243. Mi H, Huang X, Muruganujan A, Tang H, Mills C, Kang D, Thomas PD: PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. *Nucleic Acids Res* 2017, 45(D1):D183-d189.
244. Supek F, Bosnjak M, Skunca N, Smuc T: REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One* 2011, 6(7):e21800.
245. Kamburov A, Stelzl U, Lehrach H, Herwig R: The ConsensusPathDB interaction database: 2013 update. *Nucleic Acids Res* 2013, 41(Database issue):D793-800.
246. Kamburov A, Wierling C, Lehrach H, Herwig R: ConsensusPathDB--a database for integrating human functional interaction networks. *Nucleic Acids Res* 2009, 37(Database issue):D623-628.
247. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES *et al*: Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 2005, 102(43):15545-15550.
248. Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P: The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell systems* 2015, 1(6):417-425.
249. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T: Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 2003, 13(11):2498-2504.
250. Merico D, Isserlin R, Stueker O, Emili A, Bader GD: Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. *PLoS One* 2010, 5(11):e13984.
251. Kucera M, Isserlin R, Arkhangorodsky A, Bader GD: AutoAnnotate: A Cytoscape app for summarizing networks with semantic annotations. *F1000Research* 2016, 5:1717.
252. Oesper L, Merico D, Isserlin R, Bader GD: WordCloud: a Cytoscape plugin to create a visual semantic summary of networks. *Source code for biology and medicine* 2011, 6:7.
253. Janky R, Verfaillie A, Imrichova H, Van de Sande B, Standaert L, Christiaens V, Hulselmans G, Herten K, Naval Sanchez M, Potier D *et al*: iRegulon: from a gene list to a gene regulatory network using large motif and track collections. *PLoS computational biology* 2014, 10(7):e1003731.
254. Balaton BP, Cotton AM, Brown CJ: Derivation of consensus inactivation status for X-linked genes from genome-wide studies. *Biol Sex Differ* 2015, 6:35.
255. Kazi JU, Kabir NN, Ronnstrand L: Brain-Expressed X-linked (BEX) proteins in human cancers. *Biochim Biophys Acta* 2015, 1856(2):226-233.
256. Naderi A, Liu J, Bennett IC: BEX2 regulates mitochondrial apoptosis and G1 cell cycle in breast cancer. *Int J Cancer* 2010, 126(7):1596-1610.
257. Zhou X, Xu X, Meng Q, Hu J, Zhi T, Shi Q, Yu R: Bex2 is critical for migration and invasion in malignant glioma cells. *Journal of molecular neuroscience : MN* 2013, 50(1):78-87.
258. Wu J, Du W, Wang X, Wei L, Pan Y, Wu X, Zhang J, Pei D: Ras-related protein Rap2c promotes the migration and invasion of human osteosarcoma cells. *Oncology letters* 2018, 15(4):5352-5358.
259. Jones J, Wang H, Karanam B, Theodore S, Dean-Colomb W, Welch DR, Grizzle W, Yates C: Nuclear localization of Kaiso promotes the poorly differentiated phenotype and EMT in infiltrating ductal carcinomas. *Clinical & experimental metastasis* 2014, 31(5):497-510.
260. Jones J, Wang H, Zhou J, Hardy S, Turner T, Austin D, He Q, Wells A, Grizzle WE, Yates C: Nuclear Kaiso indicates aggressive prostate cancers and promotes migration and invasiveness of prostate cancer cells. *The American journal of pathology* 2012, 181(5):1836-1846.
261. van Roy FM, McCrea PD: A role for Kaiso-p120ctn complexes in cancer? *Nat Rev Cancer* 2005, 5(12):956-964.
262. Vermeulen JF, van de Ven RA, Ercan C, van der Groep P, van der Wall E, Bult P, Christgen M, Lehmann U, Daniel J, van Diest PJ *et al*: Nuclear Kaiso expression is associated with high grade and triple-negative invasive breast cancer. *PLoS One* 2012, 7(5):e37864.
263. Thompson BJ, Sahai E: MST kinases in development and disease. *The Journal of cell biology* 2015, 210(6):871-882.
264. Xiong W, Knox AJ, Xu M, Kiseljak-Vassiliades K, Colgan SP, Brodsky KS, Kleinschmidt-Demasters BK, Lillehei KO, Wierman ME: Mammalian Ste20-like kinase 4 promotes pituitary cell proliferation and survival under hypoxia. *Molecular endocrinology (Baltimore, Md)* 2015, 29(3):460-472.

265. Bublik DR, Bursac S, Sheffer M, Orsolic I, Shalit T, Tarcic O, Kotler E, Mouhadeb O, Hoffman Y, Fuchs G *et al*: Regulatory module involving FGF13, miR-504, and p53 regulates ribosomal biogenesis and supports cancer cell survival. *Proc Natl Acad Sci U S A* 2017, 114(4):E496-e505.
266. Yu L, Toriseva M, Tuomala M, Seikkula H, Elo T, Tuomela J, Kallajoki M, Mirtti T, Taimen P, Bostrom PJ *et al*: Increased expression of fibroblast growth factor 13 in prostate cancer is associated with shortened time to biochemical recurrence after radical prostatectomy. *Int J Cancer* 2016, 139(1):140-152.
267. Asano J, Nakano A, Oda A, Amou H, Hiasa M, Takeuchi K, Miki H, Nakamura S, Harada T, Fujii S *et al*: The serine/threonine kinase Pim-2 is a novel anti-apoptotic mediator in myeloma cells. *Leukemia* 2011, 25(7):1182-1188.
268. Blanco-Aparicio C, Carnero A: Pim kinases in cancer: diagnostic, prognostic and treatment opportunities. *Biochem Pharmacol* 2013, 85(5):629-643.
269. Guo S, Sun F, Guo Z, Li W, Alfano A, Chen H, Magyar CE, Huang J, Chai TC, Qiu S *et al*: Tyrosine kinase ETK/BMX is up-regulated in bladder cancer and predicts poor prognosis in patients with cystectomy. *PLoS One* 2011, 6(3):e17778.
270. Jarboe JS, Dutta S, Velu SE, Willey CD: Mini-review: bmx kinase inhibitors for cancer therapy. *Recent patents on anti-cancer drug discovery* 2013, 8(3):228-238.
271. Atanackovic D, Blum I, Cao Y, Wenzel S, Bartels K, Faltz C, Hossfeld DK, Hegewisch-Becker S, Bokemeyer C, Leuwer R: Expression of cancer-testis antigens as possible targets for antigen-specific immunotherapy in head and neck squamous cell carcinoma. *Cancer biology & therapy* 2006, 5(9):1218-1225.
272. Chen YT, Panarelli NC, Piotti KC, Yantiss RK: Cancer-testis antigen expression in digestive tract carcinomas: frequent expression in esophageal squamous cell carcinoma and its precursor lesions. *Cancer immunology research* 2014, 2(5):480-486.
273. Fang J, Yu Z, Lian M, Ma H, Tai J, Zhang L, Han D: Knockdown of zinc finger protein, X-linked (ZFX) inhibits cell proliferation and induces apoptosis in human laryngeal squamous cell carcinoma. *Molecular and cellular biochemistry* 2012, 360(1-2):301-307.
274. Weisberg SP, Smith-Raska MR, Esquilin JM, Zhang J, Arenzana TL, Lau CM, Churchill M, Pan H, Klinakis A, Dixon JE *et al*: ZFX controls propagation and prevents differentiation of acute T-lymphoblastic and myeloid leukemia. *Cell Rep* 2014, 6(3):528-540.
275. Wu S, Lao XY, Sun TT, Ren LL, Kong X, Wang JL, Wang YC, Du W, Yu YN, Weng YR *et al*: Knockdown of ZFX inhibits gastric cancer cell growth in vitro and in vivo via downregulating the ERK-MAPK pathway. *Cancer Lett* 2013, 337(2):293-300.
276. Ji X, Jin S, Qu X, Li K, Wang H, He H, Guo F, Dong L: Lysine-specific demethylase 5C promotes hepatocellular carcinoma cell invasion through inhibition BMP7 expression. *BMC Cancer* 2015, 15:801.
277. Stein J, Majores M, Rohde M, Lim S, Schneider S, Krappe E, Ellinger J, Dietel M, Stephan C, Jung K *et al*: KDM5C is overexpressed in prostate cancer and is a prognostic marker for prostate-specific antigen-relapse following radical prostatectomy. *The American journal of pathology* 2014, 184(9):2430-2437.
278. Torchy MP, Hamiche A, Klaholz BP: Structure and function insights into the NuRD chromatin remodeling complex. *Cell Mol Life Sci* 2015, 72(13):2491-2507.
279. Yeh HH, Tseng YF, Hsu YC, Lan SH, Wu SY, Raghavaraju G, Cheng DE, Lee YR, Chang TY, Chow NH *et al*: Ras induces experimental lung metastasis through up-regulation of RbAp46 to suppress RECK promoter activity. *BMC Cancer* 2015, 15:172.
280. Loukinov D: Targeting CTCFL/BORIS for the immunotherapy of cancer. *Cancer immunology, immunotherapy : CII* 2018.
281. Venny. An interactive tool for comparing lists with Venn's diagrams. [http://bioinfogp.cnb.csic.es/tools/venny/index.html]
282. Kudo-Saito C, Shirako H, Takeuchi T, Kawakami Y: Cancer metastasis is accelerated through immunosuppression during Snail-induced EMT of cancer cells. *Cancer Cell* 2009, 15(3):195-206.
283. Robert G, Gaggioli C, Bailet O, Chavey C, Abbe P, Aberdam E, Sabatie E, Cano A, Garcia de Herreros A, Ballotti R *et al*: SPARC represses E-cadherin and induces mesenchymal transition during melanoma development. *Cancer Res* 2006, 66(15):7516-7523.
284. Narayanan A, Ruyechan WT, Kristie TM: The coactivator host cell factor-1 mediates Set1 and MLL1 H3K4 trimethylation at herpesvirus immediate early promoters for initiation of infection. *Proc Natl Acad Sci U S A* 2007, 104(26):10835-10840.

285. Wysocka J, Myers MP, Laherty CD, Eisenman RN, Herr W: Human Sin3 deacetylase and trithorax-related Set1/Ash2 histone H3-K4 methyltransferase are tethered together selectively by the cell-proliferation factor HCF-1. *Genes Dev* 2003, 17(7):896-911.
286. Cavellan E, Asp P, Percipalle P, Farrants AK: The WSTF-SNF2h chromatin remodeling complex interacts with several nuclear proteins in transcription. *J Biol Chem* 2006, 281(24):16264-16271.
287. Sadeghifar F, Bohm S, Vintermist A, Ostlund Farrants AK: The B-WICH chromatin-remodelling complex regulates RNA polymerase III transcription by promoting Max-dependent c-Myc binding. *Nucleic Acids Res* 2015, 43(9):4477-4490.
288. Vintermist A, Bohm S, Sadeghifar F, Louvet E, Mansen A, Percipalle P, Ostlund Farrants AK: The chromatin remodelling complex B-WICH changes the chromatin structure and recruits histone acetyl-transferases to active rRNA genes. *PLoS One* 2011, 6(4):e19184.
289. Han D, Chen Q, Shi J, Zhang F, Yu X: CTCF participates in DNA damage response via poly(ADP-ribosyl)ation. *Sci Rep* 2017, 7:43530.
290. Hilmi K, Jangal M, Marques M, Zhao T, Saad A, Zhang C, Luo VM, Syme A, Rejon C, Yu Z *et al*: CTCF facilitates DNA double-strand break repair by enhancing homologous recombination repair. *Science advances* 2017, 3(5):e1601898.
291. Lang F, Li X, Zheng W, Li Z, Lu D, Chen G, Gong D, Yang L, Fu J, Shi P *et al*: CTCF prevents genomic instability by promoting homologous recombination-directed DNA double-strand break repair. *Proc Natl Acad Sci U S A* 2017, 114(41):10912-10917.
292. Bartek J, Lukas C, Lukas J: Checking on DNA damage in S phase. *Nat Rev Mol Cell Biol* 2004, 5(10):792-804.
293. Rogakou EP, Pilch DR, Orr AH, Ivanova VS, Bonner WM: DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *J Biol Chem* 1998, 273(10):5858-5868.
294. Rogakou EP, Boon C, Redon C, Bonner WM: Megabase chromatin domains involved in DNA double-strand breaks in vivo. *The Journal of cell biology* 1999, 146(5):905-916.
295. Falck J, Mailand N, Syljuasen RG, Bartek J, Lukas J: The ATM-Chk2-Cdc25A checkpoint pathway guards against radioresistant DNA synthesis. *Nature* 2001, 410(6830):842-847.
296. Falck J, Petrini JH, Williams BR, Lukas J, Bartek J: The DNA damage-dependent intra-S phase checkpoint is regulated by parallel pathways. *Nat Genet* 2002, 30(3):290-294.
297. Kim JS, Krasieva TB, LaMorte V, Taylor AM, Yokomori K: Specific recruitment of human cohesin to laser-induced DNA damage. *J Biol Chem* 2002, 277(47):45149-45153.
298. Kitagawa R, Bakkenist CJ, McKinnon PJ, Kastan MB: Phosphorylation of SMC1 is a critical downstream event in the ATM-NBS1-BRCA1 pathway. *Genes Dev* 2004, 18(12):1423-1438.
299. Yazdi PT, Wang Y, Zhao S, Patel N, Lee EY, Qin J: SMC1 is a downstream effector in the ATM/NBS1 branch of the human S-phase checkpoint. *Genes Dev* 2002, 16(5):571-582.
300. Davar D, Beumer JH, Hamieh L, Tawbi H: Role of PARP inhibitors in cancer biology and therapy. *Current medicinal chemistry* 2012, 19(23):3907-3921.
301. Orphanides G, LeRoy G, Chang CH, Luse DS, Reinberg D: FACT, a factor that facilitates transcript elongation through nucleosomes. *Cell* 1998, 92(1):105-116.
302. Belotserkovskaya R, Oh S, Bondarenko VA, Orphanides G, Studitsky VM, Reinberg D: FACT facilitates transcription-dependent nucleosome alteration. *Science (New York, NY)* 2003, 301(5636):1090-1093.
303. Hondele M, Stuwe T, Hassler M, Halbach F, Bowman A, Zhang ET, Nijmeijer B, Kotthoff C, Rybin V, Amlacher S *et al*: Structural basis of histone H2A-H2B recognition by the essential chaperone FACT. *Nature* 2013, 499(7456):111-114.
304. Kemble DJ, Whitby FG, Robinson H, McCullough LL, Formosa T, Hill CP: Structure of the Spt16 middle domain reveals functional features of the histone chaperone FACT. *J Biol Chem* 2013, 288(15):10188-10194.
305. Orphanides G, Wu WH, Lane WS, Hampsey M, Reinberg D: The chromatin-specific transcription elongation factor FACT comprises human SPT16 and SSRP1 proteins. *Nature* 1999, 400(6741):284-288.
306. VanDemark AP, Blanksma M, Ferris E, Heroux A, Hill CP, Formosa T: The structure of the yFACT Pob3-M domain, its interaction with the DNA replication factor RPA, and a potential role in nucleosome deposition. *Mol Cell* 2006, 22(3):363-374.
307. Foltman M, Evrin C, De Piccoli G, Jones RC, Edmondson RD, Katou Y, Nakato R, Shirahige K, Labib K: Eukaryotic replisome components cooperate to process histones during chromosome replication. *Cell Rep* 2013, 3(3):892-904.

308. Jamai A, Puglisi A, Strubin M: Histone chaperone spt16 promotes redeposition of the original h3-h4 histones evicted by elongating RNA polymerase. *Mol Cell* 2009, 35(3):377-383.
309. Yang J, Zhang X, Feng J, Leng H, Li S, Xiao J, Liu S, Xu Z, Xu J, Li D *et al*: The Histone Chaperone FACT Contributes to DNA Replication-Coupled Nucleosome Assembly. *Cell Rep* 2016, 16(12):3414.
310. Dinant C, Ampatziadis-Michailidis G, Lans H, Tresini M, Lagarou A, Grosbart M, Theil AF, van Cappellen WA, Kimura H, Bartek J *et al*: Enhanced chromatin dynamics by FACT promotes transcriptional restart after UV-induced DNA damage. *Mol Cell* 2013, 51(4):469-479.
311. Oliveira DV, Kato A, Nakamura K, Ikura T, Okada M, Kobayashi J, Yanagihara H, Saito Y, Tauchi H, Komatsu K: Histone chaperone FACT regulates homologous recombination by chromatin remodeling through interaction with RNF20. *J Cell Sci* 2014, 127(Pt 4):763-772.
312. Piquet S, Le Parc F, Bai SK, Chevallier O, Adam S, Polo SE: The Histone Chaperone FACT Coordinates H2A.X-Dependent Signaling and Repair of DNA Damage. *Mol Cell* 2018.
313. Heo K, Kim H, Choi SH, Choi J, Kim K, Gu J, Lieber MR, Yang AS, An W: FACT-mediated exchange of histone variant H2AX regulated by phosphorylation of H2AX and ADP-ribosylation of Spt16. *Mol Cell* 2008, 30(1):86-97.
314. Gao J, Wang J, Wang Y, Dai W, Lu L: Regulation of Pax6 by CTCF during induction of mouse ES cell differentiation. *PLoS One* 2011, 6(6):e20954.
315. Li T, Lu Z, Lu L: Regulation of eye development by transcription control of CCCTC binding factor (CTCF). *J Biol Chem* 2004, 279(26):27575-27583.
316. Li T, Lu Z, Lu L: Pax6 regulation in retinal cells by CCCTC binding factor. *Invest Ophthalmol Vis Sci* 2006, 47(12):5218-5226.
317. Tsui S, Wang J, Wang L, Dai W, Lu L: CTCF-Mediated and Pax6-Associated Gene Expression in Corneal Epithelial Cell-Specific Differentiation. *PLoS One* 2016, 11(9):e0162071.
318. Wu D, Li T, Lu Z, Dai W, Xu M, Lu L: Effect of CTCF-binding motif on regulation of PAX6 transcription. *Invest Ophthalmol Vis Sci* 2006, 47(6):2422-2429.
319. Muratovska A, Zhou C, He S, Goodyer P, Eccles MR: Paired-Box genes are frequently expressed in cancer and often required for cancer cell survival. *Oncogene* 2003, 22(39):7989-7997.
320. Zhao X, Yue W, Zhang L, Ma L, Jia W, Qian Z, Zhang C, Wang Y: Downregulation of PAX6 by shRNA inhibits proliferation and cell cycle progression of human non-small cell lung cancer cell lines. *PLoS One* 2014, 9(1):e85738.
321. Hellwinkel OJ, Kedia M, Isbarn H, Budaus L, Friedrich MG: Methylation of the TPEF- and PAX6-promoters is increased in early bladder cancer and in normal mucosa adjacent to pTa tumours. *BJU Int* 2008, 101(6):753-757.
322. Huang BS, Luo QZ, Han Y, Li XB, Cao LJ, Wu LX: microRNA-223 promotes the growth and invasion of glioblastoma cells by targeting tumor suppressor PAX6. *Oncol Rep* 2013, 30(5):2263-2269.
323. Robson EJ, He SJ, Eccles MR: A PANorama of PAX genes in cancer and development. *Nat Rev Cancer* 2006, 6(1):52-62.
324. Mascarenhas JB, Young KP, Littlejohn EL, Yoo BK, Salgia R, Lang D: PAX6 is expressed in pancreatic cancer and actively participates in cancer progression through activation of the MET tyrosine kinase receptor gene. *J Biol Chem* 2009, 284(40):27524-27532.
325. Bai SW, Li B, Zhang H, Jonas JB, Zhao BW, Shen L, Wang YC: Pax6 regulates proliferation and apoptosis of human retinoblastoma cells. *Invest Ophthalmol Vis Sci* 2011, 52(7):4560-4570.
326. Li Y, Li Y, Liu Y, Xie P, Li F, Li G: PAX6, a novel target of microRNA-7, promotes cellular proliferation and invasion in human colorectal cancer cells. *Dig Dis Sci* 2014, 59(3):598-606.
327. Zong X, Yang H, Yu Y, Zou D, Ling Z, He X, Meng X: Possible role of Pax-6 in promoting breast cancer cell proliferation and tumorigenesis. *BMB Reports* 2011, 44(9):595-600.
328. Nemlich Y, Baruch EN, Besser MJ, Shoshan E, Bar-Eli M, Anafi L, Barshack I, Schachter J, Ortenberg R, Markel G: ADAR1-mediated regulation of melanoma invasion. *Nat Commun* 2018, 9(1):2154.
329. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E *et al*: The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer discovery* 2012, 2(5):401-404.
330. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, Sun Y, Jacobsen A, Sinha R, Larsson E *et al*: Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Science signaling* 2013, 6(269):pl1.

331. Furuta J, Nobeyama Y, Umebayashi Y, Otsuka F, Kikuchi K, Ushijima T: Silencing of Peroxiredoxin 2 and aberrant methylation of 33 CpG islands in putative promoter regions in human malignant melanomas. *Cancer Res* 2006, 66(12):6080-6086.
332. Buenrostro JD, Wu B, Chang HY, Greenleaf WJ: ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Curr Protoc Mol Biol* 2015, 109:21 29 21-29.
333. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ: Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods* 2013, 10(12):1213-1218.
334. Ou J, Liu H, Yu J, Kelliher MA, Castilla LH, Lawson ND, Zhu LJ: ATACseqQC: a Bioconductor package for post-alignment quality assessment of ATAC-seq data. *BMC Genomics* 2018, 19(1):169.
335. Abisoye-Ogunniyan A, Lin H, Ghebremedhin A, Salam AB, Karanam B, Theodore S, Jones-Trich J, Davis M, Grizzle W, Wang H *et al*: Transcriptional repressor Kaiso promotes epithelial to mesenchymal transition and metastasis in prostate cancer through direct regulation of miR-200c. *Cancer Lett* 2018, 431:1-10.
336. Bassey-Archibong BI, Kwiecien JM, Milosavljevic SB, Hallett RM, Rayner LG, Erb MJ, Crawford-Brown CJ, Stephenson KB, Bedard PA, Hassell JA *et al*: Kaiso depletion attenuates transforming growth factor-beta signaling and metastatic activity of triple-negative breast cancer cells. *Oncogenesis* 2016, 5:e208.
337. Wang H, Liu W, Black S, Turner O, Daniel JM, Dean-Colomb W, He QP, Davis M, Yates C: Kaiso, a transcriptional repressor, promotes cell migration and invasion of prostate cancer cells through regulation of miR-31 expression. *Oncotarget* 2016, 7(5):5677-5689.
338. Wang L, Ma J, Wang X, Peng F, Chen X, Zheng B, Wang C, Dai Z, Ai J, Zhao S: Kaiso (ZBTB33) Downregulation by Mirna-181a Inhibits Cell Proliferation, Invasion, and the Epithelial-Mesenchymal Transition in Glioma Cells. *Cellular physiology and biochemistry : international journal of experimental cellular physiology, biochemistry, and pharmacology* 2018, 48(3):947-958.
339. Reinberg D, Sims RJ, 3rd: de FACTo nucleosome dynamics. *J Biol Chem* 2006, 281(33):23297-23301.
340. Winkler DD, Luger K: The histone chaperone FACT: structural insights and mechanisms for nucleosome reorganization. *J Biol Chem* 2011, 286(21):18369-18374.
341. Chaligne R, Popova T, Mendoza-Parra MA, Saleem MA, Gentien D, Ban K, Piolot T, Leroy O, Mariani O, Gronemeyer H *et al*: The inactive X chromosome is epigenetically unstable and transcriptionally labile in breast cancer. *Genome Res* 2015, 25(4):488-503.
342. Carone DM, Lawrence JB: Heterochromatin instability in cancer: from the Barr body to satellites and the nuclear periphery. *Semin Cancer Biol* 2013, 23(2):99-108.
343. Deng X, Berletch JB, Nguyen DK, Disteche CM: X chromosome regulation: diverse patterns in development, tissues and disease. *Nat Rev Genet* 2014, 15(6):367-378.
344. Kang J, Lee HJ, Kim J, Lee JJ, Maeng LS: Dysregulation of X chromosome inactivation in high grade ovarian serous adenocarcinoma. *PLoS One* 2015, 10(3):e0118927.
345. Pageau GJ, Hall LL, Ganesan S, Livingston DM, Lawrence JB: The disappearing Barr body in breast and ovarian cancers. *Nat Rev Cancer* 2007, 7(8):628-633.
346. Dunford A, Weinstock DM, Savova V, Schumacher SE, Cleary JP, Yoda A, Sullivan TJ, Hess JM, Gimelbrant AA, Beroukhim R *et al*: Tumor-suppressor genes that escape from X-inactivation contribute to cancer sex bias. *Nat Genet* 2017, 49(1):10-16.
347. Sirchia SM, Ramoscelli L, Grati FR, Barbera F, Coradini D, Rossella F, Porta G, Lesma E, Ruggeri A, Radice P *et al*: Loss of the inactive X chromosome and replication of the active X in BRCA1-defective and wild-type breast cancer cells. *Cancer Res* 2005, 65(6):2139-2146.
348. Sharp AJ, Stathaki E, Migliavacca E, Brahmachary M, Montgomery SB, Dupre Y, Antonarakis SE: DNA methylation profiles of human active and inactive X chromosomes. *Genome Res* 2011, 21(10):1592-1600.
349. Borensztein M, Okamoto I, Syx L, Guilbaud G, Picard C, Ancelin K, Galupa R, Diabangouaya P, Servant N, Barillot E *et al*: Contribution of epigenetic landscapes and transcription factors to X-chromosome reactivation in the inner cell mass. *Nat Commun* 2017, 8(1):1297.
350. Gendrel AV, Heard E: Noncoding RNAs and epigenetic mechanisms during X-chromosome inactivation. *Annu Rev Cell Dev Biol* 2014, 30:561-580.
351. Giorgetti L, Lajoie BR, Carter AC, Attia M, Zhan Y, Xu J, Chen CJ, Kaplan N, Chang HY, Heard E *et al*: Structural organization of the inactive X chromosome in the mouse. *Nature* 2016, 535(7613):575-579.
352. Brekken RA, Sage EH: SPARC, a matricellular protein: at the crossroads of cell-matrix communication. *Matrix biology : journal of the International Society for Matrix Biology* 2001, 19(8):816-827.

353. Ledda F, Bravo AI, Adris S, Bover L, Mordoh J, Podhajcer OL: The expression of the secreted protein acidic and rich in cysteine (SPARC) is associated with the neoplastic progression of human melanoma. *The Journal of investigative dermatology* 1997, 108(2):210-214.
354. Li FZ, Dhillon AS, Anderson RL, McArthur G, Ferrao PT: Phenotype switching in melanoma: implications for progression and therapy. *Front Oncol* 2015, 5:31.
355. Noguchi K, Dalton AC, Howley BV, McCall BJ, Yoshida A, Diehl JA, Howe PH: Interleukin-like EMT inducer regulates partial phenotype switching in MITF-low melanoma cell lines. *PLoS One* 2017, 12(5):e0177830.
356. Gopal SK, Greening DW, Mathias RA, Ji H, Rai A, Chen M, Zhu HJ, Simpson RJ: YBX1/YB-1 induces partial EMT and tumourigenicity through secretion of angiogenic factors into the extracellular microenvironment. *Oncotarget* 2015, 6(15):13718-13730.
357. Grande MT, Sanchez-Laorden B, Lopez-Blau C, De Frutos CA, Boutet A, Arevalo M, Rowe RG, Weiss SJ, Lopez-Novoa JM, Nieto MA: Snail1-induced partial epithelial-to-mesenchymal transition drives renal fibrosis in mice and can be targeted to reverse established disease. *Nature medicine* 2015, 21(9):989-997.
358. Chen YT, Scanlan MJ, Venditti CA, Chua R, Theiler G, Stevenson BJ, Iseli C, Gure AO, Vasicek T, Strausberg RL *et al*: Identification of cancer/testis-antigen genes by massively parallel signature sequencing. *Proc Natl Acad Sci U S A* 2005, 102(22):7940-7945.
359. Shang B, Gao A, Pan Y, Zhang G, Tu J, Zhou Y, Yang P, Cao Z, Wei Q, Ding Y *et al*: CT45A1 acts as a new proto-oncogene to trigger tumorigenesis and cancer metastasis. *Cell death & disease* 2014, 5:e1285.
360. Koop A, Sellami N, Adam-Klages S, Lettau M, Kabelitz D, Janssen O, Heidebrecht HJ: Down-regulation of the cancer/testis antigen 45 (CT45) is associated with altered tumor cell morphology, adhesion and migration. *Cell communication and signaling : CCS* 2013, 11(1):41.
361. England CG, Luo H, Cai W: HaloTag technology: a versatile platform for biomedical applications. *Bioconjugate chemistry* 2015, 26(6):975-986.
362. Fiorito E, Sharma Y, Gilfillan S, Wang S, Singh SK, Satheesh SV, Katika MR, Urbanucci A, Thiede B, Mills IG *et al*: CTCF modulates Estrogen Receptor function through specific chromatin and nuclear matrix interactions. *Nucleic Acids Res* 2016, 44(22):10588-10602.
363. Yusufzai TM, Tagami H, Nakatani Y, Felsenfeld G: CTCF tethers an insulator to subnuclear sites, suggesting shared insulator mechanisms across species. *Mol Cell* 2004, 13(2):291-298.
364. Pena-Hernandez R: Genome wide targeting of the epigenetic regulatory protein CTCF to gene promoters by the transcription factor TFll-l. 2015.
365. Andersen JS, Lam YW, Leung AK, Ong SE, Lyon CE, Lamond AI, Mann M: Nucleolar proteome dynamics. *Nature* 2005, 433(7021):77-83.
366. Boisvert FM, van Koningsbruggen S, Navascues J, Lamond AI: The multifunctional nucleolus. *Nat Rev Mol Cell Biol* 2007, 8(7):574-585.
367. Torrano V, Navascues J, Docquier F, Zhang R, Burke LJ, Chernukhin I, Farrar D, Leon J, Berciano MT, Renkawitz R *et al*: Targeting of CTCF to the nucleolus inhibits nucleolar transcription through a poly(ADP-ribosyl)ation-dependent mechanism. *J Cell Sci* 2006, 119(Pt 9):1746-1759.
368. Hernandez-Hernandez A, Soto-Reyes E, Ortiz R, Arriaga-Canon C, Echeverria-Martinez OM, Vazquez-Nin GH, Recillas-Targa F: Changes of the nucleolus architecture in absence of the nuclear factor CTCF. *Cytogenetic and genome research* 2012, 136(2):89-96.
369. Lambert SA, Jolma A, Campitelli LF, Das PK, Yin Y, Albu M, Chen X, Taipale J, Hughes TR, Weirauch MT: The Human Transcription Factors. *Cell* 2018, 175(2):598-599.
370. Allen HF, Wade PA, Kutateladze TG: The NuRD architecture. *Cell Mol Life Sci* 2013, 70(19):3513-3524.
371. Carter JH, Deddens JA, Spaulding NI, Lucas D, Colligan BM, Lewis TG, Hawkins E, Jones J, Pemberton JO, Douglass LE *et al*: Phosphorylation of eIF4E serine 209 is associated with tumour progression and reduced survival in malignant melanoma. *Br J Cancer* 2016, 114(4):444-453.
372. Robichaud N, del Rincon SV, Huor B, Alain T, Petruccelli LA, Hearnden J, Goncalves C, Grotegut S, Spruck CH, Furic L *et al*: Phosphorylation of eIF4E promotes EMT and metastasis via translational control of SNAIL and MMP-3. *Oncogene* 2015, 34(16):2032-2042.
373. Kardos GR, Robertson GP: Therapeutic interventions to disrupt the protein synthetic machinery in melanoma. *Pigment Cell Melanoma Res* 2015, 28(5):501-519.
374. Tsompana M, Buck MJ: Chromatin accessibility: a window into the genome. *Epigenetics Chromatin* 2014, 7(1):33.

375. Corces MR, Trevino AE, Hamilton EG, Greenside PG, Sinnott-Armstrong NA, Vesuna S, Satpathy AT, Rubin AJ, Montine KS, Wu B *et al*: An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat Methods* 2017, 14(10):959-962.
376. Montefiori L, Hernandez L, Zhang Z, Gilad Y, Ober C, Crawford G, Nobrega M, Jo Sakabe N: Reducing mitochondrial reads in ATAC-seq using CRISPR/Cas9. *Sci Rep* 2017, 7(1):2451.
377. Pranzatelli TJF, Michael DG, Chiorini JA: ATAC2GRN: optimized ATAC-seq and DNase1-seq pipelines for rapid and accurate genome regulatory network inference. *BMC Genomics* 2018, 19(1):563.
378. Wei Z, Zhang W, Fang H, Li Y, Wang X: esATAC: an easy-to-use systematic pipeline for ATAC-seq data analysis. *Bioinformatics* 2018, 34(15):2664-2665.
379. Fiorentino FP, Macaluso M, Miranda F, Montanari M, Russo A, Bagella L, Giordano A: CTCF and BORIS regulate Rb2/p130 gene transcription: a novel mechanism and a new paradigm for understanding the biology of lung cancer. *Mol Cancer Res* 2011, 9(2):225-233.