

# Determining the effects of DOT1L interacting proteins on cellular reprogramming

by

**Deniz Uğurlu Çimen**

A Dissertation Submitted to the  
Graduate School of Sciences and Engineering  
in Partial Fulfillment of the Requirements for  
the Degree of

**Doctor of Philosophy**

in

Molecular Biology and Genetics



**KOÇ  
UNIVERSITY**

August 31, 2018

Determining the effects of DOT1L interacting proteins  
on cellular reprogramming

Koç University

Graduate School of Sciences and Engineering

This is to certify that I have examined this copy of a doctoral dissertation by

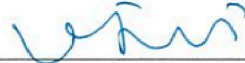
Deniz Uğurlu Çimen

and have found that it is complete and satisfactory in all respects,  
and that any and all revisions required by the final examining committee have  
been made.

Committee Members:



Associate Professor Dr. Tamer Önder (Advisor)



Associate Professor Dr. Nurhan Özlü



Associate Professor Dr. Bilal Kerman



Associate Professor Dr. Nathan Lack



Associate Professor Dr. Tolga Emre

Date: 31.08.2018

# ABSTRACT

Fully differentiated cells can be reprogrammed into induced pluripotent stem cells (iPSCs) by ectopic expression of transcription factors. However, the mechanism behind this cell fate change is not fully elucidated. Epigenetic regulators have important roles during embryonic development as well as somatic cell reprogramming. Previously, it has been shown that inhibition of DOT1L, the histone H3 lysine 79 methyltransferase, increases the efficiency of reprogramming via regulation of lineage specific genes. DOT1L is recruited to chromatin and act in concert with a number of additional chromatin regulators. However, the role of such DOT1L-interacting proteins in reprogramming remains unknown.

In the first part of this thesis, novel DOT1L interactors were identified using the BioID method in which a promiscuous BirA ligase (BirA\*) was employed to biotinylate DOT1L-proximal proteins, *in vivo*. Biotinylated proteins were pulled-down by Streptavidin and identity of the proteins was determined by LC-MS/MS. The resulting novel interaction candidates were investigated for their effects on reprogramming. Candidate genes were knocked-down in human fibroblasts via shRNAs followed by reprogramming. Our results indicated that knock-down of AF10 (*MLLT10*), significantly increased the iPSC generation efficiency, suggesting that it acts as a barrier to reprogramming similar to DOT1L. This finding was verified by CRISPR/Cas9 mediated knockout of AF10. Combining DOT1L inhibition or knockout with AF10 suppression did not result in an additive enhancement of reprogramming, suggesting that these two chromatin factors act in the same pathway.

In the second part of this thesis, known direct and functional interactors of DOT1L were curated from the literature and their effects on reprogramming was investigated through loss of function experiments. Suppression of *Mixed Lineage Leukemia 1* (*MLL1*) expression via RNA interference or CRISPR/Cas9 significantly increased reprogramming efficiency. To determine how MLL1 prevents reprogramming, RNA-sequencing was performed. MLL1 suppression resulted in downregulation of fibroblast-specific genes and accelerated the activation of pluripotency-related genes.

Taken together, this study uncovered two important chromatin factors that act as barriers to reprogramming and contributed to our understanding of epigenetic mechanisms that maintain cell identity.

Key words: Reprogramming, DOT1L, BioID, AF10, MLL1

# ÖZETÇE

Tamamen farklılaşmış hücreler ektopik transkripsiyon faktörleri ile uyarılmış plüripotent kök hücrelere (uPKH) yeniden programlanabilmektedir. Fakat bu hücre kaderi değişiminin arkasında yatan mekanizma hala tam olarak açıklanamamıştır. Embriyonik gelişim sürecinde olduğu gibi somatik hücre yeniden programlanmasında da epigenetik düzenleyicilerin büyük rolü bulunmaktadır. Bundan önce, Histon H3 lizin 79 metiltransferazı olan DOT1L'in susturulması durumunun, soya özgü genleri düzenleyerek yeniden programlama verimini arttırdığı gösterilmiştir. DOT1L, kromatine çağırılarak bir dizi diğer kromatin düzenleyiciler ile uyum halinde hareket etmektedir. Fakat bu tip DOT1L-etkileşimli proteinlerin yeniden programlama üzerindeki rolü halen bilinmemektedir.

Bu tezin ilk kısmında, DOT1L-yakınsal proteinleri biyotinlemekle görevli, seçici olmayan BirA ligazı (BirA\*) içeren BioID metodu kullanılarak; özgün DOT1L-etkileşimli proteinler in vivo olarak belirlenmiştir. Biyotinlenen proteinler Streptavidin ile çöktürülüp LC-MS/MS ile tanılanmıştır. Elde edilen özgün DOT1L-etkileşimli protein adaylarının yeniden programlama üzerindeki etkileri incelenmiştir. İnsan fibroblast hücrelerinde shRNA'lar ile aday genlerin ifadeleri baskılanıp bu hücreler yeniden programlanmıştır. Elde ettiğimiz sonuçlar, AF10'un (*MLLT10*) baskılanmasının uPKH oluşum verimini arttırdığını ve yeniden programlama üzerinde DOT1L'e benzer şekilde bariyer görevi olduğunu göstermektedir. Bu bulgu, CRISPR/Cas9 aracılığıyla AF10'un susturulması ile de doğrulanmıştır. DOT1L-baskılanması ve AF10-susturulması durumlarının bir arada uygulanmasıyla, ek bir artışın

görülmemesi, bu iki kromatin faktörünün aynı yolakta etkili olduğunu akla getirmektedir.

Bu tezin ikinci kısmında, bilinen DOT1L'e doğrudan ve işlevsel bağlanan proteinler literatürden belirlenerek, bu genlerin yeniden programlama üzerindeki etkileri işlevsel kayıp deneyleri ile incelenmiştir. RNA interferans veya CRISPR/Cas9 ile Mixed Lineage Leukemia (*MLL1*) ifadesinin baskılanması yeniden programlama verimini arttırmıştır. *MLL1*'in yeniden programlamayı nasıl engellediğini belirlemek için, RNA-dizileme analizi yapılmıştır. *MLL1*'in baskılanması fibroblastlara-özü genlerin baskılanması ve plüripotentlere-özü genlerin etkinleşme hızında ivmelenme ile sonuçlanmıştır.

Bir arada ele alındığında, bu çalışma iki önemli kromatin belirleyicisinin yeniden programlama üzerinde bariyer olarak davrandığını ortaya çıkarmış ve hücre kimliğinin devamlılığını sağlayan epigenetik mekanizmaları anlamamıza katkı sağlamıştır.

Anahtar kelimeler: Yeniden programlama, DOT1L, BioID, AF10, *MLL1*

## ACKNOWLEDGEMENT

I would like to start my thanks with my PhD advisor, Dr. Tamer Önder for his guidance throughout my PhD training. By sharing his experiences and ideas, he had made the greatest contribution to my thesis. Our scientific discussions were invaluable for shaping the scientist in me. Being in his lab was a great opportunity since he created distinguished scientific environment in Turkey. I feel very lucky for listening his talk in Bilkent which turned out to finding my dream job. I am thankful to him for every support, patience and enthusiasm.

I want to thank my thesis committee members, Dr. Nurhan Özlü for her contribution to this thesis and her help with LC-MS/MS and Dr. Bilal Kerman for his scientific discussions on my project. I also want to thank my thesis defense committee members, Dr. Nathan Lack and Dr. Tolga Emre for their questions and suggestions on this thesis.

I would like to thank to The Scientific and Technological Research Council of Turkey (TÜBİTAK) and Koç University for their financial support throughout my PhD study.

Moreover, I want to express my thanks to past and present members of Önder Lab. I appreciate every bit of help during my PhD study and I want to thank every member for their contribution. My greatest thanks are for Burcu Özçimen, a fellow recently-PhD. Her relentless support was vital and her friendship was one-of-a-kind for me. I thank my former home-mates, Filiz, Hilal and Tuğba for making living-together so fun. I want to thank all

members of Koç University School of Medicine (KUSOM) for promoting a nice workplace. Especially, I thank my friends, Fatma, Fidan, Ezgi, Kerem, Ata, Gülnihal, Ayyub and Eray. I also want to thank my friends who are living in distance, Taner, Merve, Ferhan, Ezgi, Ece and Asya. Especially, Aslmur has an immense support in this thesis.

Finally, I want to thank my family. My parents, Şemsinur and Cemil Uğurlu were supported me with their unconditional love throughout my life. And my brother, Cihan Uğurlu was always there for me whenever I had a problem. My deepest thanks are for my husband, Harun Çimen who has endless support for me during my restless becoming-a-PhD process. Thank you for being with me in this journey and sharing every aspect of life.

# ABBREVIATIONS

AML	acute myeloid leukemia
AP	Antarctic Phosphatase
APEX	ascorbate peroxidase
BD	bromodomain
BioID	Proximity dependent biotin identification
BirA*	promiscuous BirA ligase
BSA	bovine serum albumin
cDNA	complementary DNA
ChIP	chromatin immunoprecipitation
ChIP-seq	chromatin immunoprecipitation followed by sequencing
CHIR	CHIR99021
co-IP	co-immunoprecipitation
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CTD	C-terminal domain
DAB	3,3 diaminobenzidine
DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimethyl sulfoxide
DNA	deoxyribonucleic acid
DNMTs	DNA methyltransferases
Dot	<u>disruptors of telomeric silencing</u>
DOT1L	Dot1-like
DotCom	DOT1L complex
DOX	doxycycline
DSB	double strand brake
DTT	Dithiothreitol
EAPs	ENL-associated proteins
ECT spe.	ectodermal specifiers
EDTA	Ethylenediaminetetraacetic acid
EpiSCs	Epiblast Stem Cells
EPZ	EPZ004777 (DOT1L inhibitor)
ESCs	embryonic stem cells
FDR	false discovery rate
Frskln	Forskolin



FYR	phenylalanine-tyrosine rich
GFP	Green Fluorescent Protein
GSEA	gene set enrichment analysis
HDAC	histone deacetylase
HEK293T	human embryonic kidney 293T cells
HMG	high-mobility group
HRP	horseradish peroxidase
IPL	in vivo proximal labeling
IR	ionizing radiation
iDOT1L	inhibitor of DOT1L
iPSCs	induced pluripotent stem cells
KD	knock-down
kDa	kilo Dalton
KO	knock-out
KOSR	knock-out serum replacement
LBD	LEDGF binding domain
LC-MS/MS	Liquid chromatography-tandem mass spectrometry
LTR	Long terminal repeat
MBM	MENIN binding motif
ME spe.	Mesendodermal specifiers
MEF	mouse embryonic fibroblasts
MET	mesenchymal to epithelial transition
<i>MLL1</i>	<i>Mixed Lineage Leukemia 1</i>
mut	mutant
nd	not detected
neaa	Non-essential amino acids
NEB	new viingland biolabs
NES	normalized enrichment score
n.s.	not significant
NT	non-targeting
OE	overexpression
OM-LZ	Octamer motif-leucine zipper
OSKM	Oct4, Sox2, Klf4 and c-Myc
PB	PiggyBac
PBS	Phosphate-buffered saline
PFA	paraformaldehyde
PHD	plant homeodomain
PPIs	protein-protein interactions
PSM	peptide spectrum matches

PZP domain	PHD1-Zn-Knuckle-PHD2 domain
qPCR	quantitative real-time PCR
RNA	ribonucleic acid
RNAPII	RNA Polymerase II
RT	room temperature
s.d.	standard deviation
SAM	S-adenosyl-L-methionine
SCNT	somatic cell nuclear transfer
SEC	superelongation complex
SET domain	<u>S</u> uvar3-9, <u>E</u> nhancer of Zeste, <u>T</u> rithorax
sgRNA/gRNA	single guide RNA/guide RNA
shRNA	short hairpin RNA
Sir	silent information regulator
TAD	transactivation domain
Tran.	Tranylcypromine
UV	ultraviolet radiation
Vit. C	vitamin C
VPA	valproic acid
Win	WDR5 interaction motif
wt	wild-type
XEN-like cells	extraembryonic endoderm-like cells

# TABLE OF CONTENTS

ABSTRACT .....	i
ÖZETÇE.....	ii
ACKNOWLEDGEMENT .....	iv
ABBREVIATIONS.....	vi
TABLE OF FIGURES.....	xiv
TABLE OF TABLES .....	xviii
1 Introduction.....	1
1.1 Somatic cells can be reprogrammed to pluripotent stem cells.....	1
1.1.1 Induced pluripotent stem cells vs. embryonic stem cells.....	2
1.1.2 Derivation of iPSCs.....	4
1.1.3 Stages of iPSC generation .....	10
1.1.4 Transcriptional changes and chromatin reorganization during reprogramming.....	11
1.2 Epigenetic modifiers as barriers and enhancers of reprogramming...	16
1.3 Dot1L (H3K79methyltransferase) acts as a barrier to reprogramming	19
1.3.1 Discovery of Dot1 protein .....	20
1.3.2 Structure of DOT1L protein.....	20
1.3.3 DOT1L's diverse roles in the cell .....	21
1.3.4 Known protein interactions of DOT1L.....	27

1.4	Hypothesis .....	29
2	Materials and Methods .....	30
2.1	Plasmids and cloning procedures .....	30
2.1.1	BirA*-DOT1L fusion protein expression plasmid cloning.....	31
2.1.2	shRNA cloning into pSMP vector .....	34
2.1.3	gRNA cloning into lentiCRISPRv2 vector.....	37
2.1.4	T7-endonuclease assay.....	38
2.1.5	AF10 overexpression plasmids cloning.....	39
2.1.6	Cloning of GFP Plasmids .....	41
2.1.7	Sequencing of samples .....	41
2.2	Cell Culture .....	43
2.2.1	Cell dissociation, freezing and thawing procedures.....	43
2.2.2	Generation of Mitomycin-c treated MEFs.....	44
2.2.3	Generation of iPSCs from dH1fs (Reprogramming assays).....	45
2.2.4	Tra-1-60 staining of iPSC colonies .....	46
2.2.5	Generation of DOT1L-KO single cell clone .....	47
2.2.6	Transient Transfection .....	48
2.2.7	Virus production .....	49
2.2.8	Transduction of cells .....	51
2.3	Western Blots .....	51
2.3.1	Whole cell lysis method.....	52
2.3.2	Nuclear protein extraction.....	52

2.3.3	Histone acid extraction.....	53
2.3.4	Western Blotting.....	54
2.4	Pull down experiments.....	56
2.5	Mass-Spectrometry Analysis .....	56
2.6	RNA isolation, cDNA synthesis and qPCR .....	58
2.7	Microarray analysis.....	62
2.8	RNA sequencing and analysis .....	62
3	Finding proximal-protein interactions of DOT1L via BioID method and shRNA-mediated screen of DOT1L-proximity interactors for reprogramming efficiency .....	64
3.1	Introduction.....	64
3.1.1	Biotinylation based detection of protein-protein interaction methods .....	64
3.1.2	BioID is a powerful method to detect proximal protein interactions .....	66
3.1.3	AF10 is an important member of DOT1L-containing elongation complex.....	68
3.1.4	AF10 regulates histone code.....	70
3.2	Results .....	73
3.2.1	BirA*-DOT1L fusion protein successfully methylates H3K79 residue 73	
3.2.2	Pull-down of biotinylated proteins via Streptavidin beads .....	78

3.2.3	Proximal-protein interactions of DOT1L was identified after LC-MS/MS analysis of biotinylated proteins .....	80
3.2.4	shRNA-mediated knock-down of DOT1L-proximal proteins .....	84
3.2.5	shRNA-mediated screen of DOT1L proximal interactors for reprogramming efficiency .....	85
3.2.6	Knock-out of AF10 via CRISPR increases reprogramming efficiency similar to knock-down of AF10 via shRNA .....	87
3.2.7	sgAF10s decreases H3K79 methylation .....	90
3.2.8	sgAF10's effect on reprogramming can be reversed via overexpression of AF10 .....	91
3.2.9	Loss-of AF10 together with inhibition of DOT1L does not have an additive effect on reprogramming efficiency .....	93
3.2.10	Double-KO of AF10 and DOT1L during reprogramming .....	95
4	MLL1 silencing increases reprogramming efficiency .....	99
4.1	Introduction .....	99
4.1.1	MLL1 translocations in Acute Myeloid Leukemia (AML) patients 100	
4.1.2	MLL complex in transcription .....	101
4.1.3	The role of MLL1 in pluripotency .....	103
4.2	Results .....	105
4.2.1	shRNA-mediated screen of DOT1L's previously known interactors for reprogramming efficiency .....	105

4.2.2	MLL1 and DOT1L has independent roles during reprogramming	107
4.2.3	shRNA-mediated screen of MLL1 complex proteins for reprogramming efficiency .....	112
4.2.4	gMLL1s change gene expression to accelerate fibroblast to iPSC transition .....	115
5	Discussion .....	124
5.1	Proximal interactors of DOT1L was identified by BioID method..	125
5.2	AF10 blocks reprogramming through regulation of H3K79 methylation via DOT1L .....	128
5.3	MLL1 blocks reprogramming through regulation of gene expression independent from DOT1L .....	130
5.4	Final conclusion .....	133
6	Appendix-I.....	135
7	Appendix-II.....	142
8	Appendix-III .....	168
9	References.....	169

## TABLE OF FIGURES

Figure 1. Different combinations of transgenes and chemicals for reprogramming.....	6
Figure 2. Interaction partners of DOT1L during transcriptional elongation .	23
Figure 3. Schematic representation of the DOT1L protein .....	29
Figure 4. Cloning steps of BirA*-DOT1L fusion protein expressing plasmid	33
Figure 5. Sketch of shRNA cloned pSMP plasmid .....	34
Figure 6. Sketch of gRNA cloned lentiCRISPRv2 plasmid .....	37
Figure 7. Timeline of reprogramming experiment .....	46
Figure 8. Promiscuous biotin-ligase (BirA*) selectively biotinylates the proximal proteins .....	67
Figure 9. Schematic representation of the AF10 protein.....	68
Figure 10. Representative structure of the DOT1L–AF10 interaction (Adapted from Zhang <i>et.al.</i> with permission presented in Appendix-III <sup>147</sup> ) ..	70
Figure 11. Schematic of BirA*-DOT1L fusion proteins in pLEX_307 vectors .....	74
Figure 12. Time-line for the generation of DOT1L-KO HEK293T cells .....	75
Figure 13. H3K79me2 levels in DOT1L gRNA transfected HEK293T single cell clones.....	76
Figure 14. Rescue of H3K79 methylation by BirA*-DOT1L_wt fusion protein in DOT1L-KO cells.....	77
Figure 15. Biotinylated proteins in BirA*-DOT1L expressing HEK293T cells in comparison to uninfected cells. ....	79



Figure 16. Flowchart for identifying proteins specifically biotinylated by BirA*-DOT1L in LC-MS/MS analysis.....	81
Figure 17. Novel proximal interactions of DOT1L were revealed via BioID along with known direct interactions of DOT1L.....	83
Figure 18. mRNA expression levels of shRNA-targeted genes.....	85
Figure 19. Timeline of reprogramming experiments for shRNA infected fibroblasts .....	85
Figure 20. Fold change in reprogramming efficiency as a result of shRNA-mediated gene silencing.....	86
Figure 21. CRISPR/CAS9 based suppression of AF10 and its effect on reprogramming.....	89
Figure 22. gRNA mediated AF10 knock-out decreases H3K79me2 levels of dH1fs.....	90
Figure 23. AF10 overexpression rescues sgAF10 phenotypes .....	92
Figure 24. Knock-down or Knock-out of AF10 does not further increase reprogramming efficiency in DOT1L-inhibited condition .....	94
Figure 25. H3K79me2 levels in gRNA-mediated AF10 & DOT1L knock-out dH1fs.....	96
Figure 26. gRNA mediated AF10 & DOT1L double knock-out does not have additive effect on reprogramming.....	98
Figure 27. Multiple functional domains of MLL1 is depicted before and after its cleavage by Taspase1 .....	102
Figure 28. mRNA expression levels of shRNA-target genes .....	106
Figure 29. Reprogramming efficiency change as a result of shRNA-mediated gene silencing of DOT1L's previously known interactors.....	107

Figure 30. Knock-down of MLL1 in DOT1L-inhibited condition further increases reprogramming efficiency .....	109
Figure 31. Verification of MLL1 silencing experiments with knock-out MLL1 experiments.....	111
Figure 32. mRNA expression levels of shRNA-targeted genes belonging to MLL1 complexes.....	113
Figure 33. The effect of knocking down MLL1 complex members on reprogramming efficiency .....	114
Figure 34. Validation of reprogramming phenotype of cells used for RNA-sequencing analysis .....	116
Figure 35. Commonly upregulated and downregulated gene numbers among gMLL1-2 and gMLL1-3 samples at Day 0 and Day 6 samples.....	117
Figure 36. MLL1-KO accelerates the decrease in fibroblast gene sets and promotes the increase in pluripotency gene expression .....	119
Figure 37. Commonly upregulated and downregulated gene numbers among gMLL1 and EPZ004777 samples at Day 0.....	121
Figure 38. RNA-sequencing validation with q-RT-PCR analysis to compare SERPINB9, FBLN5, PRRX1 and FOXF2 gene levels.....	122
Figure 39. MLL1 knock-down does not change global H3K4 mono- and trimethylation in dH1fs.....	132
Figure 40. Fold change in reprogramming efficiency as a result of combination of shRNA-mediated silencing of ENL, AF10 and AF17 .....	135
Figure 41. mRNA expression levels of AF10, AF17, DOT1L and DOT1L-target genes.....	138
Figure 42. Commonly upregulated and downregulated gene numbers among gMLL1-2 and gMLL1-3 samples at Day 0 and Day 6 samples.....	139

Figure 43. Commonly upregulated and downregulated gene numbers among gMLL1 and EPZ004777 samples at Day 0..... 140

Figure 44. Commonly upregulated and downregulated gene numbers among gMLL1-2&3 and EPZ004777 samples at Day 0..... 141



## TABLE OF TABLES

Table 1. Diverse delivery methods for iPSC generation .....	9
Table 2. Epigenetic regulators of somatic cell reprogramming and their effects .....	18
Table 3. List of cloning primer sequences.....	30
Table 4. List of shRNA oligo sequences .....	35
Table 5. List of gRNA targeting sequences .....	38
Table 6. List of T7 specific Primers .....	39
Table 7. PCR cycling conditions of AF10 cloning experiment .....	40
Table 8. List of sequencing primers.....	42
Table 9. Table of viral transfection information for different scales.....	50
Table 10. Recipes of protein lysis buffers.....	52
Table 11. Table of antibody information that were used for western blots...	55
Table 12. Thermal cycling conditions for Q-PCR.....	59
Table 13. List of qRT-PCR Primers .....	60
Table 14. DOT1L detection values after mass spectrometry analysis from two reads of each sample (nd, not detected).....	80
Table 15. Raw data of BioID experiment: HEK293T control cells, 1 <sup>st</sup> reading .....	142
Table 16. Raw data of BioID experiment: HEK293T control cells, 2 <sup>nd</sup> reading .....	145
Table 17. Raw data of BioID experiment: BioID-DOT1L-wt infected HEK293T cells, 1 <sup>st</sup> reading .....	149

Table 18. Raw data of BioID experiment: BioID-DOT1L-wt infected HEK293T cells, 2 <sup>nd</sup> reading.....	152
Table 19. Raw data of BioID experiment: BioID-DOT1L-mut infected HEK293T cells, 1 <sup>st</sup> reading .....	156
Table 20. Raw data of BioID experiment: BioID-DOT1L-mut infected HEK293T cells, 2 <sup>nd</sup> reading.....	161



## Chapter 1- Introduction

### 1.1 Somatic cells can be reprogrammed to pluripotent stem cells

Cellular reprogramming can be defined as changing the fate of a differentiated cell into a pluripotent state. There have been three main advances that led to the discovery of somatic cell reprogramming. The first of these advances was the development of somatic cell nuclear transfer (SCNT) method as it proved that somatic cell nuclei have the same genetic material as their progenitor embryonic counterparts<sup>1</sup>. The second advance was the ability to culture and maintain pluripotent cells under in vitro conditions, beginning first with the isolation of teratocarcinoma cells<sup>2</sup> and then followed by the isolation and successful propagation of embryonic stem cells (ESCs) from early embryos<sup>3</sup>. Lastly, the appreciation of the importance of transcription factors during lineage commitment has led to the possibility of changing cell fates by ectopic expression of these factors<sup>4</sup>. All these advances culminated in the question of whether certain transcription factors can change the fate of fully differentiated cells into pluripotent state.

In 2006, Takahashi and Yamanaka demonstrated that expression of four specific transcription factors (Oct4, Sox2, Klf4 and c-Myc; OSKM in short) triggers reprogramming of mouse somatic cells into embryonic stem cell-like cells, which are called induced pluripotent stem cells (iPSCs)<sup>5</sup>. Following iPSC generation from mouse somatic cells, the same group also demonstrated

that human somatic cells can be successfully reprogrammed into pluripotent state with the same four factors which are now called Yamanaka factors<sup>6</sup>.

### **1.1.1 Induced pluripotent stem cells vs. embryonic stem cells**

ESCs are isolated from the inner cell mass of blastocysts. They are capable of differentiating into cells derived from all three germ layers; therefore ESCs are called as “pluripotent” cells. Inner cell mass cells are naturally existing pluripotent cells whereas iPSCs are generated *in vitro* from differentiated cells with ectopic expression of Yamanaka factors (OSKM).

In mice, pluripotent stem cells can be isolated from 2 different stages of embryonic development: ESCs are isolated from preimplantation inner cell mass whereas Epiblast Stem Cells (EpiSCs) are isolated from epiblast of post implantation embryo. Even though these cells are both called as pluripotent cells, they have slightly different capacities in terms of their potency. While ESCs are called as “naïve”, EpiSCs are called as “primed” because they are more committed cells than ESCs. Morphologically, EpiSCs are more flattened in contrast to dome shaped mouse ESCs. Developmentally, EpiSCs are inefficient in generating chimeras while ESCs can contribute to 3 germ-layers by forming a chimera when injected into a mouse blastocyst. Epigenetically, female ESCs have both active X chromosomes (XaXa) while EpiSCs show X chromosome inactivation (XaXi)<sup>7</sup>.

Conventional mouse iPSCs resemble EpiSCs but they can be easily converted to naïve state with certain growth conditions, however, it is harder to have

human iPSCs in the naïve state<sup>8</sup>. In addition, human ESCs resemble mouse EpiSCs even though human ESCs are also isolated from preimplanted embryos<sup>9</sup>. Still, with specific conditions, naïve state human ESCs can be isolated<sup>10</sup>. There are attempts of making naïve human iPSCs<sup>11,12</sup> but protocols remain experimentally challenging.

In definition, iPSCs are pluripotent similar to ESCs, however there are diverse studies that investigate whether iPSCs are indeed similar to ESCs. First reports demonstrated that iPSCs are different from ESCs in terms of their DNA methylation and gene expression pattern even though both cell lines have similar pluripotency properties<sup>13-15</sup>. However, these differences were abrogated as iPSCs maintained in the culture for longer period of time since fully reprogrammed iPSCs will survive while intermediate stage cells will be eliminated<sup>15,16</sup>. Moreover, human iPSCs resulted in heterogeneous status for X chromosome re-activation, whereas ESCs have both active X chromosomes. One study reported that using SNL feeder cells and iPSC derivation with retroviruses promote the re-activation of some of the X-linked genes<sup>17</sup>. This shows that iPSC derivation and their culture conditions have a role in epigenetic memory erasure as much as passaging period of iPSCs. In contrast to these comparison studies, there is lack of consensus in the ESCs epigenetic pattern. For example, two studies<sup>18,19</sup> that were published at the same time, demonstrated different DNA methylation pattern for ESCs and blastocysts<sup>20</sup>. Hence, even ESCs that were grown *in vitro* do not exhibit the same DNA methylation status as embryos that were in their natural environment. Despite differential epigenetic features of iPSCs, mouse iPSCs can generate normal mice in tetraploid complementation assay which shows that fully reprogrammed iPSCs are *bona fide* pluripotent<sup>21-23</sup>. However, this kind of



assays is not ethically suitable for human iPSCs; therefore less stringent assays (such as in vitro differentiation or teratoma formation) are performed to show pluripotency of human iPSCs.

### 1.1.2 Derivation of iPSCs

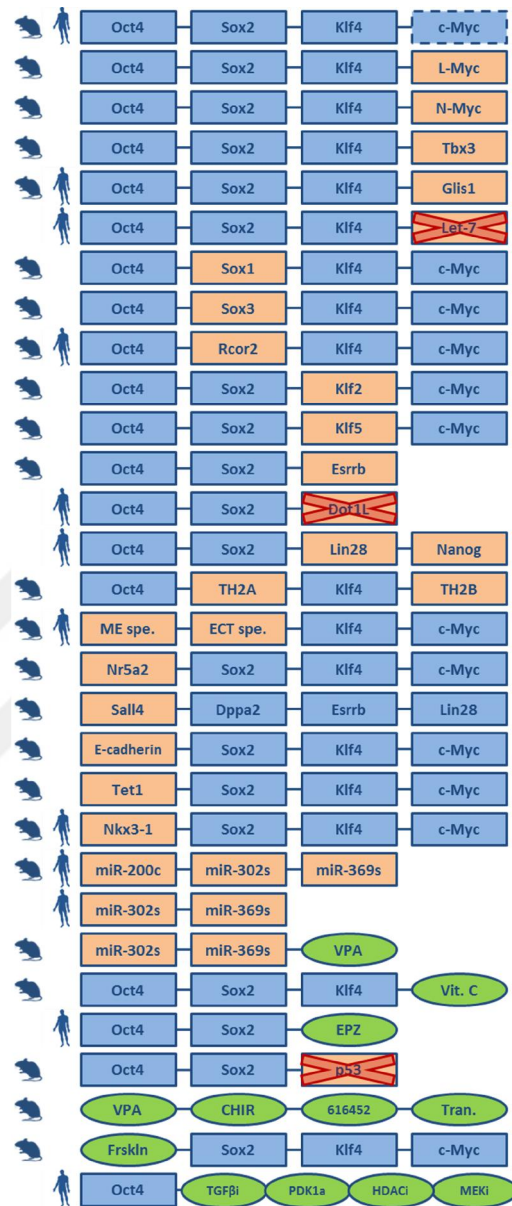
Following mouse and human iPSC derivation, somatic cells of different species were reprogrammed; such as monkey<sup>24</sup>, rat<sup>25</sup>, pig<sup>26</sup>, cattle, horse, sheep, goat, and rabbit<sup>27</sup>. Somatic cells from all these different animals were successfully reprogrammed with OSKM induction which shows the evolutionary conservation of pluripotency. In addition, reprogramming can be achieved from different source cell types such as fibroblasts, mature B cells, hepatocytes and keratinocytes<sup>28</sup>. These findings show that OSKM transcription factors are broadly effective to achieve pluripotency from various starting somatic cells.

However, OSKM is not the only route to generate iPSCs (**Figure 1**). There are several different approaches that can also result in reprogramming of somatic cells. For example, c-Myc can be replaced with L-Myc or N-Myc<sup>29</sup> or with other transcription factors such as Tbx3<sup>30</sup> or Glis1<sup>31</sup>. Besides, somatic cell reprogramming can be achieved without c-Myc albeit at much reduced efficiency<sup>5,32</sup>. Knock down of *let-7* can promote human somatic cell reprogramming as much as c-Myc<sup>33</sup>. Other replacements can be made, too: Sox1, Sox3 and Rcor2<sup>34</sup> can replace Sox2; or Klf2 and Klf5 can replace Klf4<sup>32</sup>. Moreover, orphan nuclear receptor Esrrb can replace both Klf4 and c-Myc<sup>35</sup>. Similarly, silencing Dot1L, the H3K79 methyltransferase, can also replace

Klf4 and c-Myc<sup>36</sup>. Furthermore, Lin28 and Nanog is used as an alternative to Klf4 and c-Myc for reprogramming and this cocktail is called as OSLN<sup>37</sup>. Histone variants TH2A and TH2B can replace Sox2 and c-Myc for mouse fibroblast reprogramming but they also require phosphorylated NPM (P-NPM) which is a histone chaperone<sup>38</sup>.

In addition to these alternatives, another approach to reprogramming was postulated in 2013: reprogramming with lineage specifiers. In this method, mouse cells were reprogrammed by counteracting lineage specifiers that are non-pluripotency related genes and this phenomenon was defined as “seesaw model”<sup>39</sup>. A few months later, Belmonte group replicated the similar kind of reprogramming with human cells<sup>40</sup>. They both showed that Oct4 can be replaced by mesendodermal specifiers and Sox2 can be replaced by ectodermal specifiers.

In mouse reprogramming, Oct4 can be replaced with other factors such as Nr5a2<sup>41</sup>, Sall4<sup>42</sup>, E-cadherin<sup>43</sup> or Tet1<sup>44</sup>. In human cells, Oct4 can be replaced with lineage specifiers as explained above. Recently, it was demonstrated that Oct4 can be replaced with another transcription factor Nkx3-1 in both mouse and human reprogramming<sup>45</sup>. Also, human and mouse cells can be reprogrammed with miRNA clusters without any need of exogenous expression of Yamanaka factors. For example, direct transfection of miR-200c/302s/369s mature miRNAs can reprogram both mouse and human somatic cells<sup>46</sup>. Moreover, miR-302/367 can also reprogram human fibroblasts however, it requires HDAC2 degrading valproic acid (VPA) to reprogram MEFs<sup>47</sup>.



**Figure 1. Different combinations of transgenes and chemicals for reprogramming**

Alternative reprogramming methods were curated in the image below and referred in the text. Yamanaka factors are filled with blue, chemicals are colored as green. Images on left side depict the cell type origin of given combination. (ME spe., mesendodermal specifiers; ECT spe., ectodermal specifiers; VPA, valproic acid; Vit. C, vitamin C; EPZ, EPZ004777; CHIR, CHIR99021; Tran., Tranylcypromine; Frskln, Forskolin) Adapted from Theunissen and Jaenisch, 2014<sup>48</sup>.

There are number of different ways to increase the reprogramming efficiency. Other than changing the reprogramming factors, additions can be made to enhance the efficiency. For instance, using vitamin C can boost the iPSC generation and it can also replace c-Myc<sup>49</sup>. Inhibition of a histone methyltransferase, DOT1L with small molecule EPZ004777 was identified to enhance reprogramming as well as replace Klf4 and c-Myc<sup>36</sup>. In addition, silencing of p53 significantly increases reprogramming and can replace Klf4 and c-Myc in OSKM reprogramming<sup>50</sup>. Inhibition of p53 pathway can increase the survival of cells; however, it is counter argued that reprogramming without p53 may resulted in transformed iPSCs<sup>51</sup>. This shows the complexity of production of pluripotent cells and the increase in the efficiency may decrease the quality of iPSCs. There are still many aspects to be resolved in the molecular mechanisms of reprogramming.

Another trend in reprogramming is replacing transcription factor induction with chemical treatments so that the usage of viral induction can be bypassed. For this goal, Hou *et. al.* proposed a chemical cocktail that can generate iPSCs from mice<sup>52</sup>. In their first report, they used VPA, CHIR99021, 616452 and Tranylcypromine (VC6T) chemical cocktail to reprogram MEFs with Oct4 induction<sup>53</sup>. Then, they identified Forskolin as a substitution of expression of Oct4 and included DZNep to the cocktail (VC6TFZ)<sup>52</sup>. Same group also established that there is an intermediate state of extraembryonic endoderm (XEN)-like cells in the process of chemical reprogramming which is not the case in transcription factor induced reprogramming<sup>54</sup>. In that study, they also demonstrated that two different DOT1L inhibitors (EPZ004777 and SGC0946) promote the chemical reprogramming of MEFs<sup>54</sup>. EPZ004777 was previously revealed to enhance

reprogramming of OSKM-induced human fibroblasts<sup>55</sup>. This shows that DOT1L is a barrier to reprogramming and its inhibition promotes both transcription factor induced reprogramming of human fibroblasts and chemical reprogramming of MEFs. However, chemical reprogramming has not yet been successfully achieved with human somatic cells. So far, it was shown that 3 Yamanaka factors were replaced with small molecules (TGF $\beta$ i, PDK1a, HDACi, MEKi) but the process still requires Oct4 induction<sup>56</sup>. Further investigations are continuing in this field since chemical reprogramming can be very useful for therapeutic usage of iPSCs.

There are different methods to deliver genes into cells for reprogramming (**Table 1**). The first experiment of reprogramming was done with retroviruses<sup>5</sup>. However later, lentiviral delivery was favored since its infection efficiency is higher therefore reprogramming efficiency is higher<sup>37</sup>. To use iPSCs for therapeutical purposes, integration of OSKM plasmids would be disadvantageous, because these genes may reactivate even after differentiation; or these integrations may disrupt a functional gene in the genome. Therefore, non-integrating methods was employed for iPSC generation such as adenoviruses<sup>57</sup> or Sendai viruses<sup>58</sup>. Also non-integrating OSKM DNAs can be transfected to the cells without using any viruses. Episomal plasmids are a commonly used delivery system of this kind<sup>59</sup>. Using OSKM RNA molecules<sup>60</sup> or their recombinant proteins<sup>61</sup> can also reprogram somatic cells but with recombinant proteins, the efficiency of reprogramming drops dramatically. In summary, there are many different delivery methods for OSKM factor for reprogramming. According to goal of a reprogramming experiment, one those methods can be picked. In this study, I preferred O-S and K-M containing two lentiviral plasmids since their reprogramming

efficiency is very high when compared with other methods and preparation of lentiviruses is cost-effective as well as easy.

**Table 1. Diverse delivery methods for iPSC generation**

Main viral and non-viral delivery of reprogramming factors was summarized in the table with their advantages and disadvantages. These methods can be used with different cocktails that were explained in the text. DOX, doxycycline; PB, *PiggyBac*; LTR, Long terminal repeat. (Adapted from<sup>62</sup> Gonzales et. al. 2011)

Delivery Methods			Advantages	Disadvantages	
Viral	Retrovirus		-very efficient and stable	-genomic integration -only dividing cells	
	Lenti-virus	Constitutive promoter	-very efficient and stable	-genomic integration -residual expression of transgenes	
		Inducible promoter	-very efficient and stable	-genomic integration -transgenes in genome but silenced in the absence of DOX	
		Excisable lentivirus (with LoxP sites)	-transgene free -little scar on genome	-possible LTR integration close to oncogene	
	Adenovirus		-transgene free -no genomic integration	-slow and inefficient	
	Sendai virus		-transgene free -no genomic integration	-takes too much time for viruses to completely lost in the iPSCs	
Non-viral	DNA based	Integrative	PB transposase	-transgene free -average efficiency	-genomic integration -negative selection strongly advised
			DNA with LoxP sites	-transgene free -average efficiency	-genomic integration
		Non-integrative	Non-replicative vectors	-transgene free -no genomic integration	-slow and inefficient -need to check numerous lines to find integration-free ones
			Episomal vectors	-transgene free -no genomic integration	-need to check numerous lines to find integration-free ones -Labour-intensive
	RNA based		-no transgene or genomic integration -no need to screen colonies	-Multiple transfections required	
	Protein based		-no transgene or genomic integration -no need to screen colonies	-slow and inefficient	

### 1.1.3 Stages of iPSC generation

Reprogramming has multiple steps including morphological, epigenetic and metabolic changes. The order of these events may vary or their time latency may differ but at the end of a successful reprogramming, all these changes need to be fulfilled. There are 3 main phases of events which can be listed as initiation, maturation and maintenance (or stabilization). In the first, initiation stage, molecular signatures of the starting cell type gradually disappear. Starting cell generally changes its morphology by a mesenchymal to epithelial transition (MET)<sup>63</sup>. Also, proliferation capacity of starting cell adapts to ESC-like state, becoming more resistant to apoptosis and senescence<sup>64</sup>. Also during the initial stage, metabolism of cells shifts from oxidative phosphorylation towards glycolysis. It was shown that in the hypoxic conditions, reprogramming is more efficient than under normoxic conditions<sup>65</sup>. In the maturation phase, second wave of transcriptional changes take place which includes the upregulation of endogenous pluripotency markers such as Fbx15, Sall4, Oct4, Nanog and Esrrb<sup>64</sup>. This step is a bottleneck for reprogramming because the proportion of cells that can pass through this step is quite low<sup>66</sup>. At the end of this intermediate step, transgenes are silenced which requires nascent pluripotent cells to self-renew independent of ectopic transcription factors. The main event of maintenance stage is the independence of transgenes and endogenous pluripotency genes take over the control for reprogramming to finalize<sup>64</sup>. Since cells are pluripotent at this point, epigenetic changes such as DNA methylations regulate epigenetic memory erasure which fine-tunes the detailed epigenetic

signature of new cell fate. For example in mice, cells undergo telomere elongation and in female cells X chromosome reactivates at this stage<sup>67</sup>.

#### **1.1.4 Transcriptional changes and chromatin reorganization during reprogramming**

Transcription is tightly regulated process because cell identity is directly correlated with the gene expression profile of a cell. Expression of a gene can start a cascade of intracellular events; therefore transcription is regulated in multiple steps to ensure the controlled regulation. For example, initiation of a transcription requires the transcription factor binding and these bindings recruits many other accessory proteins to transcription start site. After transcription starts, expressed genes are controlled by post-transcriptional regulations and mRNAs can be degraded by miRNAs or lncRNAs to regulate genes that are going to be silenced. All transcriptional regulation events work in harmony to fine tune gene expression so that cell state can be determined. In the case of reprogramming, the cell fate is turned back to the ESC stage by changing global gene expression profile. This includes the activation of pluripotency related genes while at the same time, erasure of lineage specific gene expression signatures.

After discovery of direct reprogramming, many researchers showed that gene expression is drastically changed when a somatic cell is reprogrammed into iPSCs<sup>68-70</sup>. Gene expression analysis shows that iPSCs have similar gene expression profile with ESCs, and they are very different from the starting cell's gene expression profile<sup>70</sup>. To activate the pluripotency genes, a positive



feedback loops play a role between endogenous OSK and induced OSK. Towards the end of reprogramming, iPSCs are independent from transgenic OSK since endogenous OSK becomes active. Since gene expression profile illustrates the characteristics of a cell, there are bioinformatics tools to depict the resemblance of an iPSC colony to ESC. For instance, CellNET<sup>71</sup> is a powerful tool to understand whether an iPSC colony is *bona fide* iPSC, or still carrying its starting cell traces. Benefiting from such bioinformatics tools can accelerate the molecular knowledge about reprogramming so that we can find an easier and more precise way to achieve *bona fide* iPSCs.

Another key regulator of gene expression is epigenetic status of the cell. Hence, cell fate is determined with epigenetic factors due to their effect on dynamic chromatin state. Chromatin state can be defined as methylation status of DNA and histone code (post-translational modifications on histone tails in nucleosome). All heritable changes in gene expression that does not change the DNA sequence are referred as epigenetics. Epigenetics contributes to differential gene expression in every cell type. Since life starts with a single cell, zygote and whole organisms form from cells that are divided from zygote and all different cell types have same genetic code; epigenetic regulations are responsible from this diversity. Formation of somatic cells from pluripotent cells during development requires many epigenetic changes and conversely somatic cell reprogramming also requires extensive epigenetic changes<sup>68</sup>. DNA methylation and histone modifications control the gene expression together and these modifications have different outcomes as they can activate or inactivate gene expression. Histone acetylation and methylation are the most studied modifications on histones, but there are many other marks on

histones such as phosphorylation, ubiquitination, sumoylation, crotonylation, etc.

***DNA methylation:*** One of the most important events in initiation of reprogramming is demethylation of promoters of endogenous OSK since pluripotency genes are silenced in somatic cells. DNA hypomethylation can be achieved actively via Tet enzymes or passively with cell division in the absence of DNA methyltransferases (DNMTs)<sup>72</sup>.

***Histone acetylation:*** Acetylation of lysines on histones neutralizes the positive charge, therefore negatively charged DNA binds loosely to the histones. Acetylation of histones correlates with active transcription as it promotes DNA accessibility to transcription factors. For example, treating cells with valproic acid (VPA), an histone deacetylase (HDAC) inhibitor, yields more iPSC colonies<sup>73</sup>. Open chromatin structure might favor the activation of pluripotency-related genes activation. As a result, OSKM can easily activate pluripotency circuit.

***Histone methylation:*** There are multiple Lysine (K) residues on histone tails that can be mono-, di-, tri-methylated. Different from acetylation of histones, methylation of each residue has distinct effect on gene expression regulation. In general, H3K4, H3K36 and H3K79 methylations are correlated with active genes while H3K9 and H3K27 methylations are related with silenced genes. There are also bivalent histones that have both active mark H3K4me3 and repressive mark H3K27me3 at the same time<sup>74</sup>. Bivalency has an essential role during development since erasure of one mark results in robust differentiation of cells.

**H3K4 methylation:** In ESCs, H3K4me3 mark is present on almost every promoter regardless of their activity, however the active genes have Polymerase II bound on the promoter whereas inactive genes have H3K27me3 marks, too<sup>75</sup>. H3K4 methylation is modified with MLL complex proteins and those proteins are effective on ESC self-renewal and differentiation regulation. For example, WDR5 protein is an essential component of MLL complex and it directly interacts with OCT4<sup>76</sup>. Also, it is known that WDR5 is required for both ESC self-renewal and iPSC generation<sup>76</sup>. Conversely, another component of MLL complex, DPY30 is required for maintaining the neuronal differentiation of stem cells<sup>77</sup>. Demethylation of H3K4 residue is also very important to maintain the balance of self-renewal and differentiation. KDM1a (LSD1) is a H3K4me2 demethylase and its deficiency triggers ESCs to differentiate<sup>78</sup>. Also it was shown that H3K4me3/2 demethylase, KDM5b is essential for self-renewal of ESCs<sup>79</sup>. Even though it is clear that H3K4 methylation is important for regulation of pluripotency, effect of other components of MLL complex need to be investigated.

**H3K9 methylation:** H3K9 methylation is associated with transcriptional silencing. Knock-down of Jmjd1a and Jmjd2c, H3K9 demethylases, blocks the self-renewal of ESCs<sup>80</sup>. When H3K9 methyltransferase Suv39H1 was knocked-down, it increases the reprogramming efficiency<sup>36</sup>. On the contrary, another H3K9 methyltransferase EHMT1 knock-down decreases reprogramming significantly<sup>36</sup>. Therefore, it is controversial how H3K9 methylation effects reprogramming.

**H3K27 methylation:** H3K27 methylation is a well-known modification for gene silencing and that also generates bivalency together with H3K4me3

mark. PRC2 complex methylates H3K27 residue and the proteins in this complex (Jarid2, Mtf2 and Esprc2p48) promote reprogramming<sup>81</sup>. Conversely, knock-down of Jarid2, Mtf2 and Esprc2p48 significantly decreases the reprogramming<sup>81</sup>. Knock-down of H3K27 demethylase, Jmjd3 increases reprogramming<sup>82</sup>. On the contrary, H3K27 demethylase, UTX was demonstrated as an essential protein for reprogramming<sup>83</sup>.

**H3K36 methylation:** H3K36 methylation has a role in transcriptional elongation and this mark is largely found in gene bodies of actively transcribed genes. H3K36 demethylase, Kdm2b enhances reprogramming and vitamin-C synergistically promotes reprogramming with Kdm2b<sup>84</sup>.

**H3K79 methylation:** H3K79 methylation found on actively transcribed genes. When H3K79 methylation decreases, reprogramming efficiency increases due to the faster silencing of fibroblast specific genes and enhanced MET<sup>36</sup>. DOT1L is the sole enzyme that methylates H3K79 residue and its knock-down or inhibition with small molecule inhibitors result in increased iPSC generation<sup>36</sup>. Even though DOT1L is the sole enzyme that has the catalytic activity to H3K79 methylation, it acts with other proteins (AF9, AF10, AF17, ENL) as a complex. Other proteins' effect on self-renewal or reprogramming is not known.

**Chromosome remodelers:** Chromosome remodelers are also key players of epigenetic regulations hence they are important during reprogramming. For example, overexpression of BAF complex proteins increases the reprogramming efficiency via promoting Oct4 binding to its target sequences in the genome<sup>85</sup>.

All these examples show that epigenetic regulation has a crucial role in reprogramming. According to all these information, it is not surprising to see that epigenetic researches gained a lot of attention in the stem cell field. To unravel the molecular basis of reprogramming, better understanding of transcriptional order of events and epigenetic regulations is required.

## **1.2 Epigenetic modifiers as barriers and enhancers of reprogramming**

As discussed previously, the importance of epigenetics on reprogramming was appreciated very quickly in the stem cell field. Many of the epigenetic modifiers were tested for their effect on reprogramming in different labs. Some of the epigenetic regulatory proteins were found to be “barriers” of reprogramming since their silencing increase the iPSC generation. On the other hand, other epigenetic regulators were identified as “enhancers” of reprogramming due to their overexpression facilitates the reprogramming. Similar to enhancers, a few proteins were characterized as “essentials” of reprogramming since their silencing blocks the reprogramming event even in the presence of OSKM factors. In summary, enhancer epigenetic regulators facilitate the reprogramming while barriers obstruct the reprogramming.

As previously described, DNMT1, the maintenance DNA methyltransferase is a barrier while Tet DNA demethylases are enhancers of reprogramming. Triple knock-out of Tet1/2/3 proteins in the mice resulted in impaired reprogramming which showed that Tet proteins are essential for reprogramming<sup>86</sup>. PRCD2 complex proteins and BAF complex are also

enhancers of reprogramming while DOT1L and HDAC are barriers of reprogramming<sup>36,73,81,85</sup>. In addition to these, there are many other barriers and enhancers identified. For example, H3K9 methyltransferase, Suv39H1 depletion increases reprogramming efficiency therefore it is a barrier<sup>36</sup>. On the contrary, H3K9 demethylase, Kdm3/4 overexpression enhances reprogramming hence Kdm3/4 is an enhancer of reprogramming<sup>87</sup>. Also, a histone variant can be a barrier or enhancer of reprogramming. For instance, MacroH2A is a barrier of reprogramming since its deletion increases the reprogramming<sup>88</sup>. It has been shown that MacroH2A occupies pluripotency genes silent in fibroblasts<sup>88</sup> therefore its deletion favors the pluripotency.

Even though many barriers and enhancers of reprogramming were identified (**Table 2**), the crosstalk between these epigenetic regulators during reprogramming should be further investigated. The reprogramming process is complicated and many proteins need to function in synchrony. All these new findings will increase the understanding of the molecular orchestra of reprogramming so that *bona fide* iPSCs can be produced more efficiently and safely.

**Table 2. Epigenetic regulators of somatic cell reprogramming and their effects**

Table is adapted from previously published review<sup>89</sup>. (OE, overexpression; KD, knock-down; KO, knock-out; +, enhancer of reprogramming; -, inhibitor of reprogramming; M, mouse; H, human)

	Protein	Phenotype	Species
<b>Chromatin Remodelers</b>	BRG1 and BAF155	OE +	M
	BRM and BAF170	KD +	M
	CHD1L	OE +; KD -	M
	CHD4	KD +	M
	INO80	KD -	M
<b>Histone Modifiers</b>	BMI1	KD -; OE +	M, H
	RING1	KD -	H
	WDR5	KD -	M
	LSD1	KD +	H
	EHMT1 and SETDB1	KD -	H
	EHMT1/2 and SETDB1	KD +	M
	SUV39H1/2	KD +	M, H
	JMJD1A/1B, 2B, 2C	KD-	M, H
	EED and SUZ12	KD -	H
	EZH1	KD +	H
	EZH2	KD -; KO -; OE +	M, H
	JMJD3	KD +; KO +; OE -	M, H
	UTX	KD -; KO -	M, H
	JHDM1A/1B	KD -; OE +	M
	DOT1L	KD +	M, H
	GCN5	KD -	M, H
	HDAC2	KD +	M
<b>DNA Modifiers</b>	DNMT1	KD +	M, H
	AID	KD -; KO -; OE +	M, H
	TET1, 2, 3	KD -; OE +	M
<b>Other Epigenetic Regulators</b>	ASF1A	KD -; OE +	H
	CHAF1A, B	KD +	M
	MacroH2A	KD +; KO +; OE -	M, H
	BRD4	KD -; OE +	M, H
	CBX3	KD +	M
	SIN3A	KD +	M

### **1.3 Dot1L (H3K79methyltransferase) acts as a barrier to reprogramming**

Reprogramming requires global remodeling of the epigenome, therefore investigation of chromatin regulators' role in reprogramming can provide more insight. In a previous study, an shRNA screen was performed to identify chromatin modifiers that have an effect on reprogramming in human origin fibroblasts<sup>36</sup>. Dot1L, along with SUV39H1 and YY1 were identified as blockers of reprogramming. It was reported that knock down of DOT1L can substitute Klf4 and c-Myc in reprogramming factors<sup>36</sup>. ChIP-seq of DOT1L inhibited cells showed that the decrease in the H3K79 methylation levels in lineage specific genes accelerated the generation of pluripotency<sup>36</sup>. Correspondingly, shRNA-mediated silencing of Dot1L in mESCs impairs the differentiation into embryoid body and DOT1L regulated genes were identified as differentiation related genes<sup>90</sup>. These studies show that DOT1L and H3K79 methylation have crucial roles in pluripotency and yet, it is unclear how DOT1L acts in pluripotency and differentiation. Further studies are required to reveal DOT1L's mechanism of action and H3K79 methylation during reprogramming.

DOT1L is a histone lysine methyltransferase that is responsible of mono-, di-, tri-methylation of H3K79 residue. Unlike other histone modifications, H3K79 is localized in the globular domain of histone and DOT1L can catalyze methylation specifically in the nucleosome structure<sup>91</sup>. DOT1L is the sole enzyme that is capable of H3K79 methylation and the same modification is carried out in other organisms with homologs of DOT1L. Even though



DOT1L was discovered as a blocker of reprogramming, its mechanism of action during reprogramming is still unclear.

### 1.3.1 Discovery of Dot1 protein

Dot1 gene was identified in a study that was investigating the disruptors of telomeric silencing (Dot) in *S. cerevisiae*<sup>92</sup>. In this genetic screen was performed to identify genes that have a role during telomeric silencing and Dot1 was the strongest candidate therefore gene was named as Dot1. The major function of Dot1 is H3K79 methylation, and DOT1L is the only H3K79 methyltransferase<sup>93</sup>. When the mammalian homolog of Dot1 was identified, it was named as Dot1-like (DOT1L) protein, also referred as KMT4, and it has been revealed that DOT1L is evolutionarily conserved from yeast to human<sup>94</sup>. When DOT1L was discovered as a H3K79 methyltransferase, it was categorized as a novel class of histone lysine methyltransferase (KMT) due to the absence of SET domain (Suvar3-9, Enhancer of Zeste, Trithorax) in DOT1L<sup>92</sup>.

### 1.3.2 Structure of DOT1L protein

Crystal structure of catalytic domain of human DOT1L protein in complex with methyl donor SAM (S-adenosyl-L-methionine) was identified in 2003<sup>91</sup>. This work demonstrated how SAM can bind to a pocket at the N-terminus of DOT1L<sup>91</sup>. It was also demonstrated that SAM cannot bind to GSG<sub>163-165</sub>RCR

mutant of DOT1L<sup>94</sup> and in structural analysis it was shown that indeed these amino acids were found in the backside of the SAM binding pocket<sup>91</sup>. GSG<sub>163-165</sub>RCR mutation is within the DxGxGxG signature motif that was previously found in SAM binding proteins<sup>95</sup>. This motif is found in hDOT1L as DLGSGVGQ and mutation of GSG residue blocks DOT1L binding to SAM<sup>94</sup>. In the structural study, it was also claimed that lysine binding channel of DOT1L is separate from SAM binding pocket; therefore they claimed that multiple rounds of H3K79 methylation can be processed without releasing from the enzyme<sup>91</sup>.

In previous studies, it was shown that DOT1L methylates H3K79 in the nucleosome form rather than free histone form<sup>94</sup>. In the structural study it was demonstrated that “*DOT1L interacts with all four histones on the same side of the nucleosome disk surface*”<sup>91</sup> which explains the reason for DOT1L preferring the nucleosome structure to perform enzymatic activity. This finding was supporting DOT1L’s binding to ubiquitinated H2B, since it was demonstrated that H2BK123 ubiquitination is required for H3K79 methylation in yeast<sup>96</sup>. Later, it was confirmed biochemically that H2B ubiquitination is required for H3K79 methylation in human, as well<sup>97</sup>.

### 1.3.3 DOT1L’s diverse roles in the cell

Researchers discovered that Histone H3 is methylated at the globular residue of lysine 79 and Dot1 and its mammalian homolog of DOT1L protein was found to be the sole responsible histone methyltransferase for H3K79 methylation<sup>94,98-100</sup>. Even though structural properties of DOT1L resembles

arginine methyltransferases<sup>93</sup>, DOT1L's arginine methyltransferase activity could not be demonstrated<sup>100</sup>. However, in recent studies, DOT1L's diverse functions in cellular activities have been discovered in telomeric silencing, transcriptional regulation, cell cycle regulation, DNA damage response, development and leukemia.

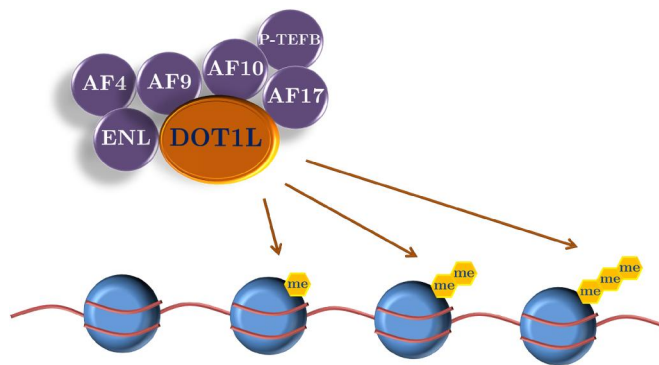
#### *1.3.3.1 Role of DOT1L in telomeric silencing*

As stated previously, yeast Dot1 plays a role in telomeric silencing. Dot1's effect on heterochromatin formation was dependent on Sir (silent information regulator) proteins, since Sir3 binds H3K79 and silences genes at telomeric region<sup>101</sup>. Since Sir3 and Dot1 compete to bind to the same residue, the interplay between H3K79 methylation and Sir Protein binding regulates the telomeric silencing. Similar mechanism was also revealed in human. Sir2 homolog SIRT1 was demonstrated to regulate histone deacetylation-dependent silencing interplaying with DOT1L<sup>102</sup>. DOT1L blocks SIRT1's binding therefore promotes open chromatin whereas SIRT1 binding to chromatin stimulates the silencing<sup>102</sup>.

#### *1.3.3.2 Role of DOT1L in transcriptional regulation*

DOT1L plays a role in elongation of transcription. DOT1L is recruited to different transcriptional elongation complexes (**Figure 2**). For example, DOT1L was found in a complex with AF4, AF9, AF10, ENL and p-TEFb,

where p-TEFb phosphorylates the C-terminal domain (CTD) of RNA Polymerase II (RNAPII)<sup>103</sup>. Phosphorylation of RNAPII from its CTD is a key event during transition from initiation to elongation of transcription. Also, DOT1L was found in core ENL-associated proteins (EAPs) along with AF4 and p-TEFb<sup>104</sup>. Later, DOT1L complex (DotCom) was identified along with AF9, AF10, AF17, ENL and Wnt pathway proteins<sup>105</sup>. On the other hand, there are other elongation complexes that do not contain DOT1L, such as superelongation complex (SEC) and AEP (AF4, ENL, p-TEFb)<sup>106,107</sup>. Presence of DOT1L not being consistently found in all the elongation complexes points to two possible reasons: (1) insufficient experimental designs, (2) dynamic elongation complexes formed for transcription of subset genes at different time points of cell cycle. Although DOT1L is not found in every elongation complex, it plays a role in elongation of at least a subset of genes.



**Figure 2. Interaction partners of DOT1L during transcriptional elongation**

### ***1.3.3.3 Role of DOT1L in DNA damage response***

A few studies have associated DOT1L with double strand break (DSB) repairs. In the first study that detected the DNA repair role of DOT1L, 53BP1 protein was found to bind to methylated H3K79<sup>108</sup>. 53BP1 protein binds to p53, which is important in DNA damage response pathways<sup>109</sup>. Later, yeast homolog of 53BP1, Rad9, was also demonstrated to bind to methylated H3K79<sup>110</sup>. It has been shown that DOT1L knock down increases sensitivity to ionizing radiation (IR) and ultraviolet radiation (UV) in both yeast and mammalian cells<sup>111,112</sup>. For example, HLA-B-associated transcript 3 (BAT3) protein recruits DOT1L to the chromatin site, and BAT3 knock-down decreases 53BP1 foci, and sensitizes cells to IR<sup>113</sup>. Even though it is unclear how H3K79 methylation regulates DNA damage response, DOT1L performs strategic H3K79 methylation marks on the genome to protect genomic stability. It was established that H3K79 methylation is enriched around the origin of replication sites in genome and loss of DOT1L disrupts replication<sup>114</sup>. Taken together, DOT1L has a primary role in maintaining genomic stability via H3K79 methylation at DNA damage site and origin of replication region. However, there are still many unknowns on how DOT1L regulates DNA damage pathways.

### ***1.3.3.4 Role of DOT1L in cell cycle regulation***

Histone methylation pattern is an important regulator of cell cycle, therefore when H3K79 methylation was found, its abundance in different cell cycle

phases were investigated. Interestingly, literature shows no consensus on cell cycle regulation effect of DOT1L, since different organisms displayed distinct H3K79 methylation patterns. In yeast, H3K79 methylation levels are low in the G1 and gradually increases through S, G2 and M phases<sup>115</sup>. On the contrary, HeLa cells have high H3K79me2 levels in G1 phase which gradually decrease through S and G2 phase and increase again during M phase<sup>94</sup>. In mouse ESCs, DOT1L knock out did not affect proliferation, however, in vitro differentiation of DOT1L-KO ESCs displayed defects in proliferation and cells were arrested at G2/M phase<sup>90</sup>. In erythroid lineage cells, DOT1L-KO resulted in G0/G1 accumulation<sup>116</sup>. Different organisms and different cell types exhibiting different patterns of H3K79 methylation shows that cell cycle regulation of DOT1L is not conserved evolutionarily. To understand the cell cycle regulation system of DOT1L, further studies are required.

#### ***1.3.3.5 Role of DOT1L in embryonic development***

In fly embryogenesis, H3K79 methylation was not detected until gastrulation. After that, increasing levels of H3K79 methylation pointed out importance of *grappa*, fly ortholog of yeast Dot1, for developmental genes<sup>117</sup>. It was also reported that *grappa* is associated with Wnt signaling pathway because knock down of *grappa* affects Wingless target genes (*senseless*, *frizzled 3* and *homothorax*) that are actively transcribed<sup>105</sup>. For mouse embryogenesis, H3K79me2 can be detected after blastocyst stage<sup>118</sup>. Dot1L knock out mouse embryo is lethal after day 10 when organogenesis of cardiovascular system starts<sup>119</sup>. Even though mouse oocyte has detectable methylation on H3K79,

remaining methylation disappears following fertilization<sup>118</sup>. This observation implies the existence of an active demethylation enzyme(s), because during fertilization, H3K79 methylation loss was independent from DNA synthesis. However, to date, a histone demethylase that specifically targets H3K79 residue has not been identified. In 2017, KDM4D was proposed to be H3K79me3 demethylation regulator<sup>120</sup>. However, it is not certain that KDM4D can demethylate H3K79me3 residue all by itself, therefore H3K79 demethylation activity needs to be further investigated.

#### ***1.3.3.6 Role of DOT1L in leukemia***

In many cases of Acute Myeloid Leukemia (AML), rearrangement of the *Mixed Lineage Leukemia (MLL)* locus takes place, resulting in the production of MLL fusion proteins. Role of DOT1L during leukemia was revealed as many of the MLL fusion proteins were found to interact with it. MLL is a H3K4 methyltransferase and has a regulatory role during hematopoiesis. However, in some cases *MLL* gene translocates to other chromosomes and translocation product translates an in frame fusion protein. Some examples of leukemia associated MLL fusion proteins are MLL-AF9, MLL-AF10, MLL-ENL. Interesting aspect of the translocations is the fact that fusion partners of MLL are usually the interactors of DOT1L<sup>121</sup>. It is hypothesized that DOT1L's aberrant recruitment via fusion partner of MLL causes the leukemia phenotype. In the case of MLL-AF10 fusion protein triggering leukemia, HoxA9 gene is upregulated and this contributes to the leukemia phenotype<sup>121</sup>. Moreover, deletion of HoxA9 in MLL-AF10 fusion prevents

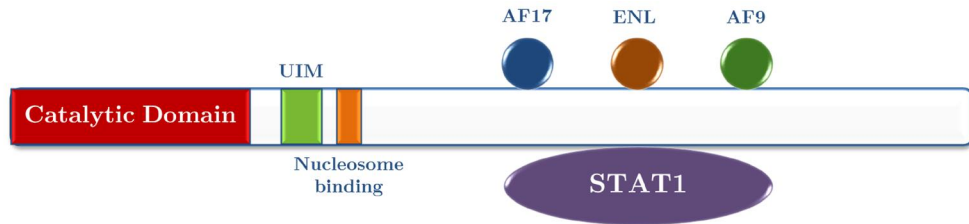
transformation. Similarly, upregulation of HoxA5 gene was responsible in CALM-AF10 translocated leukemia<sup>122</sup>. Other than AF10, there are a few protein fusions of MLL such as AF4 (*AFF4*), AF9 (*MLLT3*), AF17 (*MLLT6*) and ENL (*MLLT1*) which are all involved in elongation complex binding partners of DOT1L<sup>123</sup>. Since DOT1L has an important role during leukemia, one of the approaches of treatment for leukemia was using small molecule inhibitors for DOT1L. The most efficient DOT1L inhibitors were EPZ004777, EPZ005676 and SGC0946 which were tested for their therapeutic effect on leukemia. However, using these drugs might result in side effects such as disruption in normal hematopoiesis because DOT1L has a role during normal hematopoiesis as well<sup>124</sup>.

### 1.3.4 Known protein interactions of DOT1L

DOT1L has many protein-protein interactions that have been identified to date. The first interaction of DOT1L was found in 2005, with AF10<sup>121</sup>. This interaction was very important because it associates DOT1L with leukemia. Later, it was shown that DOT1L was interacting with AF9, AF17 and ENL as well (Figure 3)<sup>93,104,105</sup>. In 2010, affinity purification of DOT1L was performed, followed by mass spectrometry for identification of proteins that are directly interacting with DOT1L<sup>125</sup>. In that study, a few new interactions of DOT1L was detected and NPM1, HNRNPM, DDX21 interactions were also validated with pull-down assays<sup>125</sup>. In 2011, DOT1L was pulled down with STAT1<sup>126</sup> and AF4<sup>127</sup>. In 2012, DOT1L was pulled down with BAT3<sup>113</sup>. All these proteins were direct interactors of DOT1L and their interactions



were validated with immunoprecipitation methods. However, there are also functional interactions of DOT1L, where the proteins do not physically interact with it but work in parallel in similar cellular activities. MLL1 is one of those functional interactors of DOT1L. Their aberrant cooperation results in leukemia which shows that MLL1 and DOT1L's co-existence in the chromatin should be tightly regulated<sup>93</sup>. This makes MLL1 an important functional interactor of DOT1L. SIRT1 and CDK9 are also functional interactors of DOT1L where CDK9 and DOT1L work in parallel in transcriptional elongation. CDK9 is a subunit of p-TEFb and even-though p-TEFb does not interact directly with DOT1L, they coexist in elongation complexes<sup>103,128</sup>. On the other hand, SIRT1 and DOT1L work in opposite functions. DOT1L inhibits SIRT1 binding to chromatin therefore they are working in reverse<sup>102</sup>. While DOT1L makes active mark H3K79me, SIRT1 carries out deacetylation of H3K9 which is a repressive mark<sup>102</sup>. Even though some of the DOT1L interactors were identified; DOT1L-mediated H3K79 methylation mechanism is still unclear. Therefore, different approaches on protein interaction studies are required to identify not only direct interactions but also functional interactions of DOT1L.



**Figure 3. Schematic representation of the DOT1L protein**

Ubiquitin-interaction motif (UIM) is located N-terminus of DOT1L, a lysine-rich region is required for nucleosome binding and interacts with the ubiquitin H2B. Within the lysine-rich region, there is a nucleosome/DNA-binding motif. Adapted from Castelli *et. al.*<sup>129</sup>.

## 1.4 Hypothesis

It has been established that DOT1L acts as a potent blocker of reprogramming of somatic cells. However, it is not known whether the proteins that are interacting with DOT1L have any role in reprogramming. I hypothesize that DOT1L's interactome contributes to reprogramming, and investigating the key factors in DOT1L's interactome will reveal the molecular mechanism of DOT1L's action during reprogramming.

## Chapter 2- Materials and Methods

### 2.1 Plasmids and cloning procedures

For reprogramming experiments and OSKM transcription factor overexpression pSIN-O2S (Addgene # 21162) and pSIN-K2M (Addgene # 21164) vectors were used. For viral packaging of plasmids, pCMV\_VSV-G (Addgene # 8454) envelope protein expressing plasmid was used along with viral packaging plasmids. For lentivirus production, pCMV-dR8.2 ΔVPR (Addgene # 8455) packaging plasmid was used while pUMVC (Addgene # 8449) packaging plasmid was used for retrovirus production.

**Table 3. List of cloning primer sequences**

Primer Name	5' to 3' Sequence
DOT1L-BioID-Q5_top	TGGATATCTGCAGAATTCACcATGGGGGAGAAGCTGGAGCT
DOT1L-BioID-Q5_bottom	AGCTCCAGCTTCTCCCCATgGTGAATTCTGCAGATATCCA
shCloning Forward	GATGGCTGCTCGAGAAGGTATATTGCTGTTGACAGTGAGCG
shCloning Reverse	GTCTAGAGGAATCCGAGGCAGTAGGCA
XhoI-AF10 cloning-Fwd	TATAA <b>CTCGAG</b> ATGGTCTCTAGCGACC
AF10 cloning-stop-XhoI-Rev	GAAAGCTGGGTCTAGATAT <b>CTCGAG</b>

### 2.1.1 BirA\*-DOT1L fusion protein expression plasmid cloning

BirA\* cDNA was used from pcDNA3.1-mycBioID (Addgene # 35700) vector which was a gift from Nurhan Özlü (Koc University). DOT1L wildtype (wt) and mutant (GSG 163-165 RCR) cDNA plasmids were described previously<sup>36</sup>. cDNA sequence of DOT1L was confirmed by sequencing the entire DOT1L in plasmids with primers in **Table 8**. To make a fusion protein, first pcDNA3.1-mycBioID vector was cut with EcoRI (NEB) and KpnI (NEB) enzymes and then treated with Antarctic Phosphatase (AP, NEB). Then, DOT1L cDNAs were cut with EcoRI (NEB) and KpnI (NEB) enzymes and ligated (Quick Ligase, NEB) into pcDNA3.1-mycBioID. Ligation product was transformed into DH5 $\alpha$  competent bacteria.

In the second step, pcDNA3.1-mycBioID vector was cut with XhoI (NEB) and then treated with AP. Then, DOT1L cDNAs that were previously cloned into pcDNA3.1-mycBioID, were cut with XhoI (NEB) enzyme and ligated (Quick Ligase, NEB) into XhoI-cut pcDNA3.1-mycBioID vector. Ligation product was transformed into DH5 $\alpha$  competent bacteria.

In the third step, BirA\*-DOT1L fusion was cloned into pENTR1A vector. NheI cut site was necessary at the end of DOT1L cDNA therefore; an NheI cut site was added into BamHI cut site at the end of DOT1L using short double-stranded oligos phosphorylated with T4 Polynucleotide kinase. The sequences of inserted oligos are listed in the **Table 3**. Then, NheI cut site added BirA\*-DOT1L wt and mut fusion proteins were cut with NheI (NEB) and ligated with XbaI-cut AP-treated pENTR1A no ccDB (Addgene #

17398). Since XbaI and NheI has compatible restriction enzyme sites, BirA\*-DOT1L wt and mut fusion sequence was ligated into pENTR1A vector. Ligation product was transformed into Stbl3 competent bacteria.

In fourth step, 1 nucleotide was added between the junction of BirA\* and DOT1L so that codon order of fusion protein will be in frame with BirA\* and DOT1L. For this cloning, Q5 Site-Directed Mutagenesis Kit (NEB) was used with primers that are listed in the **Table 3**. This cloning procedure was confirmed with sequencing of the base addition site with primers in **Table 8**. In the last step, in frame BirA\*-DOT1L wt and mut fusion sequences cloned into an expression plasmid pLEX-307 (Addgene # 41392) via LR cloning (Invitrogen). All these cloning steps are summarized in **Figure 4**.

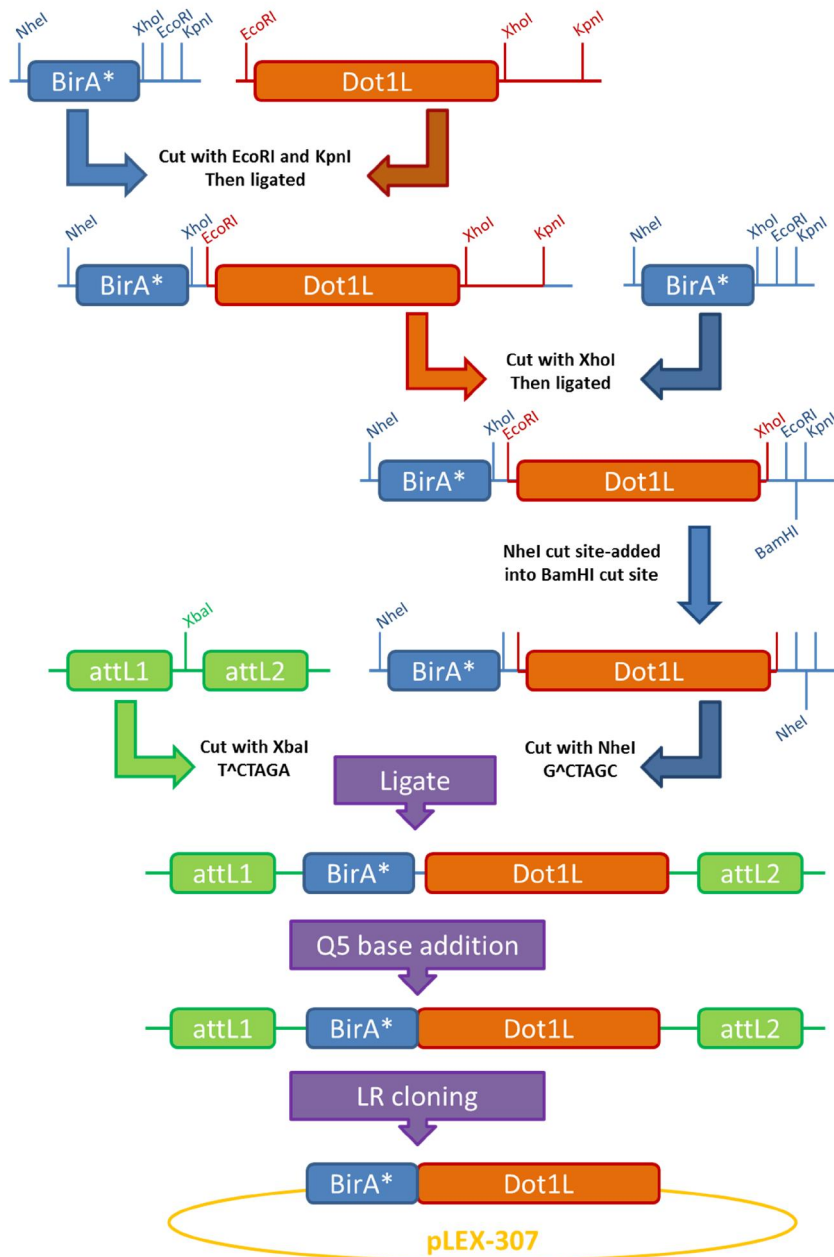
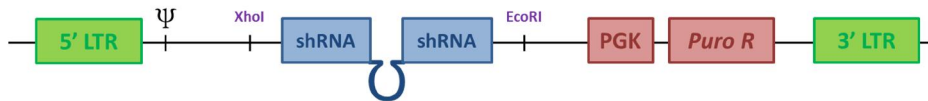


Figure 4. Cloning steps of BirA\*-DOT1L fusion protein expressing plasmid

### 2.1.2 shRNA cloning into pSMP vector

RNAi Codex (<http://cancan.cshl.edu/cgi-bin/Codex/Codex.cgi>) was used to design specific shRNAs to each gene of interest (**Table 4**). Synthetic 97-mer oligonucleotides (Macrogen Inc.) were cloned into pSMP plasmid. As a control shRNA, a firefly luciferase targeting shRNA plasmid, pSMP-Luc (shFF, Addgene # 36394) was used. shControl plasmid (shFF) was used as a backbone for cloning of other shRNAs (**Figure 5**).



**Figure 5. Sketch of shRNA cloned pSMP plasmid**

pSMP-Luc (shFF) vector was cut with EcoRI (NEB) and XhoI (NEB) enzymes and treated with AP enzyme. shRNA oligos were amplified with shCloning primers (**Table 3**) and PCR products were cut with EcoRI (NEB) and XhoI (NEB) enzymes. After enzymatic digestion, PCR products were run in a 2% agarose gel. 110bp bands were cut out and gel purified (Gel extraction kit, MN). Gel extracted oligos were ligated (Quick Ligase, NEB) into pSMP backbone. Ligation product was transformed into Stbl3 competent bacteria. All vectors were confirmed by Sanger sequencing using MSCV-fwd primer (CCCTTGAACCTCCTCGTTCGACCT).

Table 4. List of shRNA oligo sequences

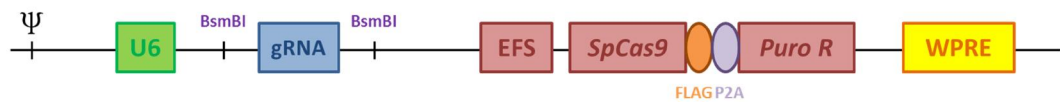
shRNA Name	5' to 3' Sequence
shAF10-1	TGCTGTTGACAGTGAGCGAGCCGAGAACCCTGGTTTATTAGTGAAGCCACAGATGTAATAAACACAGCGGTTCTCGGCCTGCCTACTGCCTCGGA
shAF10-2	TGCTGTTGACAGTGAGCGCGGTTCATATGATCAAAGTTAATAGTGAAGCCACAGATGTAATAAACCTTTCATATGACCTTGCCTACTGCCTCGGA
shNONO-1	TGCTGTTGACAGTGAGCGAAGGAAGAAATGAGGAACTATTAGTGAAGCCACAGATGTAATAGTTTCTCATTCTCTCTGCCTACTGCCTCGGA
shNONO-2	TGCTGTTGACAGTGAGCGAATGGAAGAGCTGCACAACCAATAGTGAAGCCACAGATGTAATGGTTGTGCAGCTCTCCATCTGCCTACTGCCTCGGA
shKAISO-1	TGCTGTTGACAGTGAGCGAGGCACTATTAGGAGTGAATAGTGAAGCCACAGATGTAATTTCACTCTAATAACTGCCTACTGCCTCGGA
shKAISO-2	TGCTGTTGACAGTGAGCGCCTGTAGCAAGATGCTGTTAATAGTGAAGCCACAGATGTAATAACAGCATCTTGCTACAGATGCCTACTGCCTCGGA
shSIN3B-1	TGCTGTTGACAGTGAGCGCCCGGAAATGATTATGCTTAGTGAAGCCACAGATGTAATGCATAATCAATTTCCCGCTTGCCTACTGCCTCGGA
shSIN3B-2	TGCTGTTGACAGTGAGCGCCCGTGCATCGCACTCTCAATAGTGAAGCCACAGATGTAATGAGAGTGCATGTCAGCGGATGCCTACTGCCTCGGA
shAF17-1	TGCTGTTGACAGTGAGCGGGCATTGAAGAGGACTGAATAGTGAAGCCACAGATGTAATATCAGTCTCTTCAATGCCTGCCTACTGCCTCGGA
shAF17-2	TGCTGTTGACAGTGAGCGACAGGCTGTCTCAACAGCCTTATAGTGAAGCCACAGATGTAATAAGGCTTGTGAGACAGCTGCTGCCTACTGCCTCGGA
shMRE11-1	TGCTGTTGACAGTGAGCGCCCTAATAGTTGAACAGATATTAGTGAAGCCACAGATGTAATATCTGTTCAAATATTAGTGCCTACTGCCTCGGA
shMRE11-2	TGCTGTTGACAGTGAGCGAGGCCATGAACATGAGTGAATAGTGAAGCCACAGATGTAATTTACACTCATGTTCAATGCCTGCCTACTGCCTCGGA
shENL-1	TGCTGTTGACAGTGAGCGACAGCAGATTGTGAATCTGATCTAGTGAAGCCACAGATGTAGATCAGATTCAATCTGCTGCTGCCTACTGCCTCGGA
shENL-2	TGCTGTTGACAGTGAGCGCGATTGTTTCTTCTGGATTTAGTGAAGCCACAGATGTAATCCAGGAAGAAACAATCTTGCCTACTGCCTCGGA
shNUMA1-1	TGCTGTTGACAGTGAGCGAGCACTGAAGAGGGACAGCAATAGTGAAGCCACAGATGTAATTTGCTGCCTCTTCAATGCCTGCCTACTGCCTCGGA
shNUMA1-2	TGCTGTTGACAGTGAGCGCCCTGAAGAGAAGAACAATAGTGAAGCCACAGATGTAATTCGTTCTTCTTCAAGGATGCCTACTGCCTCGGA
shTPR-1	TGCTGTTGACAGTGAGCGACCAAGTCTGCCAGAACAAATAGTGAAGCCACAGATGTAATTTGTTCTGGACAGACTGGGCTGCCTACTGCCTCGGA
shTPR-2	TGCTGTTGACAGTGAGCGCACCTCCATAGAGCTGTAATAGTGAAGCCACAGATGTAATTTACAGCTATGAGAGTGCCTGCCTACTGCCTCGGA
shDDX21-1	TGCTGTTGACAGTGAGCGCCCTCCCTTTGATTGAGAAATAGTGAAGCCACAGATGTAATTTCTCAATCAAAGGGATGGCATGCCTACTGCCTCGGA
shDDX21-2	TGCTGTTGACAGTGAGCGCCTGATCAAGTGAAGAGATTTAGTGAAGCCACAGATGTAATCTCTCCACTTGCATGATGCCTACTGCCTCGGA
shAF4-1	TGCTGTTGACAGTGAGCGCCTTACTCTGCTACTCAGAATAGTGAAGCCACAGATGTAATCTGAGTAGACAGAGTAAGCTGCCTACTGCCTCGGA
shAF4-2	TGCTGTTGACAGTGAGCGACAGCTACAAGAATTAACAAATAGTGAAGCCACAGATGTAATTTGGTTAATCTTGTAGCTGCCTACTGCCTCGGA
shAF9-1	TGCTGTTGACAGTGAGCGCCCTTCTTATTGACTTATTAGTGAAGCCACAGATGTAATAAGTCATAATCAAAGCGGATGCCTACTGCCTCGGA
shAF9-2	TGCTGTTGACAGTGAGCGCACACTGCCTTATTACATAATAGTGAAGCCACAGATGTAATGTAATAAGGCAAGTGTGTTGCCTACTGCCTCGGA
shBAT3-1	TGCTGTTGACAGTGAGCGAGCAGCAGCTCCGGTCTGATATTAGTGAAGCCACAGATGTAATATCAGACCGGAGCTGCTGCCTGCCTACTGCCTCGGA
shBAT3-2	TGCTGTTGACAGTGAGCGCCCTCACAGTATTAAGAAATAGTGAAGCCACAGATGTAATTTCTAATACTGTGAAGGATGCCTACTGCCTCGGA
shSIRT1-1	TGCTGTTGACAGTGAGCGCATCTTGCCTGATTGTAAATAGTGAAGCCACAGATGTAATTTACAAATCAGGCAAGATGCTGCCTACTGCCTCGGA
shSIRT1-2	TGCTGTTGACAGTGAGCGACCATGGAGGATGAAAGTGAATAGTGAAGCCACAGATGTAATTTCACTTTCATCTCCATGGGTGCCTACTGCCTCGGA
shSTAT1-1	TGCTGTTGACAGTGAGCGCCAGCTGTACTCAAGAAGATGAGTGAAGCCACAGATGTAATCTTCTGAGTAACAGCTGTTGCCTACTGCCTCGGA
shSTAT1-2	TGCTGTTGACAGTGAGCGCCCTAAAGGAATGATATTAGTGAAGCCACAGATGTAATATCAGTCTCTTTCAGGCAATGCCTACTGCCTCGGA
shHNRNPM-1	TGCTGTTGACAGTGAGCGCGCATAGGATTTGGAATAAATAGTGAAGCCACAGATGTAATTTTCCAAATCTTGCCTGCCTACTGCCTCGGA
shHNRNPM-2	TGCTGTTGACAGTGAGCGCGATGTATAAAGATGTTAAATAGTGAAGCCACAGATGTAATTTAAACATCTTATACATCCATGCCTACTGCCTCGGA



<b>shNPM1-1</b>	TGCTGTTGACAGTGAGCGGGAGGAAGTCTCTTTAAGAAAGTAGTGAAGCCACAGATGTAATCTTTAAAGAGACTTCTCTGCCTACTGCCTCGGA
<b>shNPM1-2</b>	TGCTGTTGACAGTGAGCGAAAGGTTCCACAGAAAAAAGTATAGTGAAGCCACAGATGTATACTTTTTCTGTGGAACCTGTGCCTACTGCCTCGGA
<b>shCDK9-1</b>	TGCTGTTGACAGTGAGCGCCCGTCAAGGGTAGTATATATAGTGAAGCCACAGATGTATATACTACCCTTGACGCGGTTGCCTACTGCCTCGGA
<b>shCDK9-2</b>	TGCTGTTGACAGTGAGCGAGCAGTTTTGTCCGTTAGAATAGTGAAGCCACAGATGATTCTAACGGACCAAACTGTGCCTGCCTACTGCCTCGGA
<b>shMLL1-1</b>	TGCTGTTGACAGTGAGCGCGTCTTATTCCGAAACCAATATAGTGAAGCCACAGATGTATTGGTTGCGAATAAGACCTGCCTACTGCCTCGGA
<b>shMLL1-2</b>	TGCTGTTGACAGTGAGCGCTGGGATCTAGTCCAGAGATATAGTGAAGCCACAGATGTATATCTCTGGAACATAGATCCCATGTGCCTACTGCCTCGGA
<b>shMLL1-3</b>	TGCTGTTGACAGTGAGCGGGACCGCTACTGATCTTGAATGATAGTGAAGCCACAGATGTACATTCAAGATCAGTACGCGTCTGCCTACTGCCTCGGA
<b>shWDR5-1</b>	TGCTGTTGACAGTGAGCGCGCAAGTTCATCTGCTGATAATAGTGAAGCCACAGATGATTATCAGCAGATGAACCTGCCATGCCTACTGCCTCGGA
<b>shWDR5-2</b>	TGCTGTTGACAGTGAGCGCCCTCAACAGCTTGCACCAATAGTGAAGCCACAGATGATTGGGTGACAAGCTGTTGAGGTTGCCTACTGCCTCGGA
<b>shRBBP5-1</b>	TGCTGTTGACAGTGAGCGCCATTAAACCGAAACTACTAGTGAAGCCACAGATGTAGTAGAGTTTCGGTTTAAATGGATGCCTACTGCCTCGGA
<b>shRBBP5-2</b>	TGCTGTTGACAGTGAGCGATGGGCACAGAATCAAGTAGAATAGTGAAGCCACAGATGATTCTACTTGTGTTGCGCCAGTGCCTACTGCCTCGGA
<b>shASH2L-1</b>	TGCTGTTGACAGTGAGCGCCGAGTAACTACTTATTAATAGTGAAGCCACAGATGTATAAATAAGTTAGTTACTCGGATGCCTACTGCCTCGGA
<b>shASH2L-2</b>	TGCTGTTGACAGTGAGCGATCCAAAGATAAAGATATTATATAGTGAAGCCACAGATGTATAAATATCTTCTTTGGAGTGCTACTGCCTCGGA
<b>shDPY30-1</b>	TGCTGTTGACAGTGAGCGCCAAATCCCATTGAATTTCTATAGTGAAGCCACAGATGTAGAAATCAATGGGATTTGGTTGCCTACTGCCTCGGA
<b>shDPY30-2</b>	TGCTGTTGACAGTGAGCGGTTAACATATTTCCCTTATTTAGTGAAGCCACAGATGAAATAAGGAAATATGTTAACCTGCCTACTGCCTCGGA
<b>shMEN1-1</b>	TGCTGTTGACAGTGAGCGACCGAGTACAGTCTGTATCAAATAGTGAAGCCACAGATGATTGATACAGACTGTAACCGGTGCCTACTGCCTCGGA
<b>shMEN1-2</b>	TGCTGTTGACAGTGAGCGACCGGGAAGACGAGGAGATCTATAGTGAAGCCACAGATGTATAGATCTCTCTCTCCCGGTGCCTACTGCCTCGGA
<b>shWDR82-1</b>	TGCTGTTGACAGTGAGCGCCAAATGATCTTAATTTGTTATTAGTGAAGCCACAGATGTAATAACAATTAAGATCATTGGTTGCCTACTGCCTCGGA
<b>shWDR82-2</b>	TGCTGTTGACAGTGAGCGCACAGTTGTTACAGCTCTATAGTGAAGCCACAGATGTATAGAGCTGAAACAACCTGTTTGCCTACTGCCTCGGA
<b>shHCFC1-1</b>	TGCTGTTGACAGTGAGCGAGCCATGCTCTCCAGAAATAGTGAAGCCACAGATGTAATTTCTGGAGAGGACATGGGCTGCCTACTGCCTCGGA
<b>shHCFC1-2</b>	TGCTGTTGACAGTGAGCGAACCGTTCACTATTGTAGAGTATAGTGAAGCCACAGATGTATACTCTACAATAGTGAACGGTGTGCCTACTGCCTCGGA
<b>shMLL2-1</b>	TGCTGTTGACAGTGAGCGAGCAGTTGGCTAGTGAACCTATAGTGAAGCCACAGATGTATAAGTTCACTAGCCAACTGCCTGCCTACTGCCTCGGA
<b>shMLL2-2</b>	TGCTGTTGACAGTGAGCGAAAGGTGTGGCTGACAGAAATAGTGAAGCCACAGATGATTCTGTACGCCACACCTTCTGCCTACTGCCTCGGA
<b>shMLL3-1</b>	TGCTGTTGACAGTGAGCGCAAGCAAGATAAGTTTAGATAAATAGTGAAGCCACAGATGATTATCTAAACTTCTTCTGTTTGCCTACTGCCTCGGA
<b>shMLL3-2</b>	TGCTGTTGACAGTGAGCGACAGGAGTAGATAGACAAAGATAGTGAAGCCACAGATGTATCTTGTCTATCTACCTCTGCTGCCTACTGCCTCGGA
<b>shMLL4-1</b>	TGCTGTTGACAGTGAGCGGGGCCAGAAACACATTGTTATTAGTGAAGCCACAGATGTATAACAATGTGTTTCTGGCCCTGCCTACTGCCTCGGA
<b>shMLL4-2</b>	TGCTGTTGACAGTGAGCGCTACCGAAGTGTGACAAAAATAGTGAAGCCACAGATGTATATTTGTACACTCCGGTATTGCCTACTGCCTCGGA
<b>shCXXC1-1</b>	TGCTGTTGACAGTGAGCGATCAGAGCAAAACATACTGTAATAGTGAAGCCACAGATGATTACAGATGTTTGTCTGAGTGCCTACTGCCTCGGA
<b>shCXXC1-2</b>	TGCTGTTGACAGTGAGCGATCCCTGGTTTTGTTAATAAATAGTGAAGCCACAGATGATTATTAACAAAACCCAGGGAGTGCCTACTGCCTCGGA
<b>shSET1B-1</b>	TGCTGTTGACAGTGAGCGCGTCTCATCCGCTCATCATTAGTGAAGCCACAGATGTAATGATGACGCGGATGAGGACGATGCCTACTGCCTCGGA
<b>shSET1B-2</b>	TGCTGTTGACAGTGAGCGCGGAGATTACCTATGACTATAATAGTGAAGCCACAGATGATTATAGTCATAGGTAATCTCTGCCTACTGCCTCGGA
<b>shFF</b>	TGCTGTTGACAGTGAGCGCCCGCTGAAGTCTGTGATAAATAGTGAAGCCACAGATGATTAAATCAGAGACTCAGGCGGTTGCCTACTGCCTCGGA

### 2.1.3 gRNA cloning into lentiCRISPRv2 vector

gRNAs that are targeting *AF10* gene are gifts from Or Gozani Lab (**Table 5**)<sup>130</sup>. gRNAs targeting MLL1 and DOT1L were cloned into lentiCRISPRv2 (Addgene # 52691) vector (**Figure 6**). Cloning of these gRNA plasmids were performed by members of the Önder lab. gNT1, gNT2 (non-targeting control) and gDOT1L (targets exon1) were cloned by Can Aztekin; gDOT1L-1 & gDOT1L-2 (targets exon 5) were cloned by Eray Enüstün; gMLL1-1, gMLL1-2 & g-MLL1-3 were cloned by Kenan Sevinç (**Table 5**). Cloning protocol of these gRNAs were carried out as described (Sanjana *et. al.*, 2014)<sup>131</sup>. This cloning procedure was confirmed with sequencing of gRNAs with U6 promoter sequencing primer in **Table 8**.



**Figure 6.** Sketch of gRNA cloned lentiCRISPRv2 plasmid

**Table 5. List of gRNA targeting sequences**

gRNA Name	5' to 3' Targeting Sequence
sgControl	CTTCGAAATGTCCGTTCCGGT
sgAF10-1	TGCAGCGTCGCGGTGCATCA
sgAF10-2	ATAAATAGTCCTTACCACTC
gNT1	ACGGAGGCTAAGCGTCGCAA
gDOT1L-1 (exon5)	GTCCACAAACAGGTCGTCGT
gDOT1L-2 (exon5)	GGTCTCCCCGTACACCTCGG
gNT2	CGCTTCCGCGGCCCGTTCAA
gDOT1L (exon1)	CTGAGCCCGCCGTCTACCCG
gMLL1-21	TTGTAGGATGAGCAATTCTT
gMLL1-22	CCACCCTGAGTGCCTTACCA
gMLL1-753	CAGCAGCCTTTAGATCTAGA

#### 2.1.4 T7-endonuclease assay

gRNA infected dH1f cells were harvested and genomic DNA were isolated using MN Nucleospin Tissue kit. gRNA targeting sites were amplified with specific primers that are listed in **Table 6**. PCR clean up (MN, PCR clean up and gel extraction kit) was performed. 400 ng from cleaned PCR products were mixed with NEB 2 buffer and shuffled via heteroduplex formation protocol (5 minutes at 95°C and ramp down to 85°C at -2°C/sec and ramp down to 25°C at -0.1°C/sec). After heteroduplex formation, samples were

treated with T7 endonuclease (NEB) for 1-2 hours at 37°C. Control samples were not treated with T7 endonuclease. Cut samples were immediately loaded on 2% agarose gel. Instead of loading buffers, 5% glycerol was added and gels were visualized via Gel Doc XR System (Bio-Rad).

**Table 6. List of T7 specific Primers**

T7 Primer Name	5' to 3' Sequence
sgAF10-1 T7 fwd	CAACTCCCTCTTAGATGGTCTC
sgAF10-1 T7 rev	GCGGAATCACATGACAGTCC
sgAF10-2 T7 fwd	GTGACAGGTGGATTAATAGGGCT
sgAF10-2 T7 rev	TCTGAAATAAGGTAACCACCCAAGT
gMLL1-1 T7 fwd	TTGGGGCTGTATGTTTCTGC
gMLL1-1 T7 rev	ATGCCCCAAGTAGTTCCCAG
gMLL1-2 T7 fwd	GATCCTCTTGTCCCAGCCTC
gMLL1-2 T7 rev	ACACAGTCTGACAGCTCTCC
gMLL1-3 T7 fwd	CCGCATGGATCACTTTACCTC
gMLL1-3 T7 rev	ACCCTTCTTCTGAAACACAAAGC

### 2.1.5 AF10 overexpression plasmids cloning

pBp-AF10 overexpression plasmid is a gift from Or Gozani Lab<sup>130</sup>. However, this plasmid did not overexpress AF10 to sufficient levels upon retroviral packaging and infection of dH1f cells (data not shown). Therefore, AF10

cDNA was cloned from pBp backbone into lentiviral, CMV promoter driven expression plasmid. Also, Hygromycin selectable plasmid backbone was used, because sgAF10 plasmids have puromycin selection and for rescue experiments, another antibiotic selection was needed.

For this purpose, AF10 cDNA was amplified with AF10 cloning primers (**Table 3**). For PCR amplification, Phusion polymerase was used with conditions depicted in **Table 7**.

**Table 7. PCR cycling conditions of AF10 cloning experiment**

Temperature (°C)	Period (min:sec)	Cycle number
98°C	00:30	1
98°C	00:10	30
53°C	00:20	
72°C	01:00	
72°C	10:00	1
4°C	∞	

PCR product was digested with XhoI (NEB) enzyme for 3 hours at 37°C. pENTR1A no ccDB (Addgene # 17398) was cut with XhoI (NEB) enzyme for 2 hours at 37°C and treated with AP enzyme 30 min at 37°C. Digested DNA samples were run on an agarose gel and expected bands were excised from gel and purified (Gel extraction kit, MN). Gel extracted DNAs were ligated (T4 DNA Ligase, NEB) into pENTR1A backbone. Ligation product was transformed into Stbl3 competent bacteria. The resulting plasmid was confirmed with sequencing of the entire AF10 sequence with primers in

**Table 8.** In the last step, AF10 sequence cloned into an expression plasmid pLenti CMV/TO Hygro DEST (Addgene # 17291) via LR cloning (Invitrogen).

### 2.1.6 Cloning of GFP Plasmids

Green Fluorescent Protein (GFP) cloned into different backbone plasmids to serve as a control in different experiments. For this purpose, GFP coding sequence was firstly cloned into pENTR1A vector. pBp GFP puro plasmid was cut with EcoRI (NEB) enzyme to excise eGFP. pENTR1A no ccDB (Addgene # 17398) was cut with EcoRI (NEB) enzyme and treated with AP enzyme for 30 min at 37°C. Digested pENTR1A was gel purified (Gel extraction kit, MN) and ligated (Quick Ligase, NEB) with eGFP insert. Ligation products were transformed into Stbl3 competent bacteria. The resulting plasmid sequence was confirmed by sequencing with pENTR1A sequencing primer (**Table 8**). In the last step, eGFP sequence was recombined into variety of destination plasmids including, pLenti CMV/TO Hygro DEST (Addgene # 17291) and pLEX-307 (Addgene # 41392) via LR cloning (Invitrogen).

### 2.1.7 Sequencing of samples

All sequencing samples were analyzed in MacroGen Europe Laboratories (Netherlands) by Sanger sequencing method (EZ-Seq). Samples were shipped

by mixing 500 ng plasmid sample and 25 pmole sequencing primers that are listed in **Table 8**. Shipping was performed according to UN3373 regulations for Category B. Sequencing results were stored as .ab1, .txt and pdf file formats.

**Table 8. List of sequencing primers**

Primer Name	5' to 3' Sequence
pENTR1A fwd seq	CTACAAACTCTTCCTGTTAGTTAG
pENTR1A rev seq	ATGGCTCATAACACCCCTTG
shRNA sequencing (MSCV) fwd	CCCTTGAACCTCCTCGTTCGACCT
U6 promoter fwd seq	ACTATCATATGCTTACCGTAAC
DOT1L sequencing rev	CGGGATTTCTTCACAGACCCA
DOT1L sequencing fwd-1	CACGATGCTGCTCATGAAAT
DOT1L sequencing fwd-2	AATTTTGCCTTTGGTCCTGA
DOT1L sequencing fwd-3	GATGCCTACAGATCCCCTCA
DOT1L sequencing fwd-4	CTGCAGCTCAAGTCCTGTGT
DOT1L sequencing fwd-5	CAGTGAGAAGGGCCTGAGAG
DOT1L sequencing fwd-6	ATTCCGGCTTCTCAGATCCT
DOT1L sequencing fwd-7	CAACCTCAACTCCATGGTCA
DOT1L sequencing fwd-8	GTGCTTCTCTTCCCCACAAG
DOT1L sequencing fwd-9	CAGTCGCTGTTTCAGCTCTGT
AF10 sequencing rev	GGGACAAAGTTCACATCTCACTC
AF10 sequencing fwd-1	GCGTCGCGGTGCATCAAG
AF10 sequencing fwd-2	AGCTGAAAAAGAGCAAACGGG
AF10 sequencing fwd-3	CTGGCAGACCCAAAGGAAACA
AF10 sequencing fwd-4	AGAGGCAGTGGAGTGAAGGA

## 2.2 Cell Culture

Cells are grown at 37°C with 5% CO<sub>2</sub>. HEK293T cells, mouse embryonic fibroblasts (MEF) and dH1f<sup>132</sup> cells were grown in D10 medium composed of 1X DMEM (Gibco) and 10% fetal bovine serum (FBS, Gibco) and 1% Penicillin/Streptomycin (Pen/Strep, Gibco). Prior to MEF seeding, plates were incubated with 0.1% Gelatin (Sigma) solution for 10 min at room temperature (RT). Induced pluripotent stem cells (iPSCs) were grown in human ES (hES) medium composed of DMEM/F12 (Stem Cell Technologies), 20% knock-out serum replacement (KOSR, Gibco), 10 ng/ml bFGF, 0.1 mM β-mercaptoethanol, 1% Non-essential amino acids (neaa), 1% Pen/Strep. For iPSC culture, MEFs were seeded at least one day before. Cells were observed with Nikon Eclipse TS100 inverted microscope and fluorescence was detected with Nikon C-SHG Mercury lamp.

### 2.2.1 Cell dissociation, freezing and thawing procedures

HEK293T cells, dH1fs and MEFs were washed with DPBS solution (Gibco) and dissociated with 0.05% Trypsin-EDTA (Gibco) at 37°C for 4 min. Trypsin was inactivated with D10 medium and cells were passaged into new plate. To freeze cells, dissociated cells were collected in a tube and centrifuged for 5 min at 1500 rpm. Freezing medium was prepared by mixing 10% DMSO (Sigma Aldrich) with 90% FBS and filtering with 0.20 μm regenerated cellulose (RC) syringe filters (Corning). After centrifugation, supernatant was removed and cell pellet was dissolved within freezing



medium. Then, cells were transferred into externally threaded cryovials (Corning) and stored in isopropanol filled Mr. Frosty freezing containers (Thermo Scientific) at  $-80^{\circ}\text{C}$  for 1-2 days. After cells were completely frozen, cryovials were transferred into a liquid nitrogen tank for long-term storage. For thawing, cryovials were quickly defrosted in  $37^{\circ}\text{C}$  water bath until half melted then transferred immediately into growth medium. The mixture was centrifuged for 5 min at 1500 rpm for quick removal of DMSO. After centrifugation, supernatant was removed and cell pellet was dissolved within fresh growth medium.

### **2.2.2 Generation of Mitomycin-c treated MEFs**

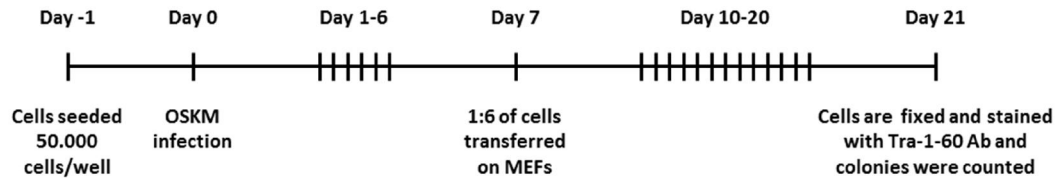
MEFs were obtained from pregnant mice at day 13 post coitum. Each embryo was separated from its placenta and surrounding membranes. Brain and dark red colored organs were removed and washed with PBS. Rest of the embryo was minced with razor blades within PBS. Minced tissue was incubated with 1-2 ml Trypsin per embryo on gentle shake at  $37^{\circ}\text{C}$  for 15 min. At the end of incubation, suspended cells were transferred into falcon tube and waited a few minutes for large pieces to settle down. Then, supernatant was carefully transferred to a falcon tube and centrifuged for 5 min at 1100 rpm. Pellet of cells were resuspended within fresh D10 medium and cells were plated as 1 embryo into 10 cm tissue plate. Fibroblasts attached to the plate and cells were passaged for 4 times and resulting cells were frozen. This procedure was performed by Dr. Tamer Önder.

Previously frozen MEFs were thawed to make a Mitomycin-c treatment. MEFs were grown in tissue culture plates that were incubated with 0.1% gelatin (Sigma) solution for 10 min. Cells were passaged within 1-2 days and waited for them to reach confluency. Prior to Mitomycin-c treatment, medium was replaced with fresh D10 medium (10 ml medium for 15 cm plates). Mitomycin-c was added at a final concentration of 10  $\mu\text{g}/\text{ml}$  into medium (Millipore 475820-10mg - dissolved in 10 ml  $\text{dH}_2\text{O}$  and filtered). Cells were incubated with Mitomycin-c for 2 hours at 37°C. Medium was removed and cells were washed with PBS, twice. Then, cells were trypsinized and frozen.

### **2.2.3 Generation of iPSCs from dH1fs (Reprogramming assays)**

For reprogramming experiments, dH1fs were counted and 50.000 cells were seeded per well into 12-well plates. Every experiment was performed in triplicate wells. Next day, dH1fs were transduced with 200-250  $\mu\text{l}$  O2S and K2M viruses and 8  $\mu\text{g}/\text{ml}$  protamine sulfate (Sigma Aldrich) in 700  $\mu\text{l}$  total volume. After overnight incubation, medium was replaced with fresh D10 medium and every two days thereafter. On post-infection day 6 or day 7, cells were trypsinized and transferred onto wells containing Mitomycin-c treated MEFs (75.000 cells/well of 12-well tissue plate) at a ratio of 1/6. Next day, medium was switched to hES medium and replaced every other day till day 14 of reprogramming. Between days 14 and 21, medium was

replaced everyday with fresh hES medium. At the end of reprogramming experiments, cells were fixed and stained with Tra-1-60 antibody (**Figure 7**).



**Figure 7. Timeline of reprogramming experiment**

In experiments where the small molecule inhibitor of DOT1L (iDOT1L), EPZ004777 (Tocris Bioscience), was used, the final concentration of the compound in cell culture medium was 3  $\mu$ M. Cells were treated with iDOT1L or DMSO as control on days 1, 3 and 5.

### 2.2.4 Tra-1-60 staining of iPSC colonies

At day 21 of reprogramming experiments, cells were stained with Tra-1-60 antibody to count the iPSC colonies. Cells in 12-well plates were washed with 1X PBS and fixed with 500  $\mu$ l 4% paraformaldehyde (PFA) solution. Fixation was performed at room temperature for 20 min on gentle shaking. Then, PFA solution was removed and fixed cells were washed with 1X PBS solution for 3-4 times. Then, cells were incubated with 300  $\mu$ l biotin conjugated Tra-1-60 antibody (BioLegend #330604) which was diluted 1:200 within staining solution (3% FBS, 0.3% Triton X in PBS). Primary antibody incubation was

performed overnight at cold room on gentle shaking. Next day, cells were washed 5 times with 1X PBS and incubated with secondary antibody solution: Streptavidin-HRP (BioLegend #405210) that was diluted within staining solution 1:500. Incubation was performed for 2 hours at room temperature on gentle shaking. Cells were washed 5 times with 1X PBS and colonies were visualized with DAB solution (1% DAB -3,3 diaminobenzidine-, 1% Nickel Ammonium Sulfate, 0,3% H<sub>2</sub>O<sub>2</sub>). Staining was performed at room temperature for 20 min on gentle shaking. Then, DAB solution was removed and cells were washed with 1X PBS solution for 2 times. Plates were developed for 1 day at room temperature before imaging. Then, PBS was decanted and wells were filled with cream (SEK or Tikveşli) to make a background on DAB stained blackish-brownish iPSC colonies. Plates were scanned to digitalize the colony images. Scanned plate images were quantified well by well via ImageJ software. While quantifying the wells, threshold for intensity was adjusted to 150 and colonies that were bigger than 20 pixels were counted as an iPSC colony.

### **2.2.5 Generation of DOT1L-KO single cell clone**

HEK293T cells were transfected with either non-targeting (gNT2) or guideDOT1L (gDOT1L) containing lenticrisprV2 plasmids and transfected cells were selected with puromycin (2 µg/ml) for 2-3 days. After selection, cells were trypsinized, diluted to a single cell suspension and seeded onto 96-well plates. Probability of events for a Poisson distribution was used to calculate the optimal cell number per plate. Accordingly, 26 cells were placed

in 100-wells. For both NT and gDOT1L samples 3 plates were prepared and their growth medium was changed within 2-3 days intervals. At the end for 1 week of growth, single cell clones were identified and transferred to 48-well plates.

As expected, not all wells had a single cell clones. Many wells had no colonies at all while 30 of them had single colonies and in 7 wells there were multiple colonies. Selected single clones were passaged and half of the each colony was deposited as frozen stock. Other half of single cell clones was expanded for histone acid extraction to check the H3K79me2 levels. Proliferated single cell clones were pelleted and histone acid extraction was performed from 19 colonies. All single clones were tested for their H3K79me2 levels via immunoblotting. As a result of this experiment single clone #10 was picked as a DOT1L-KO HEK293T cell line for following experiments.

### **2.2.6 Transient Transfection**

Transfection of HEK293T cells were performed with Fugene (Promega) transfection reagent.  $2.5 \times 10^6$  cells were counted into each 10 cm cell culture plate. Transfection was performed the following day after cells were seeded. For transfection, 500  $\mu$ l DMEM was mixed with 20  $\mu$ l Fugene and incubated at RT for 5 minutes. 5  $\mu$ g plasmid was added to DMEM/Fugene mixture and incubated at RT for 30 min. At the end of incubation, mixture was added dropwise onto HEK293T cells. Next day, medium was replaced with fresh D10 medium. At this point, fluorescent protein expression was detected with fluorescence microscope (Nikon) to determine the transfection efficiency. For

this purpose, following plasmids were used: pLenti CMV/TO-GFP-Hygro, pLenti PGK GFP puro (Addgene # 19070), pBp GFP puro, pLEX-307\_GFP and RRL\_GFP.

### 2.2.7 Virus production

Viruses were produced using HEK293T cells. Viruses can be produced in different scales from 6-cm cell culture plates to 15-cm plates. In this section, 10 cm cell culture plates will be explained but quantifications for different scales are depicted in **Table 9**.  $2.5 \times 10^6$  HEK293T cells were counted for 10 cm cell culture plates. Transfection of viral plasmids was performed the following day. For transfection, 250  $\mu$ l DMEM was mixed with 20  $\mu$ l Fugene and incubated at RT for 5 min. In another tube, 250  $\mu$ l 1xDMEM was mixed with 2.5  $\mu$ g viral transfer vector and 2.5  $\mu$ g viral packaging plasmids (2250ng PUMVC for retroviruses or pCMV-dR8.2  $\Delta$ VPR for lentiviruses and 250 ng pCMV-VSV-G). The two mixtures were combined and incubated at RT for 30 min. At the end of incubation, mixture was added dropwise onto HEK293T cells. Cells were grown with viral transfection mixture overnight and next day, medium of cells were changed with 8 ml of fresh D10.

48 hours after transfection, medium was collected in a falcon tube and stored at 4°C. 8 ml of fresh D10 medium was added on HEK293T cells for one more batch of virus production. 72 hours after transfection, medium was collected again and mixed with primary collected medium in falcon tube. All collected viral supernatants (16 ml in total) were centrifuged at 1500 rpm for 5 min to avoid any detached HEK293T cells and then filtered through a 0.45  $\mu$ m

syringe filters (Corning). Filtered viral medium was aliquoted and stored at  $-80^{\circ}\text{C}$  for long-term usage. To concentrate the viral supernatants, 50% PEG-8000 (Sigma) was dissolved (w/v) in 1xPBS (Gibco) and used as 5X. Filtered viral medium was mixed 1X PEG solution and incubated at  $4^{\circ}\text{C}$  for at least overnight or at most 2 or 3 days. At the end of PEG precipitation, tubes were centrifuged at  $4^{\circ}\text{C}$  at 2500 rpm for 20 min and most of the supernatant was decanted into 10% bleach containing bottle. Rest of the precipitate was centrifuged again at  $4^{\circ}\text{C}$  at 2000 rpm for 5 min. Viral precipitate was resuspended with 160  $\mu\text{l}$  cold 1X DPBS and aliquoted to store at  $-80^{\circ}\text{C}$  for long-term usage.

**Table 9. Table of viral transfection information for different scales**

Scale	6-cm plate	10-cm plate	15-cm plate
HEK293T seeding amount (cells)	$1 \times 10^6$	$2.5 \times 10^6$	$6.5 \times 10^6$
1X DMEM ( $\mu\text{l}$ )	100	250	500
FuGene ( $\mu\text{l}$ )	8	20	60
Amount of plasmid (ng)	1000	2500	7500
Amount of pUMVC or 8.2- $\Delta\text{vpr}$ (ng)	900	2250	6750
Amount of pCMV-VSV-G (ng)	100	250	750
Collection medium amount (ml)	3+3=6	8+8=16	20+20=40
Addition of 5X PEG amount (ml)	1.5	4	10
Amount of PBS to resuspend ( $\mu\text{l}$ )	60	160	400

### **2.2.8 Transduction of cells**

Viral transductions were performed with the addition of 8  $\mu\text{g}/\text{ml}$  protamine sulfate. Cells were incubated with transduction mixture overnight. This transduction step was repeated one more time to increase infection efficiency. After 48 hours of transduction, cells were observed for their fluorescent emission or selected with antibiotics. For HEK293T cells, 2  $\mu\text{g}/\text{ml}$  puromycin was added for 2-3 days and 300  $\mu\text{g}/\text{ml}$  Hygromycin was used for 4-5 days to complete antibiotic selection. For dH1f cells, 1  $\mu\text{g}/\text{ml}$  puromycin was added for 2-3 days and 200  $\mu\text{g}/\text{ml}$  Hygromycin was used for 4-5 days to complete antibiotic selection.

### **2.3 Western Blots**

Three different protein isolation methods were used in this thesis: whole cell lysis, cytosolic-nuclear fractionation and histone acid extraction. For all these methods, cells were prepared similarly: Cells were trypsinized and pelleted by centrifugation at 1500 rpm for 5 min. Supernatants were removed and pellets were washed with 1X PBS. Pellets were immediately frozen and stored at  $-80^{\circ}\text{C}$ . For protein extraction, cells were thawed on ice with lysis buffer.



**Table 10. Recipes of protein lysis buffers**

WHOLE CELL LYSIS BUFFER	CYTOSOLIC LYSIS BUFFER	NUCLEAR LYSIS BUFFER	TRITON EXTRACTION BUFFER
50mM Tris pH 8.0	10mM HEPES pH 7.9	20mM HEPES pH 7.9	0.5% Triton X 100
250mM NaCl	10mM KCl	0.4M NaCl	2mM PMSF
5mM EDTA	0.1mM EDTA	1mM EDTA	0.02% (w/v) NaN <sub>3</sub>
1% NP-40	0.4% NP-40	10% Glycerol	
Protease Inhibitor Cocktail	Protease Inhibitor Cocktail	Protease Inhibitor Cocktail	
dH2O	dH2O	dH2O	PBS

### 2.3.1 Whole cell lysis method

Whole cell lysis buffer was prepared with recipe in **Table 10** with the addition of cOmplete ULTRA protease inhibitor Tablets (Roche). Cell pellets were resuspended with whole cell lysis buffer and incubated for 45 min on ice with gentle shaking. At the end of incubation, tubes were centrifuged at 4°C for 10 min at 14000 rpm. Supernatant was removed into new tube and used as a whole cell lysis protein. Protein concentrations were determined via BCA assay (Thermo Scientific).

### 2.3.2 Nuclear protein extraction

Cytosolic and nuclear lysis buffers were prepared with recipe in **Table 10** with the addition of cOmplete ULTRA protease inhibitor Tablets (Roche). Cell pellets were resuspended with cytosolic lysis buffer and incubated for 15 min on ice on 50 rpm shaking plate. At the end of incubation, tubes were centrifuged at 4°C for 3 min at 3000g. Supernatant was removed into new

tube and centrifuged again at 4°C for 5 min at 3000*g*. This supernatant was reserved as the cytosolic protein fraction. The pellet was washed with half volume of cytosolic lysis buffer and centrifuged at 4°C for 3 min at 3000*g*. Supernatant was discarded and the pellet was resuspended in nuclear lysis buffer and sonicated 2 times for 10 seconds at 40 amplitude with a 10 second interval in between (QSONICA Q700 with microtip). After sonication, tubes were centrifuged at 4°C for 5 min at 15000*g*. Supernatant was removed into new tubes as a nuclear protein fraction. Both cytosolic and nuclear protein concentrations were determined via BCA assay (Thermo Scientific).

### **2.3.3 Histone acid extraction**

Cell pellets were resuspended with triton extraction buffer (**Table 10**) and incubated for 10 min on ice on 50 rpm shaking plate. At the end of incubation, the tubes were centrifuged at 4°C for 10 min at 2000rpm. Supernatant was discarded and the pellet was washed with half volume of triton extraction buffer and centrifuged at 4°C for 10 min at 2000rpm. Supernatant was discarded and the pellet was resuspended in 0.2N HCl. Tubes were incubated at 4°C for overnight on rotating wheel at 10 rpm. At the end of incubation, tubes were centrifuged at 4°C for 10 min at 2000rpm and supernatant was collected in new tube. Acid extractions were neutralized with the addition 0.1M NaOH for 1/5 volume of HCl solution. Protein concentrations were determined via BCA assay (Thermo Scientific).

### 2.3.4 Western Blotting

Equal amounts of proteins were boiled for 10 min with loading buffer (4X Laemmli sample buffer, Bio-Rad) and loaded onto 4–15% Mini-PROTEAN TGX Precast Protein Gels (Bio-Rad). Gels were run with TGS buffer (diluted from 10X stock, Bio-Rad). Precision Plus Protein Dual Color Standards (Bio-Rad) were used as a molecular weight ladder. Proteins were transferred onto Immun-Blot PVDF Membrane (Bio-Rad) via semi-dry or wet transfer technique.

For semi-dry transfer method, Trans-Blot Turbo Transfer System (Bio-Rad) was used. As a transfer buffer, 10% Ethanol added TGS buffer (Bio-Rad) was used with standard transfer protocol. Semi-dry transfer method was preferred for all the western blot methods unless the detected protein's size was more than 100 kDa. For larger proteins, wet transfer method was used (Mini Trans-Blot Electrophoretic Transfer Cell -Bio-Rad). As a transfer buffer, Towbin Buffer (25 mM Tris, 192 mM Glycine, 20% methanol (v/v) -pH 8.3) was used with 15 volt for overnight incubation in cold room.

After transfer of proteins on membrane, membrane was incubated with 5% blotting grade blocker (Bio-Rad) dissolved in TBS-T (20 mM Tris, 150 mM NaCl, 0.1% Tween 20 -pH 7.6) for 1-2 hours at room-temperature with gentle stirring. However, membranes to be incubated with Streptavidin-HRP antibody were blocked with 2% bovine serum albumin (BSA, Sigma) in TBS-T. After blocking step, membranes were incubated with primary antibody solution (1:200-1:1000 primary antibody in 2% BSA, 0.02% NaN<sub>3</sub> in TBS-T) at 4°C for 16 hours. Concentration of primary antibodies is depicted in **Table**

11. After primary antibody incubation, membranes were washed with TBS-T solution for 3 times with 15 min intervals on 50 rpm shaker at room temperature and then incubated with secondary antibody solution (1:5000 secondary antibody in 5% blotting grade blocker in TBS-T) at room temperature for 1-2 hours. Secondary antibodies were depicted in **Table 11** . Streptavidin-HRP blotted membranes were not incubated with a secondary antibody. After secondary antibody incubation, membranes were washed with TBS-T solution for 3 times with 15 min intervals on 50 rpm shaker at room temperature. Then, proteins were visualized with Pierce ECL Western Blotting Substrate (Thermo Scientific) and Odyssey Fc Imaging systems (LiCor).

**Table 11. Table of antibody information that were used for western blots**

Antibody Name	Catalog number	Concentration
Streptavidin-HRP	405210	1:10,000
H3K79me2	ab3594	1:1000
H3 total	ab1791	1:1000
H3K4me1	ab8895	1:1000
H3K4me3	ab8580	1:1000

## 2.4 Pull down experiments

HEK293T cells were infected with BirA\*-DOT1L wt or mut concentrated viruses and selected with puromycin (2  $\mu\text{g}/\text{ml}$ ) for 2-3 days. After selection, cells were treated with 50  $\mu\text{M}$  D-Biotin (Sigma, 47868) for 24 hours and cells were collected for protein isolation. Proteins were obtained via nuclear fractionation method. As a control, uninfected HEK293T cells were treated similarly. Pull-down was performed with Streptavidin beads (Thermo Scientific, 53117) as previously described<sup>133</sup>. Briefly, equal amount of nuclear fraction was incubated with Streptavidin beads at 4°C for 16 hours on rotating wheel at 10 rpm. Then supernatants were collected and beads were washed twice in 2% SDS; once with wash buffer 1 (0.2% deoxycholate, 1% Triton X, 500 mM NaCl, 1 mM EDTA, 50 mM HEPES, pH 7.5), once with wash buffer 2 (250 mM LiCl, 0.5% NP-40, 0.5% deoxycholate, 1% Triton X, 500 mM NaCl, 1 mM EDTA, 10 mM Tris, pH 8.1) and twice with wash buffer 3 (50 mM Tris, pH 7.4, and 50 mM NaCl). Eluted proteins were analyzed with Streptavidin-HRP antibodies to observe the efficiency of pull-down.

## 2.5 Mass-Spectrometry Analysis

For mass spectrometry analysis, control (uninfected) and BirA\*-DOT1L wt or mut infected HEK293T cells were treated with 50  $\mu\text{M}$  D-biotin for 24 hours and harvested. Following nuclear protein isolation, biotinylated proteins were pulled down with Streptavidin beads (Thermo Scientific, 53117). For mass-spectrometry analysis, beads were washed and bound

proteins were digested with on-bead tryptic proteolysis method as previously described<sup>134</sup>. Briefly, beads were washed (8 M urea in 0.1 M Tris-HCl, pH 8.5) and reduction and alkylation steps performed. Then beads were washed again with 50 mM ammonium bicarbonate and incubated with trypsin in 50 mM ammonium bicarbonate at 37 °C overnight. Beads were pulled-down at 1000g for 5 min and peptides were collected. Beads were rinsed with 50 mM ammonium and the supernatant was added to previously collected peptides. Peptides were acidificated and desalted. Then, all three samples were analyzed with reversed-phase nLC (NanoLC-II, Thermo Scientific) combined with orbitrap mass spectrometer (Q Exactive Orbitrap, Thermo Scientific) with data acquisition and processing steps that were as previously described<sup>134</sup>. Each sample was run for twice. On bead tryptic digestion of biotinylated proteins and their LC-MS/MS analysis was performed by Nazlı Ezgi Özkan Küçük in Nurhan Özlü Lab.

To discriminate the DOT1L-specific biotinylation, proteins detected in HEK293T control samples were subtracted from BioID samples. Rest of the proteins was selected only if they exist in both runs of mass-spectrometry. Among these common proteins, nuclear localized ones are determined via GO annotation (<http://www.geneontology.org/>) via cellular compartment analysis. UniProt protein names were converted via ID mapping tool (<https://www.uniprot.org/uploadlists/>). Determined proteins were sorted with their coverage percentage and then with PSM (peptide spectrum matches) numbers.

## 2.6 RNA isolation, cDNA synthesis and qPCR

Cells were trypsinized and pelleted by centrifugation at 1500 rpm for 5 min. Supernatant was removed and pellet was washed with 1X PBS. Resulting pellet was immediately frozen and stored at  $-80^{\circ}\text{C}$ . RNA isolation kit (MN) was used to isolate total RNA. Concentrations of RNA samples were determined with Nanodrop 2000 (Thermo scientific). For cDNA synthesis, 1  $\mu\text{g}$  RNA solution was mixed with 200  $\mu\text{M}$  dNTP (Thermo Scientific) and 4  $\mu\text{M}$  random hexamer (invitrogen) with  $\text{dH}_2\text{O}$  up to 16.5  $\mu\text{l}$  total volume. This mixture was incubated at  $65^{\circ}\text{C}$  for 5 min and then quickly chilled on ice. 5X first strand buffer (invitrogen), 8 mM DTT (invitrogen) and 20 U Rnasin (promega) was added to the mixture for up to 24  $\mu\text{l}$  total volume and incubated for 10 min at room temperature. Then, 1  $\mu\text{l}$  of M-MLV reverse transcriptase enzyme (200 U, Invitrogen) was added to the reaction and incubated at  $37^{\circ}\text{C}$  for 1 hour. Reaction was ended with inactivation at  $70^{\circ}\text{C}$  for 15 min. cDNA solutions were diluted with 75  $\mu\text{l}$  nuclease free water (NEB). From diluted cDNA mixture, 2  $\mu\text{l}$  sample was used for 1 reaction of quantitative real-time PCR (qPCR) and the rest of the cDNA was stored at  $-20^{\circ}\text{C}$  for long term.

For qPCR, 2  $\mu\text{l}$  of cDNA sample was mixed with 2.5  $\mu\text{M}$  forward and reverse primers, 10  $\mu\text{l}$  LightCycler 480 SYBR Green I Master (2X, Roche) and  $\text{dH}_2\text{O}$  up to 20  $\mu\text{l}$  total volume. Forward and reverse primers are listed in **Table 13**. Every sample was prepared in duplicates and loaded into 96-well opaque plates (Roche). Reaction was run in LightCycler 480 Instrument II (Roche) with conditions depicted in **Table 12**. For every sample, endogenous  $\beta$ -actin

levels were used as controls. Expression values were calculated by the formula “ $2^{-(C_t-C_c)}$ ”, where  $C_t$  and  $C_c$  are the average of threshold cycles after normalization to  $\beta$ -actin. The relative quantification value for a target gene was compared to the control sample.

**Table 12. Thermal cycling conditions for Q-PCR**

Temperature (°C)	Period (min:sec)	Cycle number
95°C	03:00	1
95°C	00:10	40
60°C	00:30	
72°C	00:30	
72°C	05:00	1
4°C	$\infty$	



Table 13. List of qRT-PCR Primers

qRT-PCR Primer Name	5' to 3' Sequence
RT Primer AF10 forward	GCGTCGCGGTGCATCAAG
RT Primer AF10 reverse	GGGACAAAGTTCACATCTCACTC
RT Primer NONO forward	CATCAAGGAGGCTCGTGAGAAG
RT Primer NONO reverse	TGGTTGTGCAGCTCTTCCATCC
RT Primer KAISO forward	TAGCAGAGCTTGGTGTCCCATTG
RT Primer KAISO reverse	CACCAGAATCAGGAGGTAAAGGC
RT Primer SIN3B forward	TCTGAGGACTCCACGTTTCGTCA
RT Primer SIN3B reverse	AGGTTCTGCTCCAGGACAACGT
RT Primer AF17 forward	CTGCGTATGTTTCGGACGGAGAG
RT Primer AF17 reverse	ACCTGAACGATGCCATAGCAA
RT Primer MRE11 forward	CAGCAACCAACAAAGGAAGAGGC
RT Primer MRE11 reverse	GAGTTCCTGCTACGGGTAGAAG
RT Primer ENL forward	GGTGAGGTTAGAGCTGGGG
RT Primer ENL reverse	TGGATGTCACATTGCTCGGG
RT Primer NUMA1 forward	GGTCCAGGAAGAGAGGCAGAA
RT Primer NUMA1 reverse	CTTGCTGGCTTGGTCAGAGTCA
RT Primer TPR forward	GCTCAGGTTGAGAGTCTGCGTT
RT Primer TPR reverse	CAGTTCCTCATGCTGAGCCATTG
RT Primer DDX21 forward	TGCACGTGGGTTAGACATCC
RT Primer DDX21 reverse	CGCCCGGATCGATGAATGTA
RT Primer AF4 forward	TGCATTGCAAGCACAGGCAC
RT Primer AF4 reverse	AAGGTCAAAGGCGGTAAGAACAT
RT Primer AF9 forward	TGCAGCAGATCGTGAACCTT
RT Primer AF9 reverse	ACTGTGGTTTTGTCCAGCGA
RT Primer BAT3 forward	CAGTGGTATGCCTGCCAAGA
RT Primer BAT3 reverse	AGCTCTCCTGAACCTCTGGT
RT Primer SIRT1 forward	ACAGGTTGCGGGAATCCAAA
RT Primer SIRT1 reverse	G TTCATCAGCTGGGCACCTA
RT Primer STAT1 forward	ACTCCAGGCCAAAGGAAGC
RT Primer STAT1 reverse	GACATGGGGAGCAGGTTGTC
RT Primer HNRNPM forward	TGGACGCTGAAGGAAAGTCA
RT Primer HNRNPM reverse	CATACCCATCCCACCAGTCG
RT Primer NPM1 forward	CGGTTGTGAACTAAAGCCG
RT Primer NPM1 reverse	TTTGCACCAGCCCCTAAACT
RT Primer CDK9 forward	TGCACGTGGGTTAGACATCC
RT Primer CDK9 reverse	CGCCCGGATCGATGAATGTA

RT Primer MLL1 forward	AAGCGGAAGGTGAAGGACAG
RT Primer MLL1 reverse	GGTCGGACCAGAAGAAGTCG
RT Primer WDR5 forward	AATTCAGCCCGAATGGAGAGT
RT Primer WDR5 reverse	AGGCTACATCGGATATTCCCAG
RT Primer RBBP5 forward	CATCTTTTGATAGGCGAGGGG
RT Primer RBBP5 reverse	GTTCCAGTTGTCACTCTGAAGG
RT Primer ASH2L forward	AGAATGGCCGACAGTTGGG
RT Primer ASH2L reverse	CCTTCAAGTTTGCTTGCTTCC
RT Primer DPY30 forward	GGAGGGACAAACGCAGGTT
RT Primer DPY30 reverse	GGTAGGCACGAGTTGGCAA
RT Primer MEN1 forward	GCCTGGGTAGTGTTTGGGC
RT Primer MEN1 reverse	AGCGCATGTATGATCCTTTCAG
RT Primer WDR82 forward	TTTCCTGGACATAGCAAAAGGG
RT Primer WDR82 reverse	TCCCAGAGTCGAATGGTCTTAT
RT Primer HCFC1 forward	GCAATGACCTCTACGAACTCC
RT Primer HCFC1 reverse	ACCTTGGAATGTTGTTCTTTGGG
RT Primer MLL2 forward	CTCTAAGATGTTGGTTTGCAGAGA
RT Primer MLL2 reverse	GCCTTGCCTTCCAAGAGTGA
RT Primer MLL3 forward	GGACAGAGAAAAGAACGATCTCC
RT Primer MLL3 reverse	GGCATACTCCTAGTG
RT Primer MLL4 forward	ACCCCGGCGATTTATGGATG
RT Primer MLL4 reverse	CTTCTCAGGGAGTGGAAGTGG
RT Primer CXXC1 forward	GCAAACCGGACATCAACTGC
RT Primer CXXC1 reverse	GCACTCCCGACAGTACCAC
RT Primer SET1B forward	TGAGTTTGAGTCAAGCTCCGA
RT Primer SET1B reverse	ATGCCCAACGAGTCCACTG
RT Primer $\beta$ ACTIN forward	TGAAGTGTGACGTGGACATC
RT Primer $\beta$ ACTIN reverse	GGAGGAGCAATAGATCTTGAT
RT Primer NANOG forward	TGATTTGTGGCCTGAAGAAA
RT Primer NANOG reverse	TGGTGGTAGGAAGAGTAAAG

## 2.7 Microarray analysis

Differential gene expressions between pluripotent stem cells and fibroblast cells were computed by `affy` and `limma` packages from R. Samples of dH1f and BJ fibroblasts were compared to their respective iPSCs and embryonic stem cells from GEO data series GSE55679. Genes that had a log<sub>2</sub> fold change value of 3 or more in all fibroblasts compared pluripotent cells were categorized as the fibroblast related gene set. Genes that have a log<sub>2</sub> fold change value of -3 or less in all fibroblasts compared pluripotent cells were categorized as the pluripotency related gene set. This analysis was performed by Tunc Morova (Nathan Lack lab, Koc University).

## 2.8 RNA sequencing and analysis

RNA isolation was performed on previously harvested cell pellets (Day 0 and Day 6 samples) with Direct-zol kit (Zymo Research). Isolated RNAs were separated into tubes as; 1 µg for RNA-sequencing sample, 1 µg for cDNA synthesis and the rest was stored at -80°C as a long term backup. NEBNext Poly(A) mRNA Magnetic Isolation Module from NEBNext Ultra Directional RNA Library Prep Kit for Illumina was used to enrich mRNA from RNA-sequencing samples. Samples were then validated on a Tapestation (Agilent) to determine library size and quantification prior to paired-end (2 × 41 bp) sequencing on a NextSeq 500 (Illumina) platform. Reads were mapped to hg19 built-in genome by Hisat2 after assessing their quality by FastQC. DeSeq2 package was used to find differentially expressed genes between

samples. Genes were considered to be differentially regulated based on log2 fold change 0.5 and adjusted p-value 0.05. Gene Set Enrichment Analysis (GSEA) was performed on a pre-ranked gene list based on log2 fold change<sup>135</sup>. Pluripotency-related, fibroblast-related gene sets and Wang\_MLL\_Targets<sup>136</sup> was used as a control for MLL knockout.



## Chapter 3

### 3 Finding proximal-protein interactions of DOT1L via BioID method and shRNA-mediated screen of DOT1L-proximity interactors for reprogramming efficiency

#### 3.1 Introduction

##### 3.1.1 Biotinylation based detection of protein-protein interaction methods

Finding protein-protein interactions (PPIs) is crucial for understanding the mechanistic properties of the proteins of interest. Different methods have been developed to detect PPIs, such as co-immunoprecipitation (co-IP) to detect direct interactions or yeast two hybrid to investigate *in vivo* binding of two proteins. Even though these approaches are important methods to identify PPIs, it is challenging to identify a large set of interactors with these types of experiments. Biotinylation based methods have emerged to fill this gap. Biotinylation is a naturally occurring reaction that takes place in very low abundance when compared with methylation, acetylation or ubiquitination. Therefore, ectopic biotinylation can be detected via mass-spectrometry analysis and specific proteins can be determined via comparison with control background. Biotin ligase protein, originated from *E. coli* (BirA) was utilized for such assays<sup>137</sup>. BirA enzymes can biotinylate specific peptide

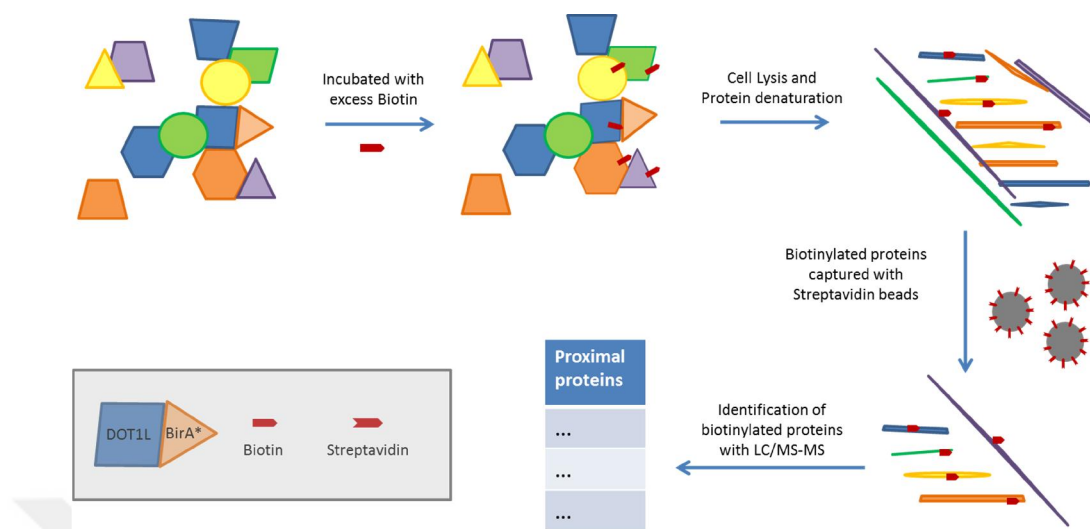
chain (GGGLNDIFEAQKIEWHE) from its lysine residue. This method is a powerful tool when there is no high quality antibody against the protein of interest or to detect a protein interaction profile at a specific time. Biotinylation-based methods are powerful since they can capture biotinylated proteins efficiently via Streptavidin pull-down. Streptavidin-biotin binding is very strong when it is compared with antibody-protein binding. Conventional antibodies bind to a protein with  $K_d$  values in nanomolar range whereas high affinity antibodies can bind with  $K_d$  values in picomolar range; in contrast, Streptavidin binds to biotin with  $K_d$  values in femtomolar range ( $K_d \sim 10^{-14}$  M)<sup>138</sup>. This method still depends on the strong binding of interacting proteins, since their detection is possible only if they are strongly bound to target protein. But this method is insufficient to detect transient interactions and functional interactions of a protein. To solve this problem, R118G mutation was generated in BirA enzyme and this mutant BirA (BirA\*) can promiscuously biotinylate proteins that bind in close proximity<sup>139</sup>. To identify binding partners of a particular protein, BirA\* can be fused, and introduced to cells so that all proximal proteins whether they are strongly interacting or transiently form a complex, will be biotinylated by BirA\*. This method was initially called BioID. Later, a smaller BirA\* enzyme was designed from *Aquifex aeolicus* with R40G mutation and referred as BioID2<sup>140</sup>. Both methods require excess amount of supplemented biotin, however BioID2 needs less<sup>141</sup>. There are various other methods that make use of biotinylation such as proximity labeling with ascorbate peroxidase (APEX) and in vivo proximal labeling (IPL). APEX can be utilized in live cells which BioID cannot<sup>142</sup>. Also, BioID cannot be applied for shorter periods of time because it requires 16-24 hours of biotinylation incubation. However, recently

proposed TurboID method only requires 10 minutes of biotinylation which is developed from a mutant biotin ligase<sup>143</sup>.

### **3.1.2 BioID is a powerful method to detect proximal protein interactions**

Proximity dependent biotin identification (BioID) method is a powerful tool to identify proximal protein interaction<sup>139</sup>. For BioID assay, fusion protein with BirA\* and protein of interest is cloned. Then, fusion protein expressing cells are incubated with excess biotin for biotinylation of proximal proteins. Later, cells are lysed and biotinylated proteins are pulled-down with streptavidin conjugated beads. Biotinylated proteins eluted from beads and identified via mass spectrometry analysis. These steps are summarized in **Figure 8**.

Previous studies showed that BioID method is an effective way of identification of proximal proteins and protein complexes<sup>133,139,144</sup>. BirA\* fusion with LaminA<sup>139</sup>, cytoskeletal protein bilobe<sup>144</sup> and centrosome proteins<sup>133</sup> showed that BioID can detect even transient protein interactions in diverse cellular compartments. BioID can be used for insoluble proteins and it can detect proteins that are in a very low abundance. Since candidate protein interactions are labeled with biotin, extremely stringent conditions can be applied during purification which can decrease contaminants.



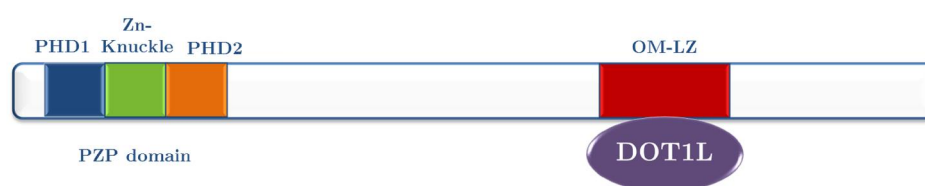
**Figure 8. Promiscuous biotin-ligase (BirA\*) selectively biotinylates the proximal proteins**

BioID also has some drawbacks that should be considered before designing any proximity interaction assay. To begin with, BioID approach relies on exogenous expression of fusion protein. Also, fusion proteins need to be generated and this will increase the size, and fusion may interfere with biological function of target protein. In BioID method, biotin is added to the lysines of proximal proteins and their biotinylation may hinder the biological modifications of that protein normally has. These limitations should be taken into account when proximity identification is planned.



### 3.1.3 AF10 is an important member of DOT1L-containing elongation complex

There are several elongation complexes reported to date such as EAP<sup>104</sup>, AEP<sup>107</sup>, SEC<sup>106</sup>, DotCom<sup>105</sup> and AF4-mediated complex<sup>103</sup>. Among these elongation complexes AF10 was identified within DotCom<sup>105</sup> and AF4-mediated complex<sup>103</sup>. DOT1L was also identified within these complexes. It was also known that DOT1L and AF10 interact through AF10's OM-LZ (octapeptide motif and leucine zipper) motif at C-terminus (**Figure 9**). It was hypothesized that AF10 recruits the elongation complex to the chromatin site since it has also a unmodified H3K27 reader domain, and this binding is abrogated by H3K27 modification<sup>145</sup>. The cross-talk between different epigenetic modifications can be explained with such cooperation within an elongation complex. Together with other protein components (AF9, AF17, ENL, p-TEFb) of the elongation complex, AF10 and DOT1L have a role in transcriptional regulation.

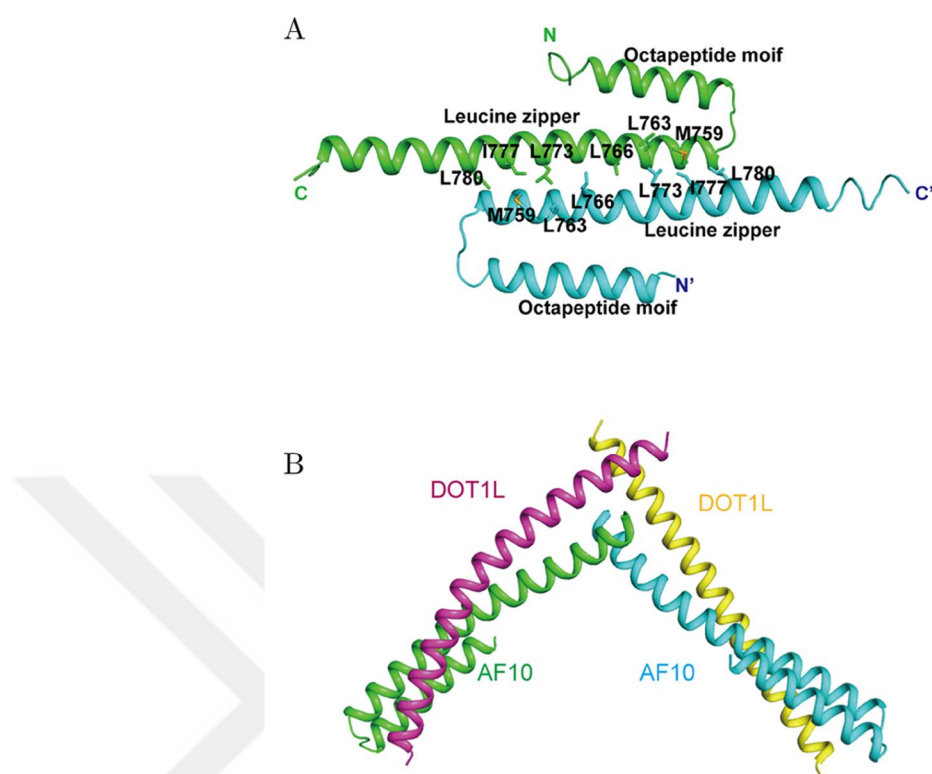


**Figure 9. Schematic representation of the AF10 protein.**

A PZP domain (PHD1-Zn-Knuckle-PHD2) is located at N-terminus of AF10, Octamer motif-leucine zipper (OM-LZ) motif interacts with DOT1L<sup>145</sup>.

In human acute myeloid leukemia (AML), MLL-AF10 fusion proteins can be generated as a result of MLL translocation  $t(10:11)(p12;q14)$ <sup>146</sup>. Main reason for transformation ability of MLL-AF10 fusion was claimed as AF10's interaction with DOT1L by recruiting DOT1L aberrantly to MLL-target sites<sup>146</sup>. In those cells, DOT1L activates genes that are supposed to be silent such as *Hoxa* cluster genes and *Meis1* through H3K79 methylation<sup>146</sup>. Aberrant expression of those genes causes transformation of MLL-AF10 fusion expressing cells. Therefore, many studies have suggested that a small molecule inhibitor that inhibits DOT1L-AF10 interaction can be used for leukemia therapy. Recently, the crystal structure of DOT1L-AF10 binding was reported (**Figure 10**)<sup>147</sup>. This progress may improve the drug designation that targets DOT1L-AF10 interaction.

Af10 knock out in mice can cause developmental defects such as development of midline facial cleft due to the reduced levels of AP2 $\alpha$  gene<sup>148</sup>. When they use chemical inhibitor of Dot1L (EPZ-5676), they also observed similar defects in embryos<sup>148</sup>. Dot1L knock out mouse embryo is lethal after day 10<sup>119</sup> and Af10-KO mice are lethal after embryonic day 16 and they are exhibiting severely decreased H3K79 methylation levels<sup>148</sup>.



**Figure 10.** Representative structure of the DOT1L–AF10 interaction (Adapted from Zhang *et.al.* with permission presented in Appendix-III<sup>147</sup>)

**A**, AF10 OM-LZ (octamer motif-leucine zipper) structure was represented as cyan and green color depicts for two chains. Dimer interface of two peptides was demonstrated as stick representation. **B**, Cartoon representation of DOT1L-CC2 (coiled-coil domain#2) interaction with AF10-OM-LZ.

### 3.1.4 AF10 regulates histone code

AF10 protein has two important domains: PZP and OM-LZ. PZP domain consists of Zn-knuckle in between two PHD (plant homeodomain) fingers. It was previously known that OM-LZ was the DOT1L-interaction domain<sup>121</sup>. PZP domain was identified as a reader of H3K27 residue by recognition of 22-

27 amino acids of H3<sup>145</sup>. It was demonstrated that PZP domain of AF10 can bind to unmodified H3K27 therefore it can recruit DOT1L, leading to H3K79 methylation<sup>145</sup>. On the other hand, AF10 cannot bind to methylated H3K27, hence DOT1L cannot recruited to methylated H3K27 sites via AF10; which promotes transcriptional silencing<sup>145</sup>. It was previously claimed that AF10 was required for aberrant expression of *HOX* genes in AML cases (acute myeloid leukemia)<sup>146,149,150</sup>. Even though it was claimed that continuous expression of *HOX* gene is dependent on AF10<sup>146,149,150</sup>, in another study it was claimed that AF10's activity was required in cytoplasm rather than gene expression regulation<sup>151,152</sup>. These controversial views need to be supported with further experiments to find out the mechanism of AF10.

AF10 was also thought to regulate H3K79 methylation through H2B ubiquitination<sup>153</sup>. *C. elegans* homolog of AF10, ZFP-1 was demonstrated to negatively regulate essential highly expressed genes via interacting with DOT-1.1<sup>153</sup>. They claimed that H2Bub1 modification was read by ZFP-1 and recruitment of DOT-1.1 to the actively transcribed site results in negative feedback via Pol-II pausing<sup>153</sup>. Since there are different hypotheses on how AF10 regulates gene expression, further studies are required to understand the mechanism and consequences of AF10 - DOT1L interaction.

In this chapter, novel DOT1L interactors were identified using the BioID method in which a promiscuous BirA ligase (BirA\*) was employed to biotinylate DOT1L-proximal proteins, *in vivo*. Biotinylated proteins were pulled-down by Streptavidin and identity of the proteins was determined by LC-MS/MS. The resulting novel interaction candidates were investigated for their effects on reprogramming. Candidate genes were knocked-down in

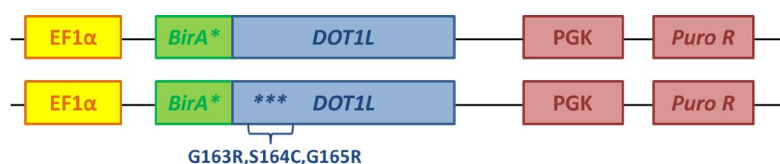
human fibroblasts via shRNAs followed by reprogramming. Our results indicated that knock-down of AF10, significantly increased the iPSC generation efficiency, suggesting that it acts as a barrier to reprogramming similar to DOT1L. This finding was verified by CRISPR/Cas9 mediated knockout of AF10. Combining DOT1L inhibition or knockout, with AF10 suppression did not result in an additive enhancement of reprogramming, suggesting that these two chromatin factors act in the same pathway.



## 3.2 Results

### 3.2.1 BirA\*-DOT1L fusion protein successfully methylates H3K79 residue

DOT1L is an established barrier of reprogramming<sup>55</sup>. However, what role, if any, its interaction partners play in reprogramming is not known. Therefore, we wanted to identify protein interactions of DOT1L. A biotinylation-based proteomics approach, BioID<sup>139</sup>, was used in this study to investigate the protein interaction network of DOT1L. In this method, promiscuous biotin ligase (BirA\*) was fused to DOT1L with the assumption that it will biotinylate proteins that come into close proximity (10 nm radius) with DOT1L (BirA\*-DOT1L). Fusion proteins were generated with either wild-type DOT1L (DOT1L wt) or a catalytically inactive mutant (DOT1L mut). Mutant DOT1L has 3 point mutations (**Figure 11**) that renders it unable to bind the methyl donor S-adenosylmethionine (SAM) and blocks the H3K79 methyltransferase activity<sup>121</sup>. Both wt- and mut- constructs were cloned into a mammalian lentiviral expression vector (pLEX\_307) which contains an E1Falpha promoter and puromycin antibiotic resistance gene (**Figure 11**).



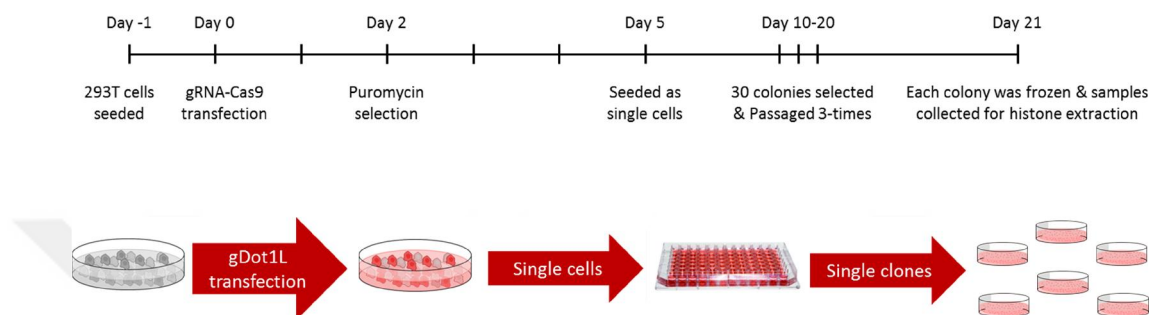
**Figure 11. Schematic of BirA\*-DOT1L fusion proteins in pLEX\_307 vectors**

Catalytically inactive DOT1L was created by 3 point mutations<sup>121</sup>. Both inserts were cloned into pLenti-pLEX-307 vector which has EF1 $\alpha$  promoter and puromycin resistance gene.

To confirm the fusion proteins did not affect DOT1L, the catalytic activity was tested. To do this, DOT1L knockout (DOT1L-KO) cells were generated. Since DOT1L is the sole enzyme that catalyzes H3K79 methylation, functionality of DOT1L can be tested via investigating the H3K79me2 levels. Aim of this experiment was to express BirA\*-DOT1L fusion plasmids in DOT1L-deficient cells and observe the rescue phenotype by H3K79 methylation. DOT1L-KO cells are devoid of H3K79me2 and if ectopically expressed DOT1L fusion proteins are functional, H3K79me2 levels is expected to increase.

DOT1L-KO HEK293T cells were generated by CRISPR-Cas9 mediated genome editing. The first exon of *DOT1L* was targeted with a guideDOT1L plasmid which also contains a Cas9 protein (LentiCRISPRv2). HEK293T cells were transiently transfected with either non-targeting (NT) or guideDOT1L containing Cas9 plasmids and selected with puromycin. After selection was completed, cells were trypsinized and diluted to single cells. Then, single cells were seeded onto 96-well plates (**Figure 12**). Proliferated

single cell clones were pelleted and histone acid extraction was performed for 19 colonies.

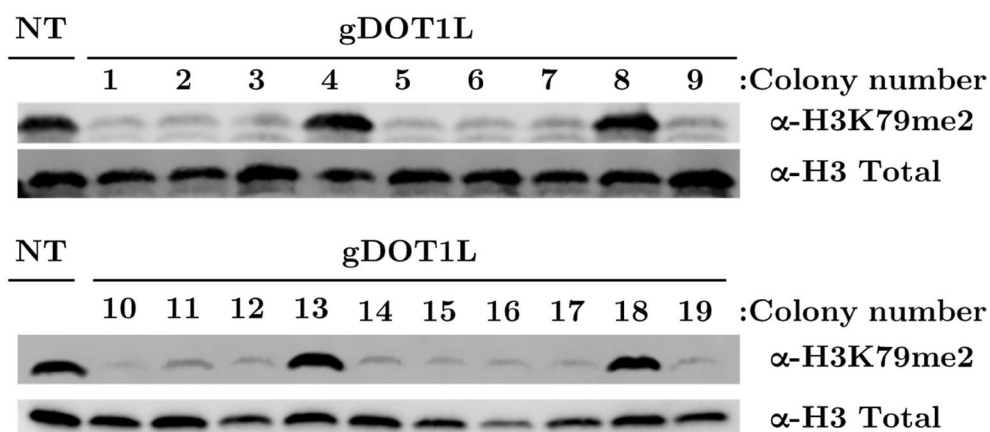


**Figure 12. Time-line for the generation of DOT1L-KO HEK293T cells**

HEK293T cells were transfected with gDOT1L along with Cas9. After puromycin selection, cells were diluted as single cells and transferred to 96-well plates. Single clones were expanded for histone extraction to test their H3K79me2 levels.

All single clones were tested for their H3K79me2 levels via immunoblotting to select a single cell clone that is H3K79me2-deficient. Out of 19 clones, 4 clones retained high levels of H3K79 methylation similar to gNT transfected cells while H3K79me2 levels were significantly downregulated in 15 clones (**Figure 13**). As a result of this experiment, single clone #10 was selected as DOT1L-KO 293T cells for further experiments since it has the lowest levels of H3K79me2 (**Figure 13**). Residual H3K79me2 levels are most likely due to the early passage of the clones, as H3K79me2 bands were invisible in cells passaged 2-3 additional times (**Figure 14**).

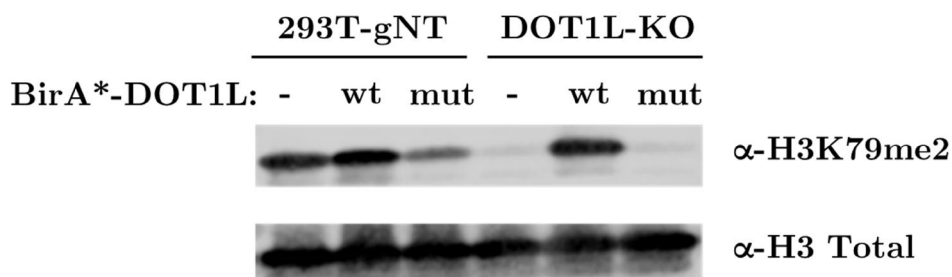




**Figure 13. H3K79me2 levels in DOT1L gRNA transfected HEK293T single cell clones**

Immunoblot results of selected clones. Histone extracted protein lysates were incubated with either H3K79me2 antibody or H3 total antibody. Colony #10 was picked as DOT1L-KO HEK293T due to the drastically decreased H3K79me2 levels. NT is a non-targeting gRNA used as a control.

Having generated DOT1L-KO cells, BirA\*-DOT1L fusion constructs could be tested for their ability to rescue diminished H3K79me2 levels. To test the functionality of the BirA\*-DOT1L fusion proteins; DOT1L-KO HEK293T clone and gNT transfected HEK293T clones were transduced with BirA\*-DOT1L-wt or mutant expressing lentiviral vectors. As a control, untransduced cells were included as well. After puromycin selection, cells were pelleted and histone acid extraction was performed. All samples were tested for their H3K79me2 levels via immunoblotting and H3 total immunoblot was performed as a loading control (Figure 14).



**Figure 14. Rescue of H3K79 methylation by BirA\*-DOT1L\_wt fusion protein in DOT1L-KO cells**

Both wt and mutant fusion protein expressing plasmids were packed into lentiviruses. HEK293T-NT (non-targeting) and DOT1L-KO cells were infected with wt and mutant fusion protein expression vectors. After puromycin selection, cells were collected for histone extraction. Histone extracts were incubated with either H3 total antibody or H3K79me2 antibody.

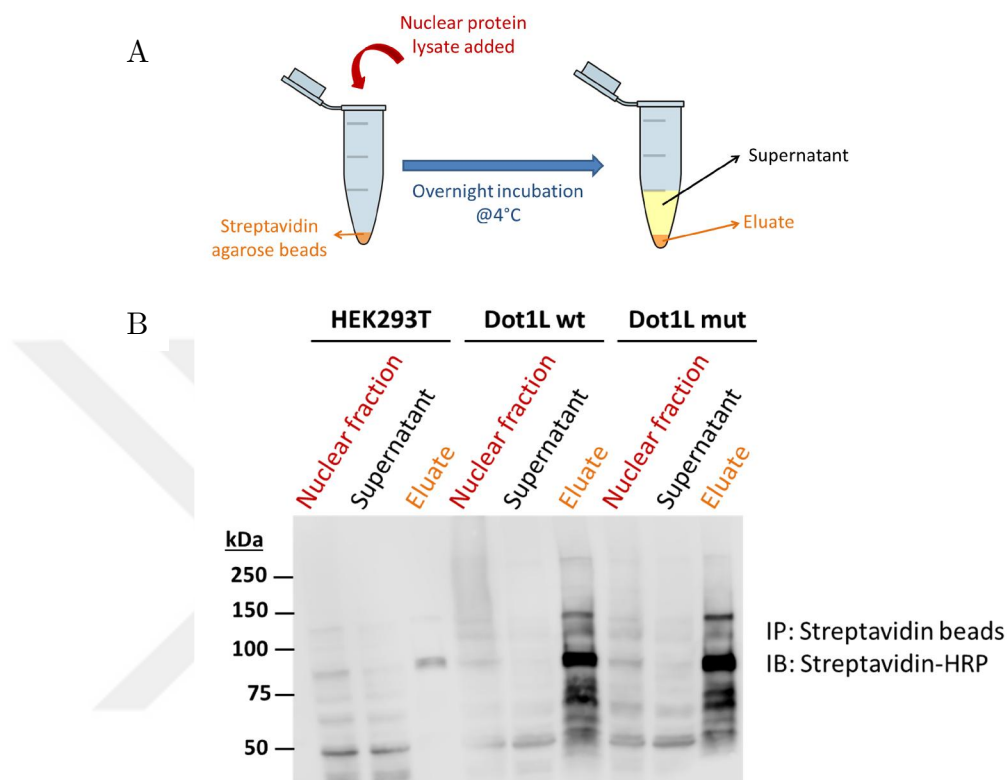
As a result of this assay, BirA\*-DOT1L-wt could rescue the H3K79me2 levels in DOT1L-KO cells while BirA\*-DOT1L-mut had no effect (**Figure 14**). Interestingly, overexpression of the mutant protein caused a decrease in H3K79 methylation in gNT transfected cells, suggestive of a dominant negative effect. This experiment shows that wt DOT1L in BirA\*-DOT1L-wt fusion is enzymatically active, whereas mut-DOT1L fusion is inactive as expected.

### 3.2.2 Pull-down of biotinylated proteins via Streptavidin beads

Next, biotinylation activity of BirA\*-DOT1L fusion proteins were tested to ensure that BirA\* in fusion proteins can successfully biotinylate. Pull-down experiments were performed to test the biotinylation activity of BirA\* in fusion proteins. For pull-down experiments, HEK293T cells were infected with BirA\*-DOT1L wt or mutant viruses and selected with puromycin for 3 days. Cells were treated with 50  $\mu$ M D-Biotin for 24 hours and biotinylated cells were collected for protein isolation. Proteins were obtained via nuclear fractionation method as DOT1L is a nuclear protein. As a control, uninfected HEK293T cells were treated similarly. Pull-down was performed with Streptavidin beads since Streptavidin can strongly bind to biotin.

After pull-down, eluted proteins were analyzed with Streptavidin-HRP antibodies to visualize biotinylated proteins and observe the efficiency of pull-down. Total nuclear fraction and unbound proteins (supernatant) were also analyzed (**Figure 15A**). As a result, it was shown that BirA\* in fusion protein can biotinylate proteins, since many different protein bands were observed in the eluate lane of fusion protein infected cells (**Figure 15B**). In contrast, there were only a few bands present in the eluate of HEK293T control cells which are naturally biotinylated proteins in the cell (**Figure 15B**). In addition, the absence of bands in the corresponding supernatants indicated that biotinylated proteins were efficiently captured by beads and pull-down protocol worked successfully (**Figure 15B**). This experiment shows that BirA\* in fusion proteins can successfully carry out biotinylation reactions. In conclusion, BirA\*-DOT1L fusion proteins can be used for BioID

assay since functionality of both DOT1L and BirA\* have been successfully demonstrated.



**Figure 15. Biotinylated proteins in BirA\*-DOT1L expressing HEK293T cells in comparison to uninfected cells.**

**A**, Nuclear proteins were isolated from uninfected and BirA\*-DOT1L wt and mut infected HEK293T cells. Equal amounts of nuclear lysate were incubated with Streptavidin-agarose beads. Unbound proteins (supernatant) were cleared from tube and the pelleted beads were washed. Biotinylated proteins were eluted from beads. **B**, Biotinylated protein levels were analyzed with Streptavidin-HRP antibody.

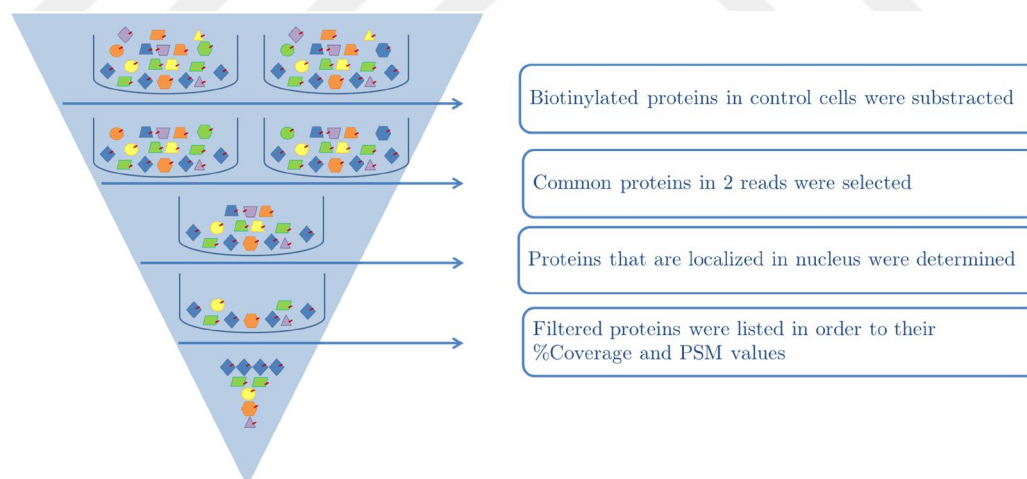
### 3.2.3 Proximal-protein interactions of DOT1L was identified after LC-MS/MS analysis of biotinylated proteins

For mass spectrometry analysis, uninfected control HEK293T cells and BirA\*-DOT1L wt or mut infected HEK293T cells were treated with 50  $\mu$ M biotin for 24 hours and harvested. Following nuclear protein isolation and over-night incubation with Streptavidin beads, biotinylated proteins were captured as explained in section 3.2.2. Biotinylated proteins were released from beads via on-bead trypsin proteolysis and identified by LC-MS/MS. As a result of LC-MS/MS analysis, DOT1L protein was detected with the highest PSM (peptide spectrum matches) values and very high coverage percentage (~30%) in BioID samples; whereas none was detected in control samples (Table 14). This results shows that DOT1L was successfully overexpressed and biotinylated only in fusion vector transduced cells.

**Table 14. DOT1L detection values after mass spectrometry analysis from two reads of each sample (nd, not detected)**

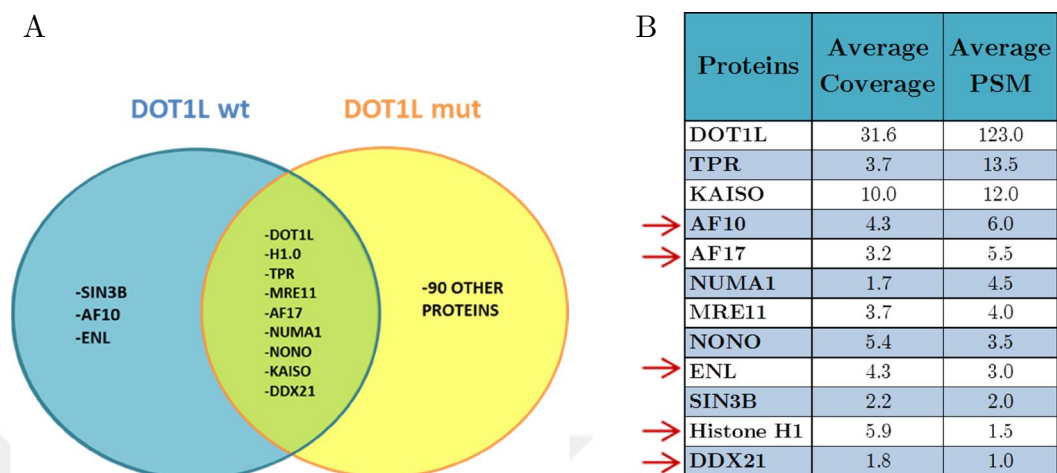
Sample	Coverage %	PSM
Uninfected HEK293T	nd	nd
	nd	nd
HEK293T BirA*-DOT1L WT	28.12	101
	35.02	145
HEK293T BirA*-DOT1L mut	25.13	87
	26.74	101

The biotinylated proteins detected in HEK293T control samples were excluded from BioID samples to obtain the list of DOT1L-proximal proteins specifically biotinylated by BirA\*-DOT1L fusions. These eliminated proteins are either naturally biotinylated proteins or non-specifically bound to Streptavidin beads. Among remaining proteins, common ones from two readings of LC-MS/MS results were determined. Then, common proteins in wt- and mut-DOT1L samples were selected according to their localization in the cell via GO annotation. Extra-nuclear proteins were disregarded with the assumption that they were contaminants from nuclear fractionation. Remaining proteins were sorted according to their coverage % and PSM (peptide spectrum matches) values. The analysis process is summarized in **Figure 16**.



**Figure 16. Flowchart for identifying proteins specifically biotinylated by BirA\*-DOT1L in LC-MS/MS analysis**

As a result of LC-MS/MS analysis, 160 and 251 proteins were detected in HEK293T cells from 2 readings. When 2 readings of LC-MS/MS were combined, 118 proteins were common in both readings and in total 224 proteins were identified in control cells as non-specific background. When these background proteins were omitted from list of fusion protein infected samples, there were 35 and 89 proteins remaining in 2 readings of wt-DOT1L whereas there were 171 and 324 proteins in 2 readings of mut-DOT1L. When 2 readings of each sample were combined, 14 proteins were common in wt-DOT1L and 129 proteins in mut-DOT1L sample. Among these proteins, extranuclear ones were eliminated and 12 proteins remained in wt-DOT1L and 99 proteins in mut-DOT1L sample. 9 of these proteins are common in wt- and mut-DOT1L samples (**Figure 17A**). Among these 9 common proteins, AF17 and DDX21 were previously reported to interact with DOT1L<sup>125</sup>. AF10 and ENL were detected in only wt-DOT1L samples and their direct interaction with DOT1L was previously reported<sup>125</sup>. Novel interactions of DOT1L were also detected. WT-DOT1L sample had 7 novel interactions (TPR, KAISO, NUMA1, MRE11, NONO, SIN3B, Histone H1) whereas mut-DOT1L samples had 96 novel-proximal interactions (**Figure 17**). Whether these novel proteins are directly interacting with DOT1L requires further investigation but these proteins can be referred as DOT1L-proximal proteins.



**Figure 17. Novel proximal interactions of DOT1L were revealed via BioID along with known direct interactions of DOT1L.**

**A**, Wt and mut DOT1L's proximal interactors. **B**, Possible novel interaction partners of WT DOT1L that are ranked according to their PSM (peptide spectrum matches) and coverage% scores. Red arrows indicate previously known interactions of DOT1L.

In conclusion, BioID analysis identified five proteins (AF10, AF17, ENL, Histone1 and DDX21) that were previously reported as direct interactors of DOT1L. In addition, novel proximal proteins such as TPR, KAISO, NUMA1, MRE11, NONO and SIN3B were identified as potential interaction partners of DOT1L. On the other hand, mut-DOT1L had dozens of proximal protein interactions. This may be due to a defect in chromatin localization of the mutant protein, a notion that needs further experimental verification.

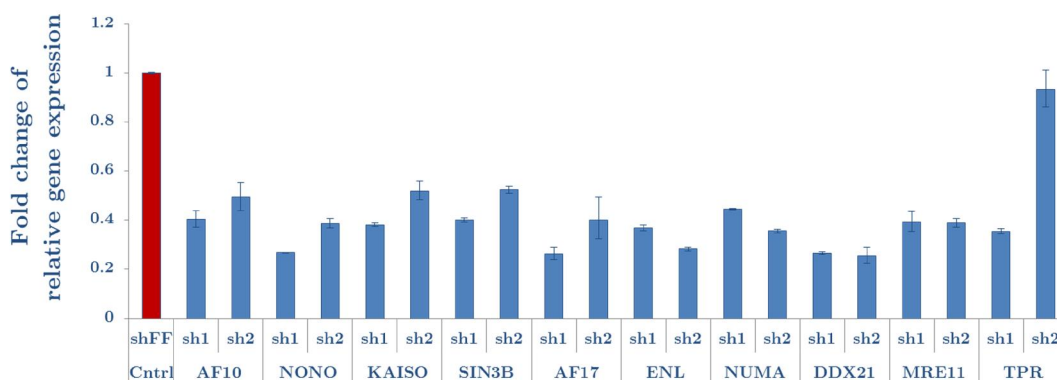


### 3.2.4 shRNA-mediated knock-down of DOT1L-proximal proteins

To determine the effect of DOT1L-interacting proteins on reprogramming, a loss-of-function based reprogramming screen was designed using short hairpin RNAs (shRNAs). Potential DOT1L-interacting proteins were curated from the BioID screen (**Figure 17B**). 12 proteins were found as potential interactors of wt-DOT1L including 4 proteins that were previously known to directly interact with DOT1L. Among 12 proteins, one of them was DOT1L protein itself and another one was histone protein, Histone H1.0, therefore no shRNA was targeted for it. For the remaining 10 proteins in DOT1L's interaction network (AF10, NONO, KAISO, SIN3B, AF17 MRE11 ENL, NUMA1, TPR, DDX21); 2 shRNAs were designed to target each gene. All shRNAs were cloned into puromycin resistance gene containing retroviral pSMP vector.

Successfully cloned shRNAs were packaged into retroviral particles and transduced to dH1f human fibroblast cells<sup>132</sup>. As a control (shControl), firefly luciferase targeting pSMP\_sh-Luc (shFF) was used. Upon completion of puromycin selection, RNA isolation was performed and the expression levels of shRNA targeted genes were quantified by qRT-PCR. mRNA levels of each gene was normalized to shControl (shFF) infected cells.

All shRNAs achieved at least 50% knock-down of their respective target genes, except shTPR#2 (Figure 18). As a result, 2 functional shRNAs were designed, cloned and tested against 10 proximal proteins of DOT1L except TPR [*it has only 1 functional shRNA*].

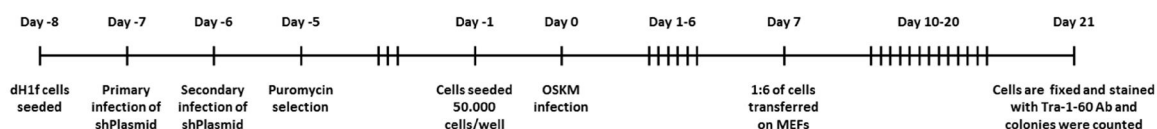


**Figure 18. mRNA expression levels of shRNA-targeted genes**

qRT-PCR was performed as duplicate samples and  $\beta$ -actin was used as an internal control gene. Every genes expression level is normalized to shControl (shFF) infected cells. (n=2; error bars represents s.d.)

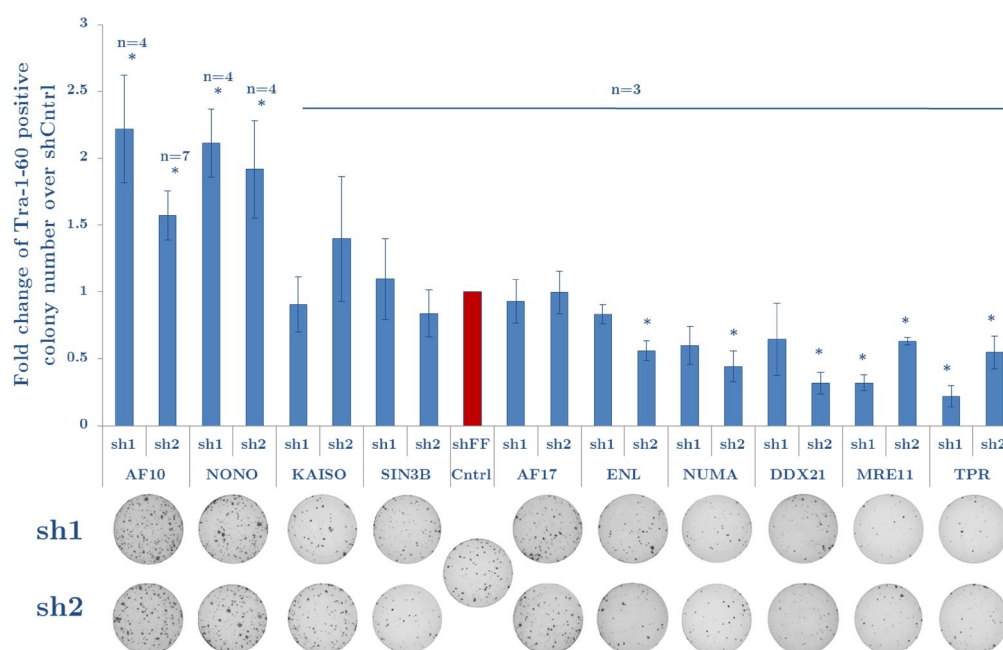
### 3.2.5 shRNA-mediated screen of DOT1L proximal interactors for reprogramming efficiency

To test the loss-of function effects of DOT1L proximal proteins on reprogramming, dH1f cells were infected with shRNA viruses and reprogrammed with OSKM (**Figure 19**).



**Figure 19. Timeline of reprogramming experiments for shRNA infected fibroblasts**

Seven days after OSKM transduction, cells were transferred onto MEFs. At the end of the reprogramming procedure, cells were stained with Tra-1-60 (embryonic cell surface marker) antibody to identify iPSC colonies. Quantifications were carried out with Image-J and results were compared with shControl sample with respect to Tra-1-60 positive colony number (**Figure 20**).



**Figure 20. Fold change in reprogramming efficiency as a result of shRNA-mediated gene silencing**

Reprogramming experiment for shRNA screen was performed. Average colony number of each experiment was calculated and normalized to shControl to calculate fold change. Average of fold change from different experiments was calculated and standard error is depicted in error bars. ( $P$  values were determined by a one sample t-test. \*,  $P < 0.05$ ; n is indicated above the bars and denotes independent biological replicates) Representative Tra-1-60 stained well images are displayed underneath the bar graph.

As a result of this shRNA screen, it was observed that knock-down of AF10 and NONO significantly increased the number of iPSC colonies, resulting in 1.5 to 2 fold greater reprogramming efficiency. (**Figure 20**). On the other hand, knock-down of MRE11 and TPR decreased reprogramming drastically while knock-down of KAISO, SIN3B and AF17 did not significantly affect reprogramming (**Figure 20**). In the case of ENL, NUMA and DDX21, one of the shRNAs did not affect the reprogramming whereas other shRNA decreased the iPSC colony numbers.

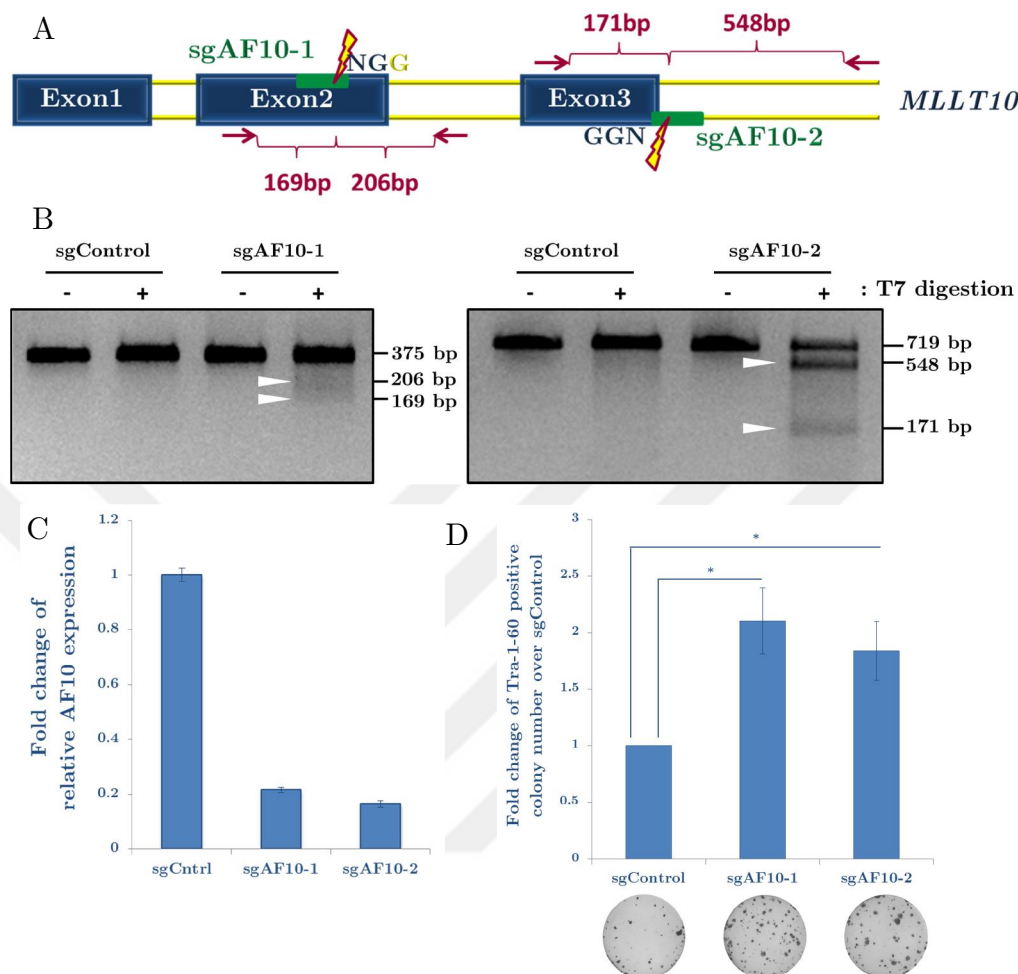
Effect of AF10 and NONO on reprogramming has not been studied in the literature and in this screen it was demonstrated that knock down of AF10 and NONO increase reprogramming. In this project, molecular mechanism behind AF10's effect on reprogramming was further investigated. However, effect of NONO on reprogramming would also warrant further investigation since its knock-down increases reprogramming significantly.

### **3.2.6 Knock-out of AF10 via CRISPR increases reprogramming efficiency similar to knock-down of AF10 via shRNA**

The strongest candidate protein identified from the shRNA-mediated reprogramming screen was AF10, which is known to be a well-established direct interactor of DOT1L<sup>154</sup>. Follow up experiments were performed to better understand the mechanism of action of AF10 during reprogramming. First, CRISPR/Cas9-mediated knock-out of AF10 was tested to verify the increase in the reprogramming via shAF10. *AF10 (MLLT10)* genomic region

was targeted by 2 independent sgRNAs targeting exon 2 or exon 3 (**Figure 21A**). Non-targeting control gRNA (sgControl) and sgAF10-1 & -2 containing lentiviral vectors were transduced to dH1f cells which were then selected with puromycin. CRISPR/Cas9-mediated mutations of the sgAF10 target sites were demonstrated via T7 endonuclease assay (**Figure 21B**). In addition, sgAF10 expressing dH1fs had lower *AF10* mRNA levels compared to sgControl dH1fs as assessed by q-RT-PCR (**Figure 21C**). These experiments demonstrate that sgAF10s disrupts the *AF10* gene.

Next, sgAF10 infected dH1fs were reprogrammed with OSKM transduction. At the end of reprogramming, iPSC colonies were quantified with Tra-1-60 antibody staining and normalized to sgControl infected dH1fs (**Figure 21D**). sgAF10 expressing fibroblasts generated approximately 2-fold greater number of iPSC colonies compared to control cells indicating that loss of AF10 increases reprogramming efficiency (**Figure 21D**). This experiment shows that knock-out of AF10 increases reprogramming efficiency similar to knock-down of AF10.

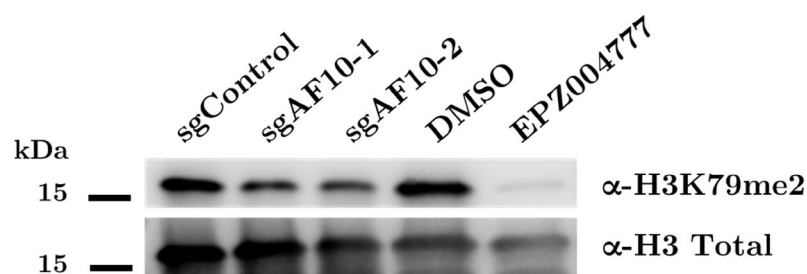


**Figure 21. CRISPR/CAS9 based suppression of AF10 and its effect on reprogramming**

**A**, sgAF10 targeting sites on *AF10* (*MLLT10*) gene. Green bars show gRNA targeting site and NGG shows the PAM sequence that Cas9 targets. Red arrows show the T7 assay primers and expected DNA fragments were depicted. **B**, T7-endonuclease assay for sgAF10 targeting sites. **C**, sgRNA-mediated AF10 knock-out decreases expression levels of *AF10* mRNA. (n=2; error bars represents s.d.) **D**, sgRNA-mediated AF10 knock-out increases reprogramming. Tra-1-60 colony numbers were quantified and normalized to sgControl to calculate fold change. Representative Tra-1-60 stained well images are displayed underneath the bar graph. (*P* values were determined by a one sample t-test; \* *P* < 0.05; n=5 and denotes independent biological replicates)

### 3.2.7 sgAF10s decreases H3K79 methylation

It has previously been reported that AF10 is responsible for the recruitment of DOT1L to most of its target genes<sup>149</sup> and that loss of AF10 decreases overall H3K79 methylation levels. To confirm these findings in our system, H3K79me2 levels were measured in sgAF10 infected dh1fs. dh1fs treated with a small molecule inhibitor of DOT1L (EPZ004777) were used as a positive control for H3K79me2 depletion. Since EPZ004777 was dissolved in DMSO, control dh1fs were treated with DMSO. Immunoblot with H3K79 dimethyl specific antibody demonstrated a reduction in total H3K79me2 levels in sgAF10 expressing cells though not as extensive as small molecule inhibition of DOT1L (**Figure 22**). This result indicates that AF10 expression is needed for H3K79 methylation.



**Figure 22. gRNA mediated AF10 knock-out decreases H3K79me2 levels of dh1fs**

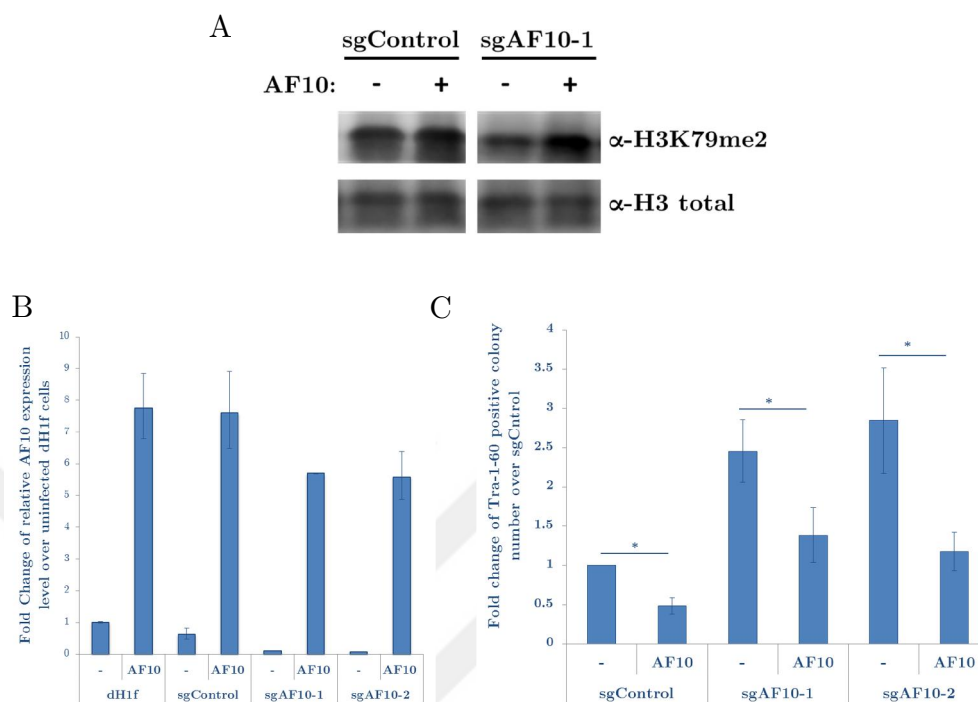
EPZ004777, a small molecule inhibitor of DOT1L, was used as a positive control of H3K79me2 depletion. dh1f cells were treated with DMSO or 3  $\mu$ M EPZ004777 for 10 days. sgAF10 infected dh1fs were selected with puromycin and cultured for 1 week.

### 3.2.8 sgAF10's effect on reprogramming can be reversed via overexpression of AF10

It was shown that silencing of AF10 via shRNAs or sgRNAs can increase reprogramming efficiency. To prove that these results are due specifically to AF10's silencing and not to off-target effects, a rescue experiment was designed. For this purpose, AF10 cDNA was cloned into a mammalian expression plasmid that has Hygromycin resistance gene. Since sgAF10 plasmids have puromycin resistance, both plasmids can be sequentially infected and selected with different antibiotics. First, dH1f cells were transduced with sgControl and sgAF10 vectors and selected with puromycin. After puromycin selection was completed, cells were infected with AF10 overexpression plasmid and selected with Hygromycin. sgAF10s do not target the AF10 overexpression construct because they are complementary to the splice junctions of *AF10* gene which are absent in exogenous AF10 cDNA. Successfully selected dH1f cells were used in histone extraction, RNA isolation or reprogramming experiments (Figure 23).

qRT-PCR confirmed that *AF10* mRNA levels were increased 5 to 7-fold upon AF10 overexpression in control and sgAF10 cells (Figure 23B). Histone extraction and H3K79me2 immunoblot indicated that AF10 overexpression increased the overall H3K79me2 levels in dH1fs (Figure 23A). When AF10 is overexpressed, the reprogramming efficiency decreased approximately by half (Figure 23C). Taken together, these experiments showed that the increased reprogramming phenotype upon AF10 silencing can be rescued by the overexpression of AF10 cDNA, which further strengthened the notion that AF10 is a barrier to reprogramming.





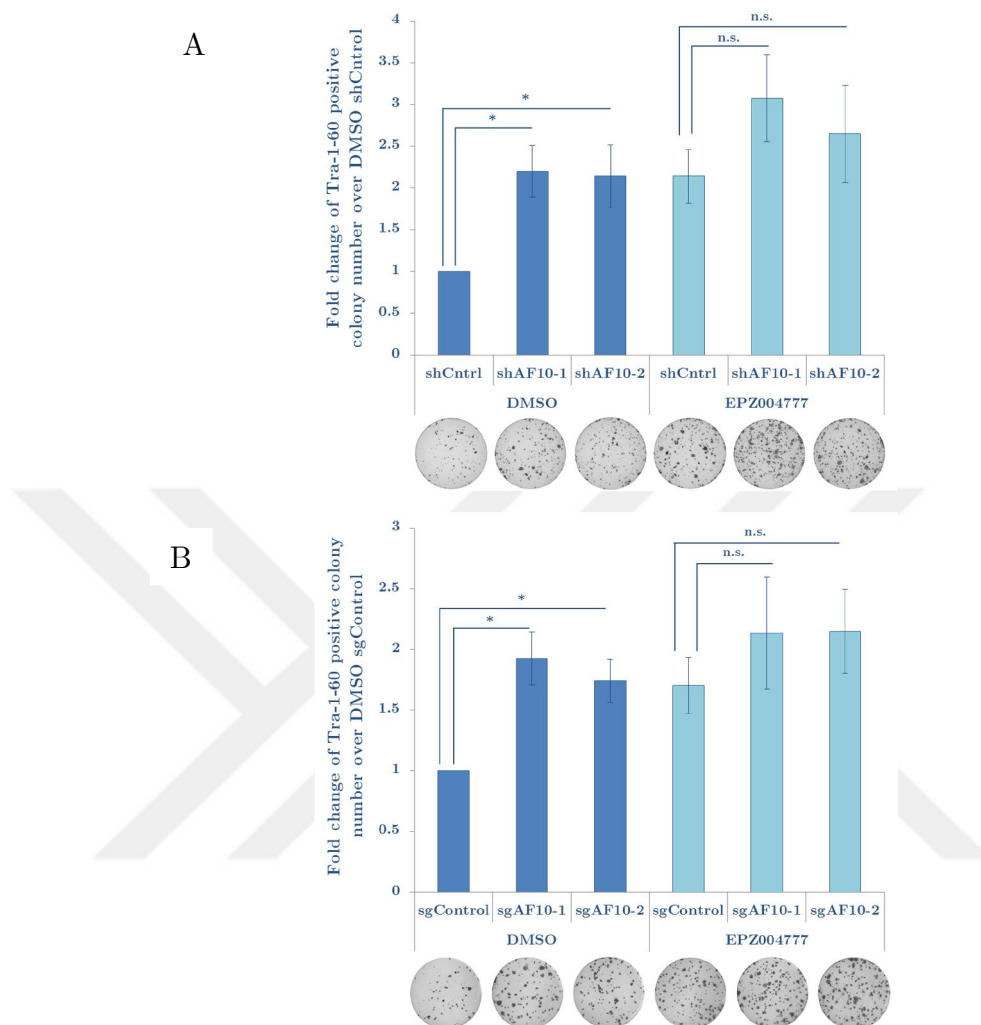
**Figure 23. AF10 overexpression rescues sgAF10 phenotypes**

**A**, H3K79me2 levels in sgAF10 infected cells and AF10 rescued cells. gRNA mediated AF10 knock-out decreases H3K79me2 levels of dH1fs. Conversely, overexpression of AF10 increases H3K79me2 levels. **B**, sgRNA-mediated AF10 knock-out decreases expression levels of *AF10* mRNA whereas overexpression of AF10 increases *AF10* expression. (n=2; error bars represents s.d.) **C**, Tra-1-60 colony numbers were quantified and normalized to sgControl to calculate fold change. sgRNA-mediated AF10 knock-out increases reprogramming while overexpression of AF10 significantly decreases reprogramming for both sgControl and sgAF10 infected cells. (*P* values were determined by a one sample t-test; \*, *P* < 0.05; n=3 and denotes independent biological replicates)

### 3.2.9 Loss-of AF10 together with inhibition of DOT1L does not have an additive effect on reprogramming efficiency

After demonstrating that genetic suppression of AF10 expression via shRNA or sgRNAs increases reprogramming efficiency, we next considered the possibility that AF10 blocks reprogramming through interfering with DOT1L function. To test this hypothesis, shAF10 and sgAF10 treated cells were reprogrammed with or without addition of a DOT1L inhibitor (EPZ004777). sgAF10 or shAF10 infected dH1fs were treated with 3  $\mu$ M EPZ004777 for 6 days after OSKM transduction and control cells were treated with DMSO. At the end of reprogramming, Tra-1-60 positive colonies were counted and their fold change over DMSO treated shControl/sgControl cells was calculated.

The results revealed that, consistent with previous experiments, AF10 knock-down or knock-out on their own increases reprogramming. However, this effect is lost upon Dot1L inhibition, suggesting that the AF10 phenotype is dependent on the presence of H3K79 methylation activity of DOT1L. In other words, in DOT1L-inhibited cells loss of AF10 does not increase reprogramming efficiency further (**Figure 24**). These results suggest that AF10 and DOT1L may be acting in the same pathway to block reprogramming.



**Figure 24. Knock-down or Knock-out of AF10 does not further increase reprogramming efficiency in DOT1L-inhibited condition**

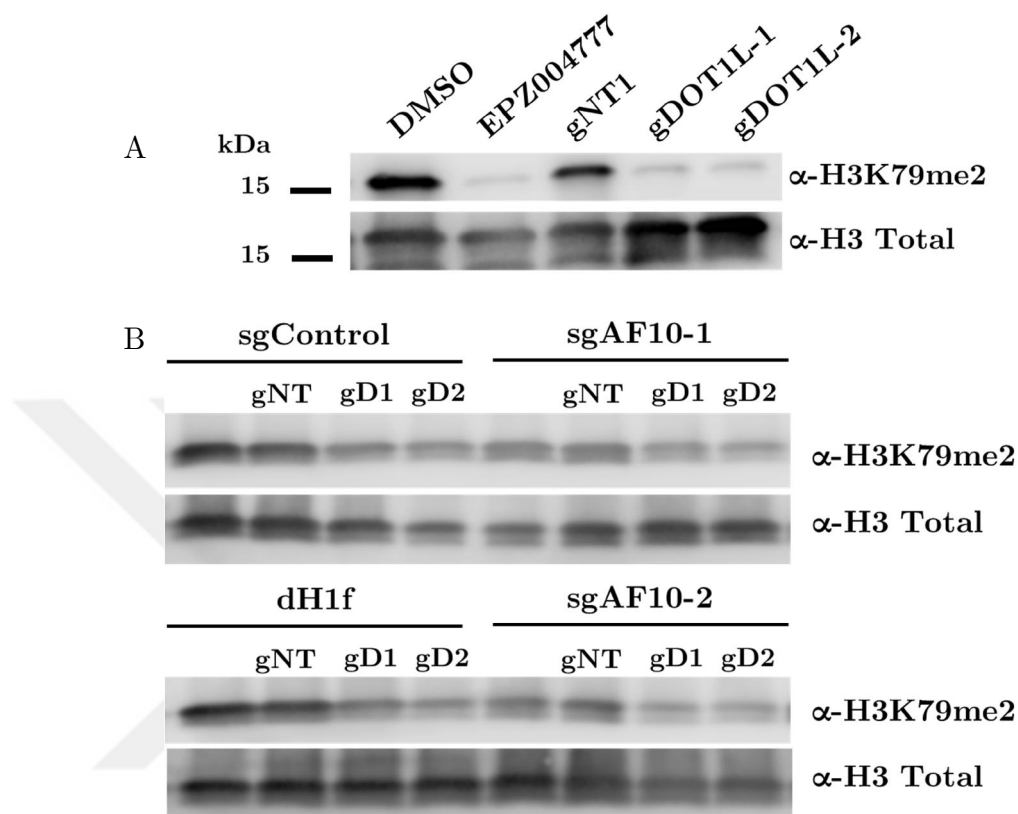
**A**, shAF10 infected cells were reprogrammed with DMSO or 3  $\mu$ M EPZ004777 treatment for 6 days. Average colony number of each experiment was calculated and normalized to shControl to calculate fold change. Average of fold change from  $n=5$  experiments was calculated and standard error is depicted in error bars. **B**, sgAF10 infected cells were reprogrammed with DMSO or 3  $\mu$ M EPZ004777 treatment for 6 days. Average colony number of each experiment was calculated and normalized to sgControl to calculate fold change. Average of fold change from  $n=4$  experiments was calculated and standard error is depicted in error bars. ( $P$  values were determined by a one sample t-test; \*,  $P < 0.05$ ) Representative Tra-1-60 stained well images are displayed underneath the bar graph.

### 3.2.10 Double-KO of AF10 and DOT1L during reprogramming

Since AF10 and DOT1L may work together during reprogramming, AF10 and DOT1L double knock-outs were generated to test their effect on reprogramming. DOT1L-KO cells were generated by gDOT1L-1 & gDOT1L-2 CRISPR plasmids both targeting exon 5 of *DOT1L gene (KMT4)*. sgAF10s and sgControl infected cells were knocked-out with non-targeting gRNA (gNT1) and gDOT1Ls (gD1 and gD2). Double-KO cells were reprogrammed to compare the reprogramming efficiency.

First, dH1fs were infected with non-targeting gRNA (gNT1) and gDOT1Ls (gD1 and gD2) and selected with puromycin. To test the gRNAs efficiency on silencing DOT1L, H3K79me2 levels were measured. DMSO and 3  $\mu$ M EPZ004777 treated dH1fs were used as a control as in **Figure 22**. gDOT1L infected cells exhibited drastic reduction in H3K79me2 levels similar to EPZ004777 treated cells (**Figure 25A**). This experiment demonstrates that sgRNA-mediated DOT1L suppression causes depletion of H3K79me2.

For reprogramming experiment, dH1f cells were infected with sgAF10 and sgControl viruses and selected with puromycin. Successfully selected cells were then transduced with non-targeting gRNA (gNT1) and gDOT1Ls (gD1 and gD2) without further puromycin selection. Cells were passaged for 10 days and half of the cells were harvested for histone extraction and the other half of the cells were utilized for reprogramming experiments. To test double-KO dH1fs, collected cells were investigated for their H3K79me2 levels (**Figure 25B**).

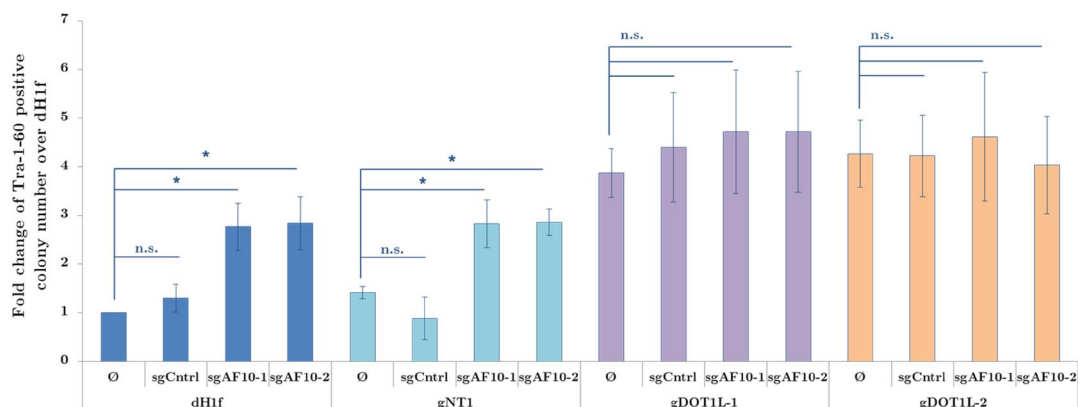


**Figure 25. H3K79me2 levels in gRNA-mediated AF10 & DOT1L knock-out dh1fs**

**A**, gDOT1L infected cells have depleted H3K79me2 levels. EPZ (EPZ004777) is a small molecule inhibitor of DOT1L, used as a positive control of H3K79me2 depletion. dh1f cells were treated with DMSO or 3  $\mu$ M EPZ004777 for 10 days. gDOT1L infected dh1fs were selected with puromycin and cultured for 10 days. **B**, gDOT1L and sgAF10 double knock-out cells display decreased H3K79me2 levels. (H3 total immunoblot depicts histone loading)

We observed that sgAF10 transduced cells exhibited reduction in H3K79 methylation and gDOT1Ls caused a further decrease in H3K79 methylation (**Figure 25B**). The decrease in H3K79me2 levels are not as drastic as in **Figure 25A** since no puromycin selection was done after gDOT1L transduction.

Double-KO cell populations were reprogrammed with OSKM transduction. At the end of reprogramming experiment, Tra-1-60 positive colony numbers were quantified and normalized to the uninfected dH1fs (**Figure 26**). This reprogramming experiment demonstrated that AF10-KO cannot further increase the reprogramming when DOT1L was knocked-out even though sgAF10 and gDOT1L can increase reprogramming 3 to 4-fold, by themselves (**Figure 26**). These results suggest that silencing AF10 increases reprogramming through DOT1L.



**Figure 26. gRNA mediated AF10 & DOT1L double knock-out does not have additive effect on reprogramming**

sgAF10 and gDOT1L infected cells were reprogrammed. Average colony number of each experiment was calculated and normalized to uninfected dH1fs to calculate fold change. Average of fold change from  $n=3$  experiments were calculated and standard error is depicted in error bars. ( $P$  values were determined by a one sample t-test; \*,  $P < 0.05$ ; n.s., not significant,  $P > 0.05$ )

Experimental work described in this chapter indicated that BioID analysis identified TPR, KAISO, NUMA1, MRE11, NONO and SIN3B as novel proximal proteins of DOT1L. Taken together, these data are the first to demonstrate that knock-down of AF10, significantly increased iPSC generation efficiency, suggesting that it acts as a barrier to reprogramming similar to DOT1L. This finding was verified by CRISPR/Cas9 mediated knockout of AF10. In addition we showed that combining DOT1L inhibition or knockout, with AF10 suppression did not result in an additive enhancement of reprogramming, suggesting that these two chromatin factors act in the same pathway.

## Chapter 4

### 4 MLL1 silencing increases reprogramming efficiency

#### 4.1 Introduction

*Mixed Lineage Leukemia 1* (MLL1/KMT2A) gene was discovered by its association with leukemia through genetic analyses<sup>155</sup>. Homozygous *Mll* knockout mice are embryonic lethal at day 11.5-14.5 and exhibit decreased amount of hematopoietic cells in the liver<sup>156</sup>. Further investigations revealed that MLL1 regulates growth of hematopoietic precursors<sup>156</sup>. MLL1 catalyzes H3K4 methylation and it mainly regulates *Hox* genes expression<sup>157</sup>. MLL1 also regulates cell cycle through cyclins, CDK inhibitors and transcription factor GATA3<sup>158,159</sup>. Therefore, MLL1 is a master regulator that has important roles in gene expression regulation during development, cell cycle and hematopoiesis.



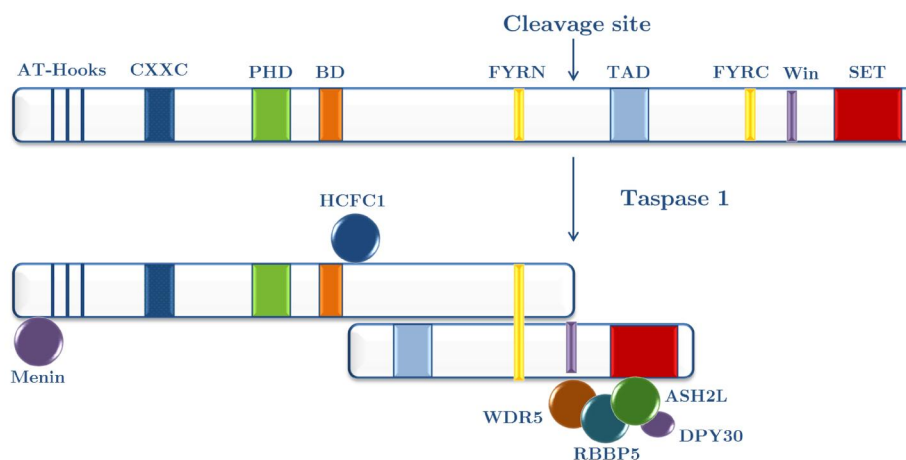
### 4.1.1 MLL1 translocations in Acute Myeloid Leukemia (AML) patients

MLL1 (KMT2A) is a histone H3 lysine 4 (H3K4) methyltransferase and H3K4 methylation on chromatin is associated with active gene transcription. MLL1 enzyme is well-known for its translocations which leads to 70% of the infant acute leukemias<sup>160</sup>. There are more than 60 fusion partners of MLL1<sup>160</sup>. Among MLL1 translocations the top five subtypes are: MLL1-AF4, MLL1-ENL, MLL1-AF9, MLL1-AF10, MLL1-AF6<sup>161</sup>. In the MLL1 translocation, MLL1 loses its SET domain which is the H3K4 methyltransferase domain; however the fusion protein still causes increased expression of some target genes<sup>162</sup>. One of the most prominent outcomes of MLL1 translocations is that aberrant transcriptional regulation of HOX genes<sup>123</sup>. Even though MLL1 loses its SET domain, HOX genes' expression is upregulated. It was shown that many of the fusion partners of MLL1 such as AF9, AF10, AF17 and ENL are direct interactors of DOT1L, the H3K79 methyltransferase. MLL fusion target genes were high in H3K79 methylation which supports the common interactor DOT1L as an important player of leukemia<sup>163</sup>. Since H3K79 methylation is found in actively transcribed genes, DOT1L was found as a strongest suspect of HOX genes' aberrant transcriptional upregulation due to DOT1L's interaction with MLL1 fusion partners. Therefore, there have been many attempts to use small molecule inhibitors of DOT1L in AML patients to inhibit DOT1L methyltransferase activity<sup>123</sup>. However, using these drugs might result in side effects such as disruption in normal hematopoiesis because DOT1L has a role during normal hematopoiesis as well<sup>124</sup>. Therefore, inhibition of interactions between MLL fusion partners and DOT1L to block

aberrant recruitment of DOT1L to MLL1 target genes could be an alternative route to ameliorate AML.

### 4.1.2 MLL complex in transcription

MLL1 has multiple functional domains including AT-hook domains and a CXXC motif that bind to DNA, plant homeodomains (PHDs), a bromo domain (BD), transactivation domain (TAD) and a SET domain that catalyzes the H3K4 methylation<sup>164</sup>. MLL1 is cleaved by Taspase1 into N-320 kDa and C-180 kDa fragments that are forming heterodimers to stabilize complex<sup>165</sup>. There are two phenylalanine-tyrosine rich regions (FYR) in MLL1 that re-associates the MLL1-N and MLL1-C fragments<sup>166</sup> (**Figure 27**). MLL1 has evolutionary conserved protein-binding sites such as MENIN binding motif (MBM), WDR5 interaction motif (Win) and LEDGF binding domain (LBD)<sup>167</sup>. MLL1<sup>-/-</sup> mice have a fetal liver hematopoiesis defect which is associated with decreased expression of HOX genes<sup>156</sup>. Dysregulation of HOX genes resulted in acute myeloid leukemia (AML) where there are MLL1 translocations from its break point region<sup>168</sup>. Therefore, MLL1 has a crucial role during development of hematopoietic precursors.



**Figure 27. Multiple functional domains of MLL1 is depicted before and after its cleavage by Taspase1**

DNA binding AT-hooks, zinc finger containing CXXC motif, plant homeodomain (PHD) fingers, bromodomain (BD), phenylalanine-tyrosine rich regions (FYR), transactivation domain (TAD), WDR5 interaction (Win) motif, and the histone methyltransferase SET domain are highlighted. The full-length MLL1 protein is cleaved by Taspase 1 into MLL-N (300 kDa) and MLL-C (180 kDa) fragments that then re-associate through FYRN and FYRC motifs to form stable complex. Adapted from *Dharmarajan and Cosgrove*<sup>167</sup>.

MLL1 forms a dynamic complex that regulates H3K4 methylation. The core subunits of this complex are WDR5, RBBBP5, ASH2L and DPY30 which are referred as WRAD<sup>169</sup>. It has been demonstrated by structural biologists that RBBP5-ASH2L heterodimer activates the catalytic function of MLL family histone methyltransferases<sup>170</sup>. In addition to core components there are accessory proteins in the complex which are CXXC1, WDR82, HCFC1 and MENIN<sup>171</sup>. These interactors of MLL proteins are important to recruit MLL complex to the specific gene sites therefore assist in regulation of H3K4 methylation. MENIN is thought to bind DNA directly and recruit MLL

complex to specific sites<sup>172</sup>. Also MENIN binding domain (MBD) is retained in the translocated MLL1 fusion proteins therefore MENIN drives the aberrant recruitment of fusion partners of MLL1 at promoters of HOX genes<sup>172</sup>.

MLL1 can catalyze mono-, di-, tri-methylation of H3K4 residue with its SET domain. There are catalytic subunits other than MLL1 as a H3K4 methyltransferase in dynamic MLL complex such as SET1A, SET1B, MLL2, MLL3 and MLL4 which are all MLL family proteins<sup>170</sup>. H3K4 methylation is correlated with transcriptionally active genomic sites. And each H3K4 methylation has distinct transcriptional outcomes: H3K4 tri-methylation found in actively transcribed genes, H3K4 di-methylation is associated with poised chromatin and H3K4 mono-methylation is enriched at enhancers, ribosomal DNA<sup>167</sup>. H3K4 methylation is tightly regulated process since methylation degree has different functions in chromatin.

### **4.1.3 The role of MLL1 in pluripotency**

During reprogramming, H3K4 methylation levels are increased globally<sup>173,174</sup>. MLL complex is mainly responsible from H3K4 methylation. It was previously shown that Yamanaka factors are interacting with proteins in MLL complexes<sup>175</sup>. SOX2 strongly binds ASH2L and WDR5 through its high-mobility group (HMG) domain<sup>175</sup>. DPY30 and RBBP5 knock-down in MEF decreases the OCT4 expressing cells<sup>175</sup>. In another study Wdr5 was shown as an essential protein for reprogramming and WDR5 is overexpressed during reprogramming process<sup>76</sup>. On the other hand, it was shown that an inhibitor

that blocks WDR5-MLL1 interaction is sufficient for reprogramming of epiblast stem cells (EpiSCs) into naïve pluripotency<sup>176</sup>. They also showed that inhibition of WDR5-MLL1 interaction results in global redistribution of H3K4me1 marks on enhancers<sup>176</sup>. This study shows the importance of H3K4 methylation mark on pluripotency however MLL1's role on somatic cell reprogramming is unknown.

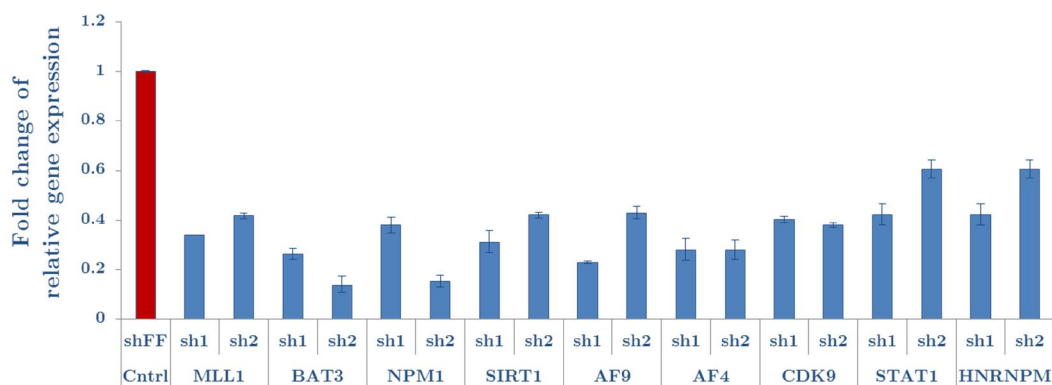
In this chapter, known direct and functional interactors of DOT1L were curated from the literature and their effect on reprogramming was investigated through loss of function experiments. Suppression of *Mixed Lineage Leukemia 1 (MLL1)* expression via RNA interference or CRISPR/Cas9 significantly increased reprogramming efficiency. To determine how MLL1 prevents reprogramming, RNA-sequencing was performed. MLL1 suppression resulted in downregulation of fibroblast-specific genes and accelerated the activation of pluripotency-related genes.

## 4.2 Results

### 4.2.1 shRNA-mediated screen of DOT1L's previously known interactors for reprogramming efficiency

To test the loss of function effect of DOT1L's previously known interactors on reprogramming; an shRNA-mediated screen was carried out. For this purpose, proteins that are known to interact with DOT1L directly or functionally were curated from the literature. As a result, 13 proteins were found to be closely related with DOT1L, however among those proteins 4 of them (AF10, AF17, ENL and DDX21) have been already investigated, which were also identified in our BioID assay. For the 9 remaining proteins (MLL1, BAT3, NPM1, SIRT1, AF9, AF4, CDK9, STAT1 and HNRNPM) in the interaction network of DOT1L, 2 shRNAs were designed to target each gene. All shRNAs were cloned into puromycin resistance gene containing mammalian expression vector (pSMP) and verified by sequencing.

Successfully cloned shRNA plasmids were packaged into retroviruses. dH1fs were infected with each shRNA viruses twice and selected with puromycin for 2-3 days. As a control shRNA (shControl), firefly luciferase targeting pSMP\_sh-Luc (shFF) was used. Successfully selected cells were collected for RNA isolation and q-RT-PCR was performed to quantify the expression levels of targeted genes. While quantifying the results, mRNA level of each gene was compared with shControl infected cells (**Figure 28**). As a result, every shRNA successfully knocked-down its target gene at least 40%.

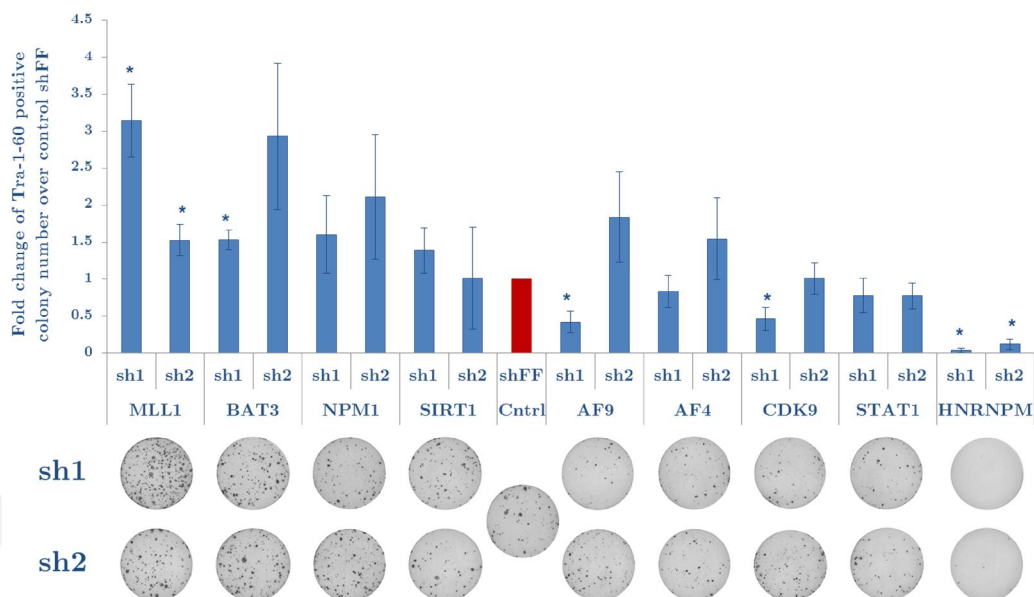


**Figure 28.** mRNA expression levels of shRNA-target genes

qRT-PCR was performed as duplicate samples and  $\beta$ -actin was used as an internal control gene. Every genes expression level is normalized to shControl (shFF) infected cells. (n=2; error bars represents s.d.)

shRNA infected dH1f cells were reprogrammed as previously summarized in **Figure 19**. At the end of the reprogramming procedure, cells were stained with Tra-1-60 (embryonic cell surface marker) antibody. Quantifications were carried out with Image-J and results were compared to shControl sample with respect to Tra-1-60 positive colony number (**Figure 29**).

Both shRNAs targeting MLL1 and one shRNA targeting BAT3 significantly increased reprogramming efficiency in this screen (**Figure 29**). On the other hand, one shRNA of AF9, CDK9; and both shRNAs of HNRNPM significantly decreased reprogramming, while knock-down of other genes did not affect reprogramming significantly (**Figure 29**). MLL1 knock-down generated approximately 2-fold more iPSC colonies compared to shControl cells indicating that loss of MLL1 increases reprogramming (**Figure 29**). As a next step, the role of MLL1 on reprogramming was investigated.



**Figure 29. Reprogramming efficiency change as a result of shRNA-mediated gene silencing of DOT1L’s previously known interactors.**

Reprogramming experiment for shRNA screen was repeated for  $n=3$  times and for each experiment, samples were in triplicates. Average colony number of each experiment was calculated and normalized to shControl to calculate fold change. Average fold change from independent  $n=3$  experiments is shown and error bars indicate standard error. ( $P$  values were determined by a one sample t-test; \*,  $P < 0.05$ ) Representative Tra-1-60 stained well images are displayed underneath the bar graph.

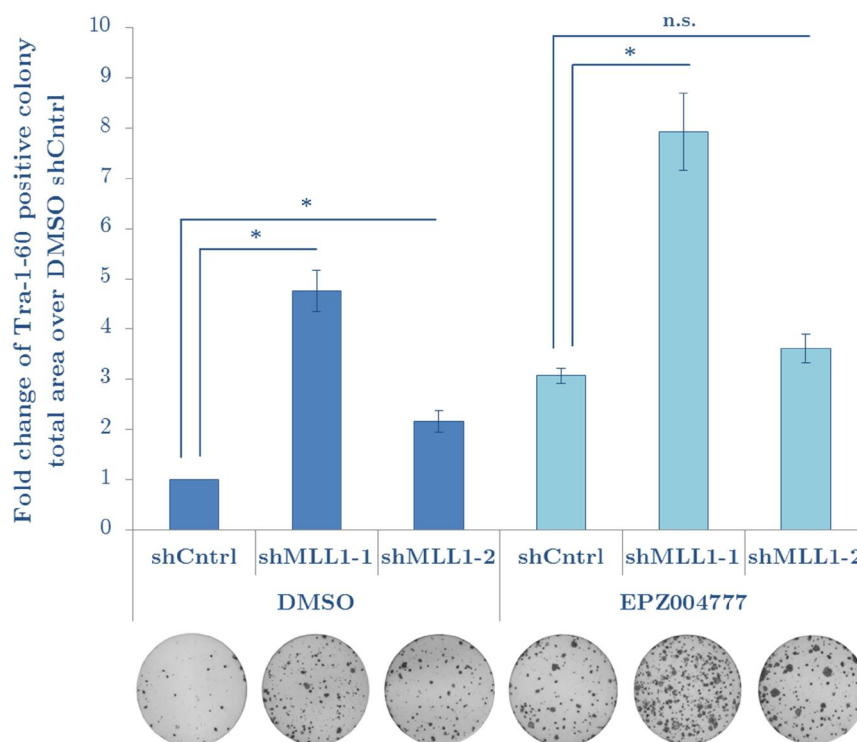
#### 4.2.2 MLL1 and DOT1L has independent roles during reprogramming

The strongest candidate protein identified from previous screen was MLL1 and it is known to be a functional interactor of DOT1L<sup>177</sup>. Follow up experiments were performed to better understand the mechanism of MLL1 during reprogramming. First, we considered the possibility that MLL1 blocks



reprogramming through DOT1L, similar to AF10. To test this hypothesis, shMLL1 treated cells were reprogrammed with or without DOT1L inhibitor (EPZ004777). shRNA infected dH1fs were treated with 3  $\mu$ M EPZ004777 for 6 days after OSKM transduction and control cells were treated with DMSO. At the end of reprogramming, total area of Tra-1-60 positive colonies was measured and their fold change over DMSO treated control cells was calculated (**Figure 30**). In this experiment, quantification was done by comparison of Tra-1-60 positive colonies total area instead of colony numbers; because EPZ004777 treated colonies had grown too much and were interfering with each other.

This result revealed that one of the shRNAs targeting MLL1 (shMLL1-1) significantly increased reprogramming 2-fold in DOT1L inhibitor-treated cells, while another shRNA (shMLL1-2) only slightly increased the efficiency (**Figure 30**). This result suggested that DOT1L and MLL1 may operate independently during the reprogramming process. Since we observed a strong phenotype in only one of the shMLL1s, this hypothesis was tested in alternative experimental set up. For this purpose, knock-out MLL1 cells were reprogrammed with DOT1L inhibitor treatment condition instead of shMLL1. 3 gRNAs were designed against *MLL1* to have more confident results.



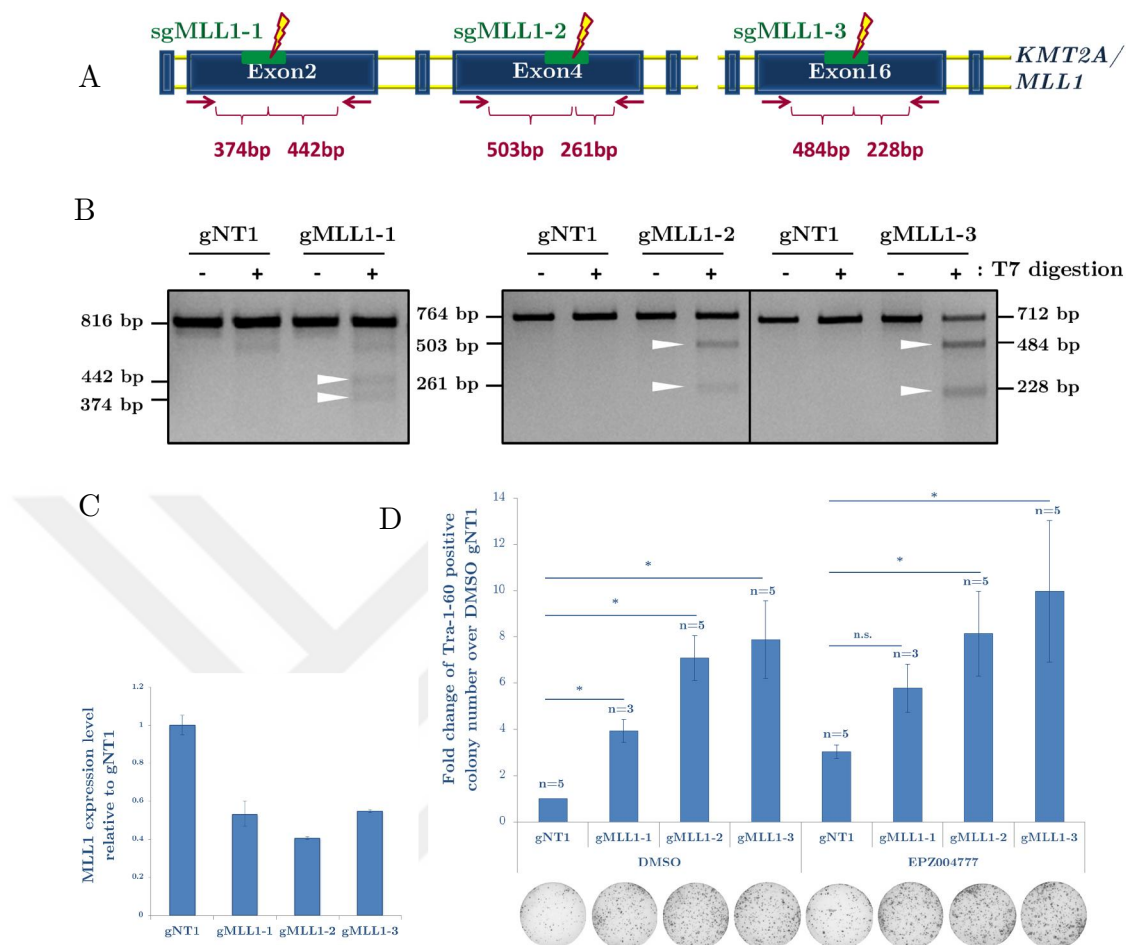
**Figure 30. Knock-down of MLL1 in DOT1L-inhibited condition further increases reprogramming efficiency**

shMLL1 infected cells were reprogrammed with DMSO or 3  $\mu$ M EPZ004777 treatment for 6 days. Average Tra-1-60 positive colonies total area of each experiment was calculated and normalized to shControl to calculate fold change. Average of fold change from  $n=7$  experiments were calculated and standard error is depicted in error bars. ( $n=7$  indicates independent biological replicates;  $P$  values were determined by a one sample t-test; \*,  $P < 0.05$ ) Representative Tra-1-60 stained well images are displayed underneath the bar graph.

MLL1 knock-out reprogramming was performed to verify the increase in the reprogramming efficiency via shMLL1s. CRISPR/Cas9 method was used to knock-out MLL1. *MLL1* genomic region was targeted by 3 sgRNAs (**Figure 31A**). Non-targeting control gRNAs (gNT1 and gNT2) and gMLL1s (gMLL1-

1, gMLL1-2 and gMLL1-3) were cloned into lentiCRISPRv2 plasmids and packed into lentiviruses. CRISPR/Cas9-mediated mutation of gMLL1 target sites was demonstrated via T7 endonuclease assay (**Figure 31B**). T7 endonuclease assay of gMLL1 infected cells exhibited expected band sizes depicted in **Figure 31A** (**Figure 31B**). In addition, gMLL1 infected dh1fs had lower *MLL1* mRNA levels compared to gNT1 dh1fs as assessed by qPCR (**Figure 31C**). These experiments show that *MLL1* can be successfully targeted by gRNAs.

gMLL1 infected dh1fs were reprogrammed via OSKM transduction with or without DOT1L inhibitor (EPZ004777). gMLL1 infected dh1fs were treated with 3  $\mu$ M EPZ004777 for 6 days after OSKM transduction and control cells were treated with DMSO. At the end of reprogramming, Tra-1-60 positive colonies were measured and their fold change over DMSO treated control cells (gNT1) was calculated. Three replicates of this reprogramming experiment were performed by Gülben Gürhan in Tamer Önder's group. This result revealed that all of the gMLL1s significantly increased reprogramming in DMSO, control condition whereas 2 out of 3 gMLL1s were significantly increased (2 to 3-fold) iPSC colony number in the DOT1L inhibitor treated cells (**Figure 31D**). Thus, it supports the hypothesis that DOT1L and MLL1 independently operate during reprogramming process. However, this result does not rule out the presence of overlapping functions of DOT1L and MLL1 since there is not any significant increase between DMSO and EPZ004777 treated gMLL1 samples.



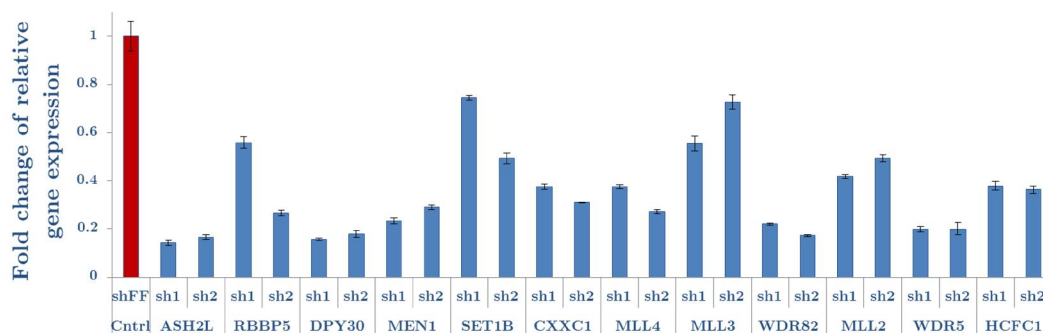
**Figure 31. Verification of MLL1 silencing experiments with knock-out MLL1 experiments**

**A**, gMLL1 targeting sites on *MLL1* (*KMT2A*) gene. Green bars shows the gRNA targeting site. Red arrows show the T7 assay primers and expected DNA fragments were depicted. **B**, T7-endonuclease assay for gMLL1 targeting sites. **C**, gRNA-mediated MLL1 knock-out decreases expression levels of *MLL1* mRNA. (n=2; error bars represents s.d.) **D**, gRNA-mediated MLL1 knock-out increases reprogramming with or without treatment with 3  $\mu$ M EPZ004777. (*P* values were determined by a one sample t-test; \*, *P* < 0.05; n is indicated above the bars and denotes independent biological replicates) Representative Tra-1-60 stained well images are displayed underneath the bar graph.

### 4.2.3 shRNA-mediated screen of MLL1 complex proteins for reprogramming efficiency

Having shown that MLL1 is a barrier to reprogramming, we next asked which proteins in MLL complex have an effect on reprogramming. MLL1 functions in large, multi-protein complexes in transcription regulation. We therefore sought to examine the role of MLL1-interacting proteins in reprogramming and investigated whether their inhibition would phenocopy MLL1 loss in reprogramming. An shRNA-mediated screen was planned to target MLL1 complex proteins in reprogramming. 12 proteins were curated as MLL1 complex proteins according to literature (ASH2L, RBBP5, DPY30, MENIN1, SET1B, CXXC1, MLL2, MLL3, MLL4, WDR82, WDR5, and HCFC1). Two shRNAs were designed to determine 12 proteins in MLL1 complex. All shRNAs were cloned into puromycin resistance gene containing, pSMP vector.

Successfully cloned shRNAs were packed with retroviral packaging plasmids to produce viruses. As a control shRNA (shControl), firefly luciferase targeting pSMP\_sh-Luc (shFF) was used. Successfully selected cells were collected for RNA isolation and followed by q-RT-PCR to quantify the expression levels of shRNA targeting genes. While quantifying the results, mRNA levels of each gene was compared with shControl infected cells. As a result of qRT-PCR experiments, all shRNAs were successfully knocked-down their target gene (**Figure 32**).



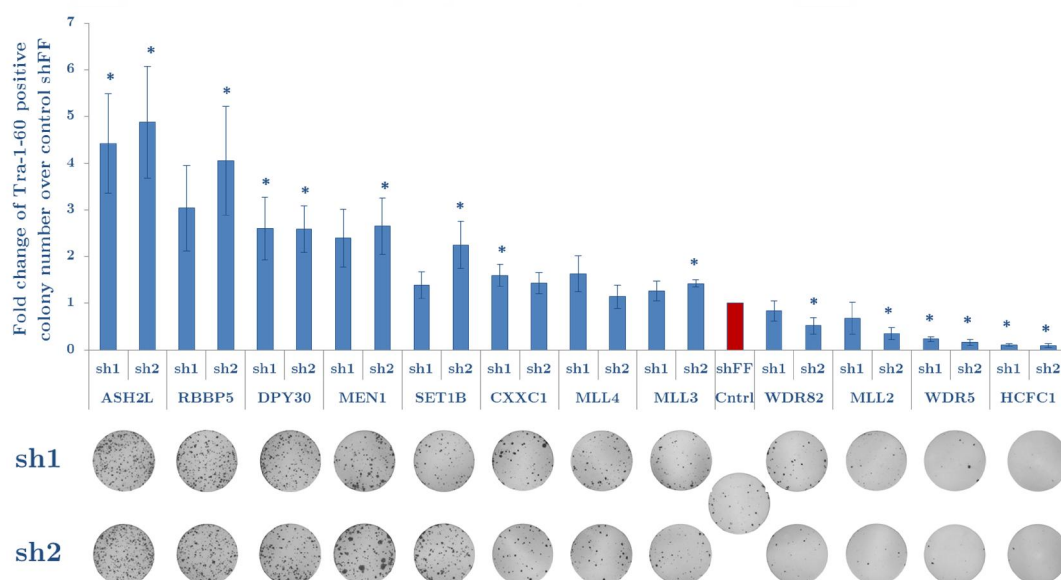
**Figure 32. mRNA expression levels of shRNA-targeted genes belonging to MLL1 complexes**

qRT-PCR was performed as duplicate samples and  $\beta$ -actin was used as an internal control gene. Every genes' expression level is normalized to shControl (shFF) infected cells. (n=2; error bars represents s.d.)

shRNA infected dH1f cells were reprogrammed with OSKM to test the loss-of function effect of MLL complex proteins on reprogramming. At the end of the reprogramming, cells were stained with Tra-1-60 (embryonic cell surface marker) antibody. Quantifications were carried out with Image-J and results were compared to shControl sample with respect to Tra-1-60 positive colony number.

As a result of this shRNA screen, it has been concluded that some of the proteins in MLL1 complex (ASH2L, RBBP5, DPY30, MEN1, SET1B and CXXC1) act as a barrier of reprogramming similar to MLL1. Whereas, WDR5 and HCFC1 are essential for reprogramming since the reprogramming efficiency drastically decreases when they are knocked-down (**Figure 33**). These results can be interpreted as knock-down of MLL1 core complex

increases the efficiency, except WDR5 and HCFC1. These exceptions may be due to the differential functions of WDR5 and HCFC1 since WDR5 interacts with OCT4<sup>76</sup> and HCFC1 interacts with Sin3 histone deacetylase (HDAC)<sup>178</sup>. This result reveals that MLL1 complex has an undeniable effect on reprogramming since the complex consist of both enhancers and barriers of reprogramming.



**Figure 33. The effect of knocking down MLL1 complex members on reprogramming efficiency**

Reprogramming experiment for shRNA screen was repeated for  $n=4$  times and triplicate technical replicates for each experiment were used. Average of each experiment was calculated and normalized to shControl to calculate fold change. Average of fold change from different experiments was calculated and standard error was depicted in error bars. ( $P$  values were determined by a one sample t-test; \*,  $P < 0.05$ ;  $n=4$  and denotes independent biological replicates) Representative Tra-1-60 stained well images are displayed underneath the bar graph.

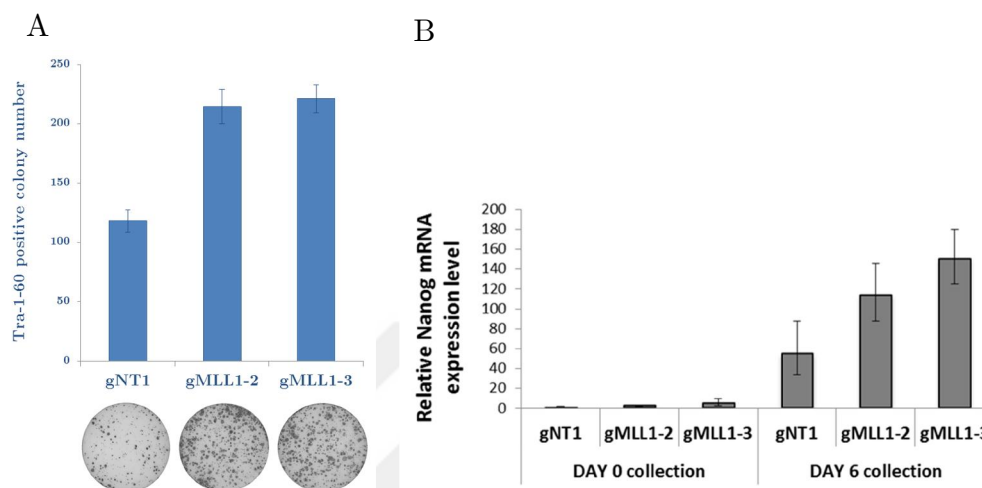
#### 4.2.4 gMLL1s change gene expression to accelerate fibroblast to iPSC transition

MLL complex is known to play an important role during transcriptional regulation<sup>157</sup>. Since we have shown that MLL1 as a complex has a role during reprogramming, we next determined the global gene expression differences upon MLL1 downregulations. gMLL1-2 and gMLL1-3 were picked for further experiments since they increase reprogramming efficiency more than gMLL1-1 (Figure 31D). RNA-sequencing (RNA-seq) was performed on gMLL1 expressing cells with and without reprogramming factor transduction (OSKM).

dH1f cells were infected twice with gNT1, gMLL1-1 and gMLL1-2 and selected with puromycin. Successfully selected cells were passaged for 10 days and then seeded onto triplicate wells in two replicate plates. Samples in one of the plates were infected with OSKM while other samples were not reprogrammed and referred as Day 0 samples of RNA-seq experiment. At day 6 of reprogramming, infected cells were trypsinized and 1/6 of them were transferred on MEFs to continue the reprogramming experiment while the remaining cells were harvested as Day 6 samples of RNA-seq experiment. At the end of reprogramming, gMLL1's reprogramming efficiency was analyzed via Tra-1-60 staining (**Figure 34A**). As a result, almost 2-fold increase in Tra-1-60 positive colony number was observed. RNA isolation was performed on previously harvested cell pellets (Day 0 and Day 6 samples). Firstly, cDNA synthesis and q-RT-PCR analysis were performed for quality control of RNA samples. *NANOG* mRNA levels were checked to confirm that



isolated RNA samples represent the increase in reprogramming efficiency (Figure 34B).

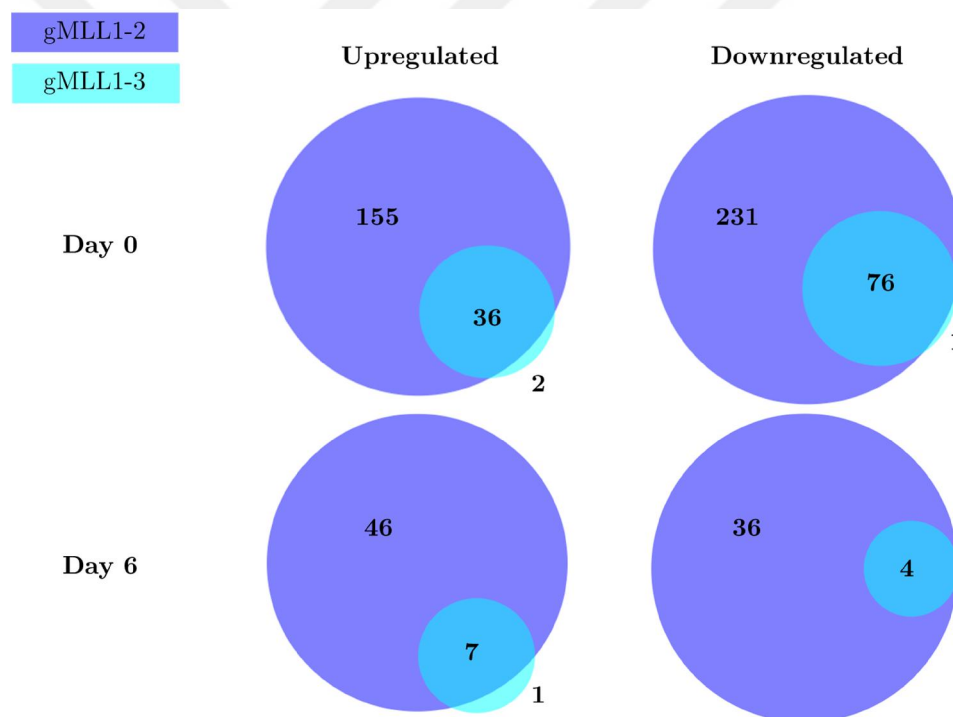


**Figure 34. Validation of reprogramming phenotype of cells used for RNA-sequencing analysis**

**A**, Tra-1-60 colony numbers were quantified via Image J software. gRNA mediated MLL1 knock-out increases reprogramming. Representative Tra-1-60 stained well images are displayed underneath the bar graph. (error bars represents standard error of  $n=3$  technical replicates) **B**, OSKM treated cells have increased expression levels of *NANOG* mRNA at Day 6. ( $n=2$ ; error bars represents s.d.)

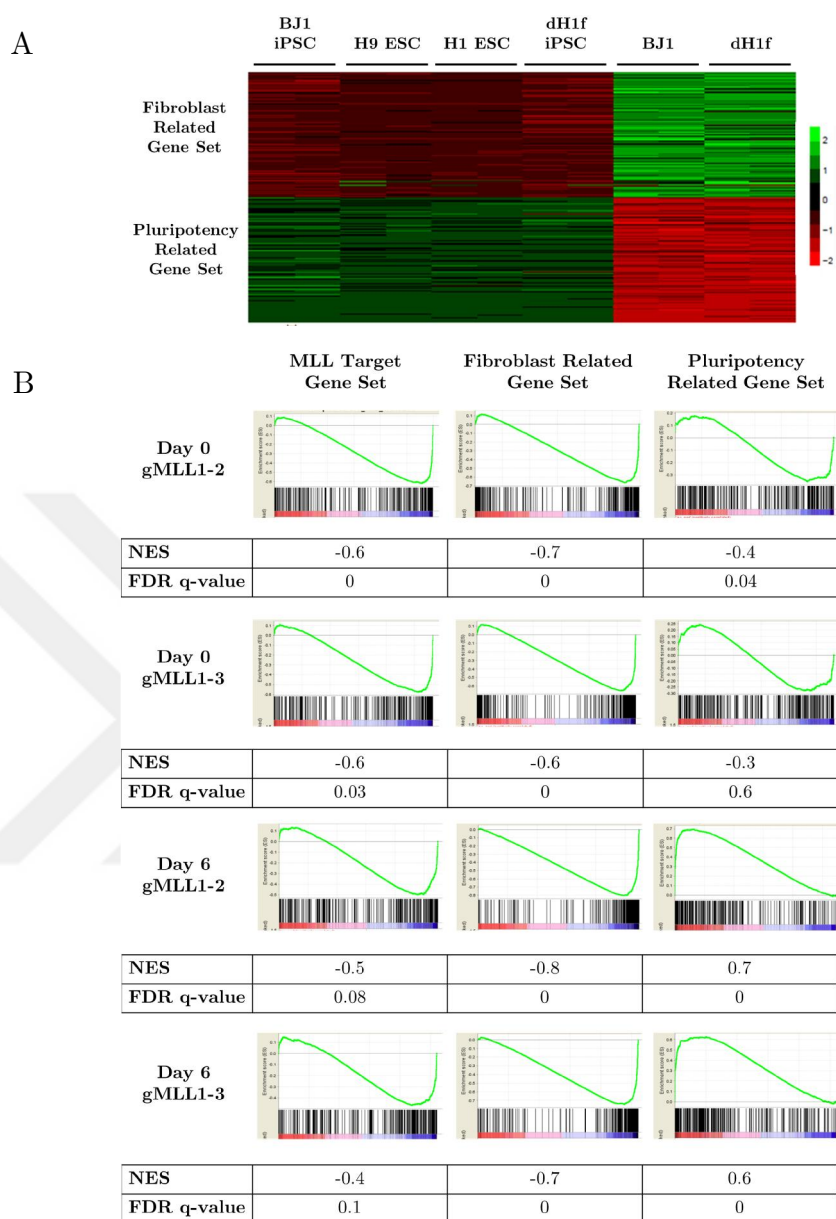
After quality control of RNA samples; RNA sequencing and its analysis were performed by Kenan Sevinç and Tunc Morova. After RNA sequencing result were obtained, differentially expressed genes among gMLL1 treated cells were identified via DESeq2 software. Significant genes were selected if their  $p$ -value  $< 0.05$  and upregulated-downregulated genes were separated if their  $\log_2\text{fold\_change} > 0.5$  or  $\log_2\text{fold\_change} < -0.5$ , respectively. According to these cutoffs, upregulated and downregulated genes of gMLL1-3 are identified

as a subset of gMLL1-2 results. At day 0, 36 genes out of 38 upregulated genes of gMLL1-3 were common with 191 upregulated genes of gMLL1-2; 76 genes out of 77 downregulated genes of gMLL1-3 were common with 307 downregulated genes of gMLL1-2. At day 6, 7 genes out of 8 upregulated genes of gMLL1-3 were common with 53 upregulated genes of gMLL1-2; all 4 downregulated genes of gMLL1-3 were common with 42 downregulated genes of gMLL1-2 (Figure 35). Similar comparison was also performed with p-adjusted values  $<0.05$  and differentially expressed genes were separated if their  $\log_2\text{fold\_change} >0$  or  $\log_2\text{fold\_change} <0$  (Appendix-I; Figure 42).



**Figure 35.** Commonly upregulated and downregulated gene numbers among gMLL1-2 and gMLL1-3 samples at Day 0 and Day 6 samples. Differentially expressed genes were selected if their  $p\text{-value} < 0.05$ . Upregulated genes were selected if their  $\log_2\text{fold\_change} > 0.5$  and downregulated genes were selected if  $\log_2\text{fold\_change} < -0.5$ .

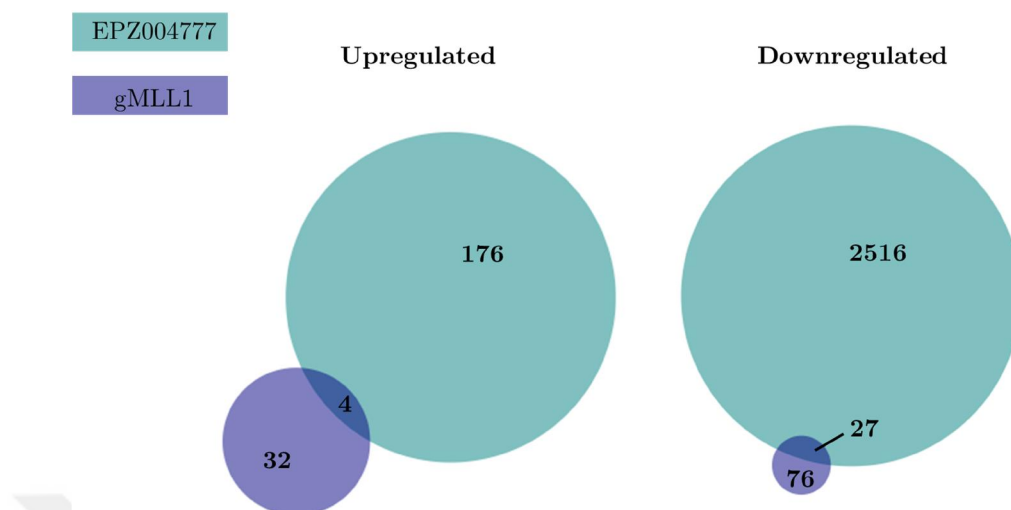
Differentially expressed genes upon gMLL1 infection were analyzed with gene set enrichment analysis (GSEA) to investigate the statistically significant correlations of these genes with set of fibroblast-specific genes, pluripotency-specific genes and MLL-target genes. First, a set of fibroblast and iPSC specific genes were determined by analyzing microarray data to compare fibroblasts with ESCs and iPSCs (**Figure 36A**). Then, established gene sets were related with differential genes between gNT1 and gMLL1 samples (**Figure 36B**). Also, publically available MLL target gene set (Wang\_MLL\_Targets)<sup>136</sup> was compared with gMLL1 samples (**Figure 36B**). Same analysis was performed for samples that are induced with OSKM and these are referred as Day 6 samples (**Figure 36C**).



**Figure 36. MLL1-KO accelerates the decrease in fibroblast gene sets and promotes the increase in pluripotency gene expression**

**A**, Microarray analysis of fibroblasts and pluripotent cells to determine fibroblast related genes and pluripotency related gene sets. GEO data series GSE55679. **B**, RNA sequencing analysis by using MLL target genes, fibroblast related genes and pluripotency related gene sets. (NES, normalized enrichment score; FDR, false discovery rate)

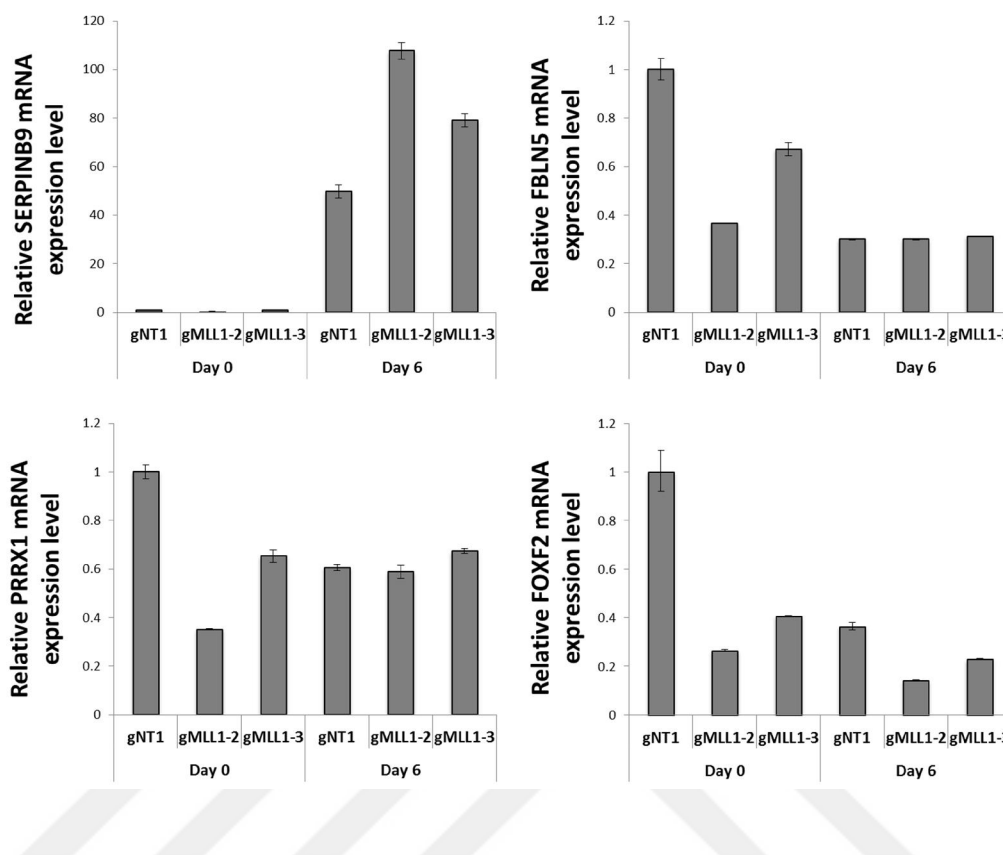
As a result of RNA-sequencing experiments, gMLL1 infected cells display accelerated decrease in fibroblast gene sets and increase in pluripotency gene expression. When we compared differentially expressed genes of gMLL1 treatment with DOT1L inhibitor treated cells, only a small number of genes were appeared as common (RNA-sequencing data not shown for DOT1L inhibitor treated cells). To be exact; among 76 downregulated genes, 27 of them was common with 2543 downregulated genes of DOT1L inhibitor treated cells and among 36 upregulated genes, only 4 of them was common with 180 upregulated genes of DOT1L inhibitor treated cells (Figure 37). This result supports the hypothesis that DOT1L and MLL1 independently operate during reprogramming process. Similar analysis was performed with p-adjusted values  $< 0.05$  and differentially expressed genes were separated if their  $\log_2\text{fold\_change} > 0$  or  $\log_2\text{fold\_change} < 0$  (Appendix-I; Figure 43). In addition, each gMLL1 infected samples were compared with EPZ004777 samples (Appendix-I; Figure 44).



**Figure 37. Commonly upregulated and downregulated gene numbers among gMLL1 and EPZ004777 samples at Day 0**

Differentially expressed genes were selected if their  $p$ -value  $< 0.05$ . Upregulated genes were selected if their  $\log_2$ fold\_change  $> 0.5$  and downregulated genes were selected if  $\log_2$ fold\_change  $< -0.5$ .

To validate the RNA-sequencing results, a few commonly downregulated (FBLN5, PRRX1, and FOXF2) or upregulated (SERPINB9) genes were analyzed with q-RT-PCR (**Figure 38**). As a result of q-RT-PCR validation of RNA-sequencing samples, SERPINB9 demonstrated increasing levels of mRNA with gMLL1 samples and FBLN5, PRRX1 and FOXF2 exhibits decreasing levels as expected from RNA-sequencing results (**Figure 38**). These set of experiments validate the results of the RNA sequencing analyses.



**Figure 38. RNA-seq validation with q-RT-PCR analysis to compare SERPINB9, FBLN5, PRRX1 and FOXF2 gene levels**

qRT-PCR was performed as duplicate samples and  $\beta$ -actin was used as an internal control gene. Every genes' expression level is normalized to gNT1 infected cells. (n=2; error bars represents s.d.)

Taken together, experimental work described in this chapter indicated that these data are the first to demonstrate that knock-down of MLL1, significantly increased iPSC generation efficiency, suggesting that it acts as a barrier to reprogramming similar to DOT1L. This finding was verified by CRISPR/Cas9 mediated knockout of MLL1. In addition we showed that combining DOT1L inhibition with MLL1 knockout resulted in an increased enhancement of reprogramming, suggesting that these two chromatin factors

act independently. Our findings suggest that MLL1 suppression resulted in downregulation of fibroblast-specific genes and accelerated the activation of pluripotency-related genes.





## Chapter 5

### 5 Discussion

DOT1L is the sole H3K79 methyltransferase and it catalyzes mono-, di-, and tri-methylation<sup>177</sup>. It is also known that H3K79 methylation is linked to active gene transcription<sup>177</sup>. DOT1L catalyzes this modification during elongation of transcription and it works in a complex during transcriptional regulation<sup>105</sup>. DOT1L is an established barrier of reprogramming<sup>55</sup>. It is known that H3K79 methylation can affect lineage specific genes' expression; therefore, knock-down or inhibition of DOT1L increases reprogramming efficiency<sup>55</sup>. However, other proteins in DOT1L complex have not been investigated for their roles on reprogramming. In this study, our aim was to identify novel interaction partners of DOT1L and investigate their effect on reprogramming. In this thesis, three main findings have been presented: (1) Novel proximal interactors of DOT1L were found via BioID assay, (2) AF10 was found as a barrier of reprogramming through regulation of H3K79 methylation with DOT1L, and (3) MLL1 was found as a barrier of reprogramming through regulation of gene expression.

## 5.1 Proximal interactors of DOT1L was identified by BioID method

AF9, AF10, and ENL are the proteins that were identified as direct interactors of DOT1L<sup>105,121,179</sup>. Later, DOT1L's interactors were identified by affinity pull-down method and these interactions were verified by co-IP experiments<sup>125</sup>. In that study<sup>125</sup>, interaction of DOT1L with AF9, AF10 and ENL was verified and a few novel proteins were suggested as DOT1L interactors; including NPM1, HNRNPM, DDX21, etc. In our study, we focused on finding DOT1L's interactions as a complex rather than direct interactors. Therefore, we chose a method that will reveal proximal interaction partners: BioID.

BioID is a powerful method to find out proximal protein interactions of a particular protein<sup>139</sup>. In our study, we aimed to identify DOT1L's proximal interactors, therefore, we cloned a fusion protein of BirA\* (promiscuous biotin ligase) and DOT1L. This method allowed us to biotinylate proximal proteins that are in 10 nm radius so that we can pull-down these proteins with Streptavidin beads and identify via LC-MS/MS. As a result of this analysis, some of the known interactors (AF10, AF17, ENL, DDX21) of DOT1L were acquired as well as some novel interactions were suggested (TPR, NONO, NUMA, KAISO, MRE11, SIN3B, Histone H1.0).

Since BioID captured some of the known interactions of DOT1L, we can say that BioID method successfully worked in our experimental design. However, not all of the known interactions were found in LC-MS/MS analysis. This can be due to the algorithm that we used for identification of interactions. For

example, NPM1 and HNRNPM proteins, which are also known interactions of DOT1L<sup>125</sup>, were detected via LC-MS/MS in BirA\*-DOT1L samples; however, they were eliminated from the list because they were also detected in control/background sample (Appendix: BioID raw data). In the case of NPM1, detected PSM values were similar in both samples and control. On the other hand, HNRNPM was highly enriched in both reads of BirA\*-DOT1L-mut sample (1<sup>st</sup> reading 34 PSM, 2<sup>nd</sup> reading 37 PSM), whereas it was detected in only one reading of BirA\*-DOT1L-wt sample and control sample with a low PSM value. This result suggests that HNRNPM might be an interaction partner of DOT1L when DOT1L is catalytically inactive. Another possibility is that, DOT1L mutant might have a different set of interactors in the cell as a result of an aberrant localization of mut-DOT1L on chromatin. Moreover, AF9 is a well-established interaction partner of DOT1L, but it was not in the final list of DOT1L-proximal interactors. In LC-MS/MS analysis, AF9 was detected in only one of the readings of BirA\*-DOT1L-wt sample. Therefore, we eliminated this interaction when we selected the proteins that are common in both readings. This can be interpreted as follows: *some of the novel interactions may be eliminated due to the stringent scanning that was used.* On the other hand, we revealed 4 of the known interactions of DOT1L along with 7 possible novel proximal interactions, even though we used this stringent scanning. The scanning that we followed increased the confidence in our suggested novel interactions despite a few possible interactions might have eliminated during analysis.

When we analyze the BioID results, DOT1L was the top hit detected in samples and none was detected in control cells. This is the first step that shows the BioID method worked efficiently in our study. During BioID assay,

proteins other than DOT1L were at their basal levels in the cell since they were not overexpressed. Therefore, we might have missed some of the low expressed proteins that are interacting with DOT1L. One way of overcoming this issue might be repeating the assay. In this study, we performed BioID assay once with 2 readings of each sample. When replicated, the confidence of detection could be increased.

In this study, we fused BirA\* to the N-terminus of DOT1L. Even though BirA\* is a relatively small protein with 35 kDa weight, it might interfere with the N-terminus specific interactions of DOT1L. Alternatively, another fusion protein could be used for BioID assay. If BirA\* was fused to the C-terminus of DOT1L, possible interactions of DOT1L that requires intact N-terminus structure could be captured. Similarly, BirA\* might interfere with the catalytic activity of DOT1L since the catalytic domain of DOT1L is on the N-terminus. However, we tested the catalytic activity of DOT1L and found that it did not lose its H3K79 methyltransferase activity. Since fusion of BirA\* was close to the catalytic domain, interactions of DOT1L that are important for its catalytic activity could be identified.

In future work, the novel proximal proteins identified here can be investigated to understand the interaction dynamics whether they directly interact or occupy the same complex or transiently interact at a certain time point. For this purpose, a few techniques can be performed such as co-immunoprecipitation (co-IP), mammalian two hybrid, proximity ligation assay, etc. It would be insightful to understand the details of interactions that DOT1L has.

## 5.2 AF10 blocks reprogramming through regulation of H3K79 methylation via DOT1L

Our shRNA-mediated screen against DOT1L-proximal proteins revealed that AF10 is a barrier of reprogramming. Although AF10's recruitment of DOT1L to the target genes<sup>130</sup> was known before, AF10's effect on reprogramming has not been investigated. In this study, we showed the increase in reprogramming when AF10 is silenced via shRNA or sgRNA. Rescue of AF10-KO with sgAF10-resistant overexpression plasmid indicates that the increase in reprogramming is indeed through AF10-loss and not via off-target effect of sgAF10s. When combinatory knock-down of AF10 with other DOT1L-interacting genes ENL and AF17 was tried, reprogramming efficiency was not increased more than only AF10 knock-down (Appendix-I, Figure 40). In addition, our results suggest that AF10's effect on reprogramming is through DOT1L. It was reported that DOT1L inhibition decreases H3K79 methylation and increases reprogramming<sup>55</sup>. Previously, it was demonstrated that AF10 has an important role in the regulation of H3K79 methylation<sup>130</sup>, as well. These findings are consistent with our observation that AF10 silencing increased reprogramming through DOT1L. On the contrary, sgAF10 treated cells does not exhibit decreased levels of expression in DOT1L-target genes (Appendix-I, Figure 41).

In shRNA-mediated screen of DOT1L-proximal interactions on reprogramming, knock-down of NONO was increased reprogramming significantly. In the future, effect of NONO on reprogramming can be investigated. Previously, it was reported that loss of Nono had a significant role in mouse ESC self-renewal through Erk pathway activation<sup>180</sup>.

Therefore, an evolutionarily conserved effect of NONO could be revealed since our reprogramming experiments were conducted with human fibroblasts. NONO is found in RNA regulatory paraspeckles in nucleus of differentiated cells<sup>181</sup>. However, paraspeckle formation is not observed in ESCs<sup>182</sup>. Therefore, it is correlated to observe that loss of NONO increasing the reprogramming. These observations and our results are consistent with the finding that NONO is a possible barrier in reprogramming. Thus, further investigation of NONO's effect on reprogramming could be a promising future project.

In this study, it was demonstrated that knock-down of MRE11 and TPR significantly decreased iPSC generation in shRNA-mediated screen of DOT1L-proximal interactions on reprogramming. TPR is a member of nuclear pore complex and its function is to regulate mRNA transportation to the cytoplasm via regulating the quality control of spliced mRNAs<sup>183</sup>. However, TPR's effect on reprogramming or pluripotency had not been investigated before. MRE11 is a DNA damage repair protein that has a functional role in the cell against double stranded breaks (DSBs)<sup>184</sup>. Another report shows that fibroblasts have lower levels of MRE11 while hESCs and iPSCs have elevated levels of MRE11<sup>185</sup>. These observations are consistent with our finding that shows knock-down of MRE11 significantly decreased iPSCs generation.

Taken together, findings in our study suggest that knock-down of AF10 increases reprogramming through DOT1L by regulation of H3K79 methylation and knock down of NONO also facilitates reprogramming. Conversely, knock-down of MRE11 and TPR impede reprogramming,

however we do not know by which mechanism these proteins operate during reprogramming. Nevertheless, interaction partners of DOT1L has a certain role during reprogramming, therefore, it was important to pursue their effect on reprogramming.

Recently, it was reported that KDM4D might be responsible for H3K79 demethylation<sup>120</sup>. Before that study, H3K79 specific demethylase had not been identified. Since it is known that H3K79 methylation is important for reprogramming process and now, we are suggesting that regulation of H3K79 methylation has a role during reprogramming; it would be worthwhile to investigate the effect of KDM4D on reprogramming, as well.

### **5.3 MLL1 blocks reprogramming through regulation of gene expression independent from DOT1L**

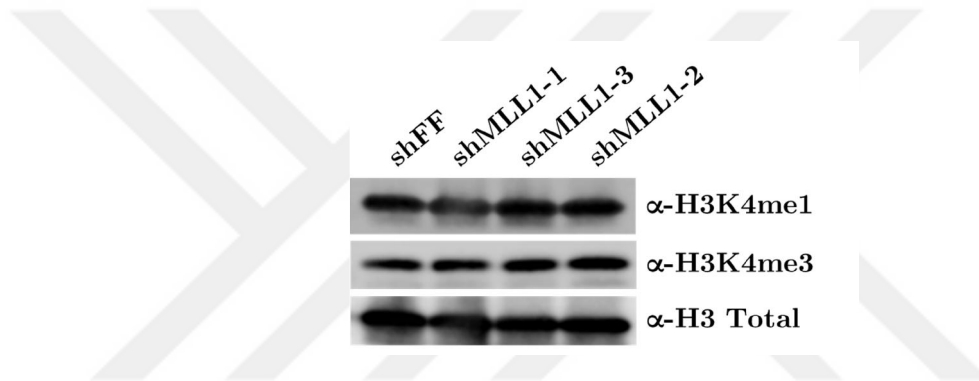
As result of BioID assay, a few known interactions of DOT1L were detected along with a few novel proximal interactions of DOT1L. However, in the literature there are a few more known interactions of DOT1L: some are direct interactions (AF4, AF9, HNRNPM, NPM1, STAT1, BAT3) and some of them are functional interactions (CDK9, MLL1, SIRT1). Since we have obtained crucial information by investigating the BioID-identified DOT1L proximal interactors on reprogramming; we decided to adopt a similar strategy with known interactions of DOT1L. As a result of this shRNA-mediated screen of reprogramming, knock-down of MLL1 demonstrated a significant increase in iPSC generation.

Knock-down of MLL1 was associated with the increased pluripotency for the first time; therefore, further investigation to find out how knock-down of MLL1 increased reprogramming was carried out. First, we tested whether MLL1 is functioning through DOT1L and the results indicated that MLL1 and DOT1L appear to have overlapping as well as distinct functions in reprogramming. Then, we investigated the effect of proteins in MLL1 complex during reprogramming process. Our findings indicated that knock-down of core components (ASH2L, DPY30, RBBP5) of MLL complex increased reprogramming. This finding contradicts with a previous study, where they demonstrated that knock-down of Rbbp5 and Dpy30 decreased reprogramming in MEFs<sup>175</sup>. These adverse outcomes could be as a result of differential epigenetic regulations between mouse and human. Nevertheless, further investigation is required to make a definite conclusion on this issue. Knock-down of WDR5, another core protein of MLL complex, decreased reprogramming significantly. In fact, it was previously known that WDR5 is essential for self-renewal of ESCs and therefore it is also essential for reprogramming<sup>76</sup>. It could be due to the fact that WDR5 interacts with OCT4<sup>76</sup>. Knock-down of HCFC1 also decreased reprogramming significantly. This could be as a result of HCFC1's interaction with histone deacetylase, Sir3<sup>178</sup>. Taken together, these findings suggest that MLL1 complex has an important role during reprogramming.

MLL1 complex is known to play an important role during transcriptional regulation. Catalytic function of MLL1 complex is associated with H3K4 methylation<sup>169</sup>. In previous reports, H3K4 methylation was also associated with lineage specificity and demonstrated that down-regulation of H3K4 methylation promote pluripotency<sup>186,187</sup>. In one study, H3K4 tri-methylation



was suggested as a barrier to efficient nuclear reprogramming due to its ability to maintain the memory of somatic cell identity<sup>186</sup>. In another report, downregulation of H3K4 methylation via MLL1 inhibitor was associated with improved somatic cell nuclear transfer<sup>187</sup>. However, in our study we did not observe a significant decrease in the global levels of H3K4me1, 2 or 3 when MLL1 is knocked-down (Figure 39). Therefore, we decided to investigate the action of MLL1 complex through comparing the gene expression dynamics.



**Figure 39. MLL1 knock-down does not change global H3K4 mono- and tri-methylation in dH1fs**

H3 Total was used as a loading control.

RNA sequencing was performed to gMLL1-2, gMLL1-3 and gNT1 infected cells, with or without OSKM transduction. As a result of RNA sequencing analysis, it was shown that gMLL1 treated cells exhibited down-regulated expression of MLL1-target genes which suggests that MLL1 silencing successfully achieved. When differentially expressed genes of gMLL1 treated cells were compared with fibroblast-specific gene sets, significant decrease was observed in the fibroblast-specific genes' expression. In addition, when

gMLL1 infected cells were reprogrammed, pluripotency-specific genes' expression was significantly enriched in differentially expressed genes. These two observations suggest that MLL1 might promote reprogramming by accelerating the transition of fibroblasts to pluripotency state by down-regulating the fibroblast-specific gene expression and accelerating the activation of pluripotency-related genes' expression. Furthermore, when down-regulated genes upon gMLL1 were compared with down-regulated genes upon DOT1L inhibition, only a subset of genes was common. This observation also supports our previous claim that MLL1 and DOT1L have distinct functions as well as overlapping functions.

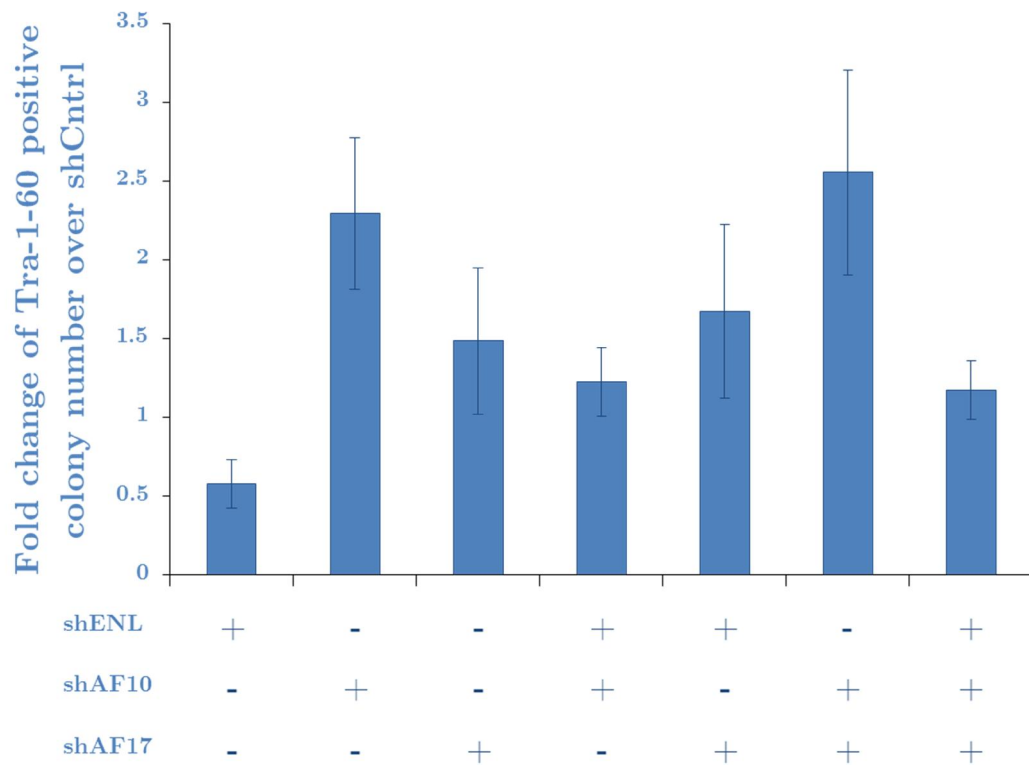
Taken together, findings in this study suggest that silencing MLL1 increased reprogramming through regulation of fibroblast-specific genes expression. Similar to MLL1, knock-down of MLL1 complex proteins, especially ASH2L and DPY30, was demonstrated to facilitate reprogramming. Conversely, knock-down of WDR5 and HCFC1 was inhibited the iPSC generation. These results suggest that MLL1 as a complex has an important regulatory role during reprogramming. In the future, chromatin immunoprecipitation (ChIP) experiments can be performed to determine the localization of MLL1 marks, H3K4 methylation, during reprogramming.

## 5.4 Final conclusion

In this thesis, we showed that DOT1L's interaction partners have an effect on reprogramming. First, we identified NONO, KAISO, NUMA1, TPR, MRE11, and SIN3B as novel proximal interactions of DOT1L via BioID

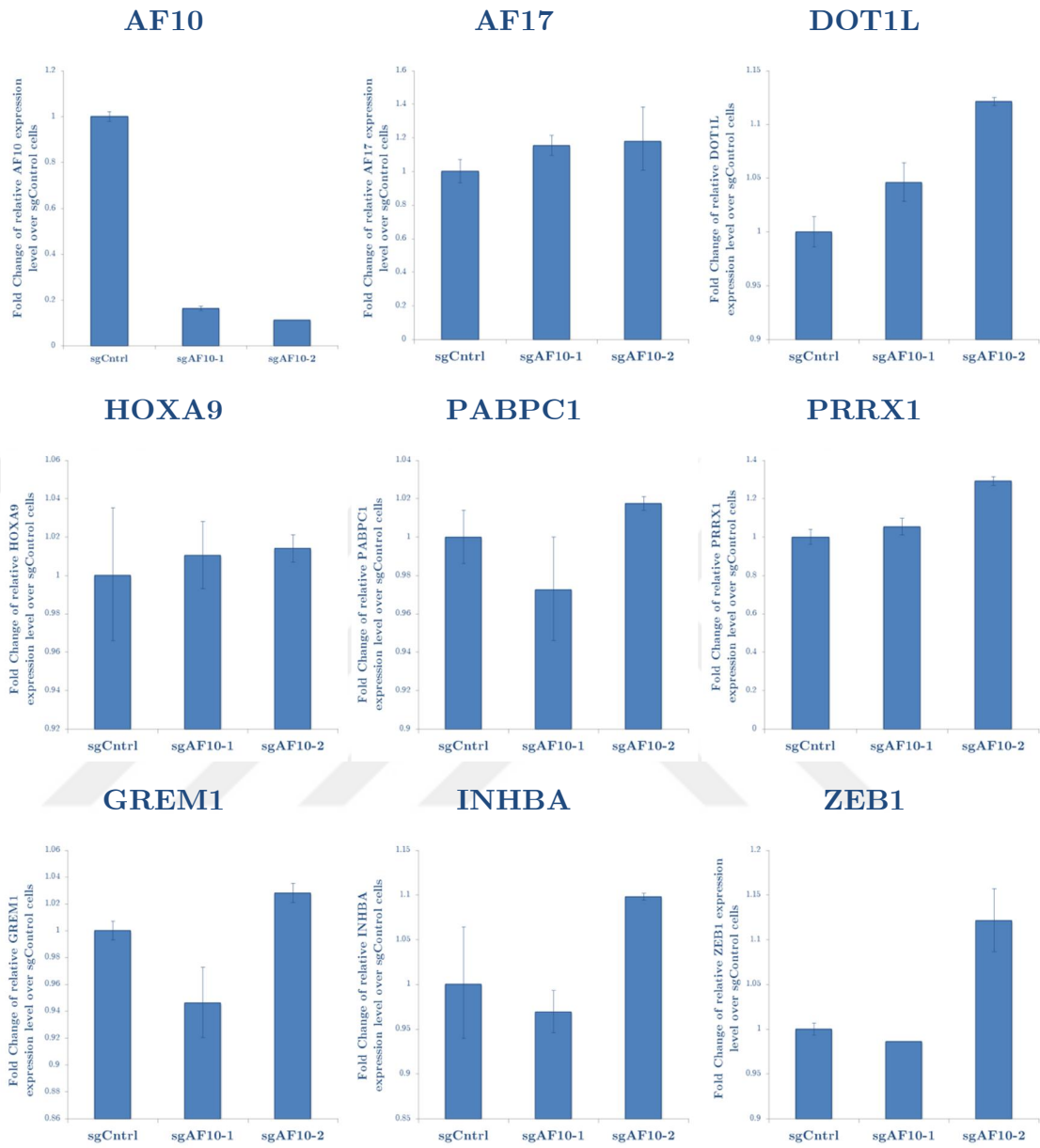
method. Several known interactions of DOT1L (AF10, AF17, ENL and DDX21) were detected via BioID method. Second, we investigated the effect of proximal interactions of DOT1L on reprogramming. As a result, AF10 was found as a barrier to reprogramming since shRNA or sgRNA-mediated silencing of AF10 increased reprogramming efficiency. Our findings suggest that AF10 increases reprogramming through DOT1L. Third, we tested the effect of known interaction partners of DOT1L on reprogramming. Consequently, MLL1 was identified as another barrier to reprogramming. It was also proposed that MLL1, as a complex, regulates gene expression of cells. sgRNA-mediated silencing of MLL1 decreased fibroblast-specific gene set and promoted pluripotency genes' expression. Taken together, we identified novel proximal interactors of DOT1L and revealed that AF10 and MLL1 are novel barriers to reprogramming.

## APPENDIX-I

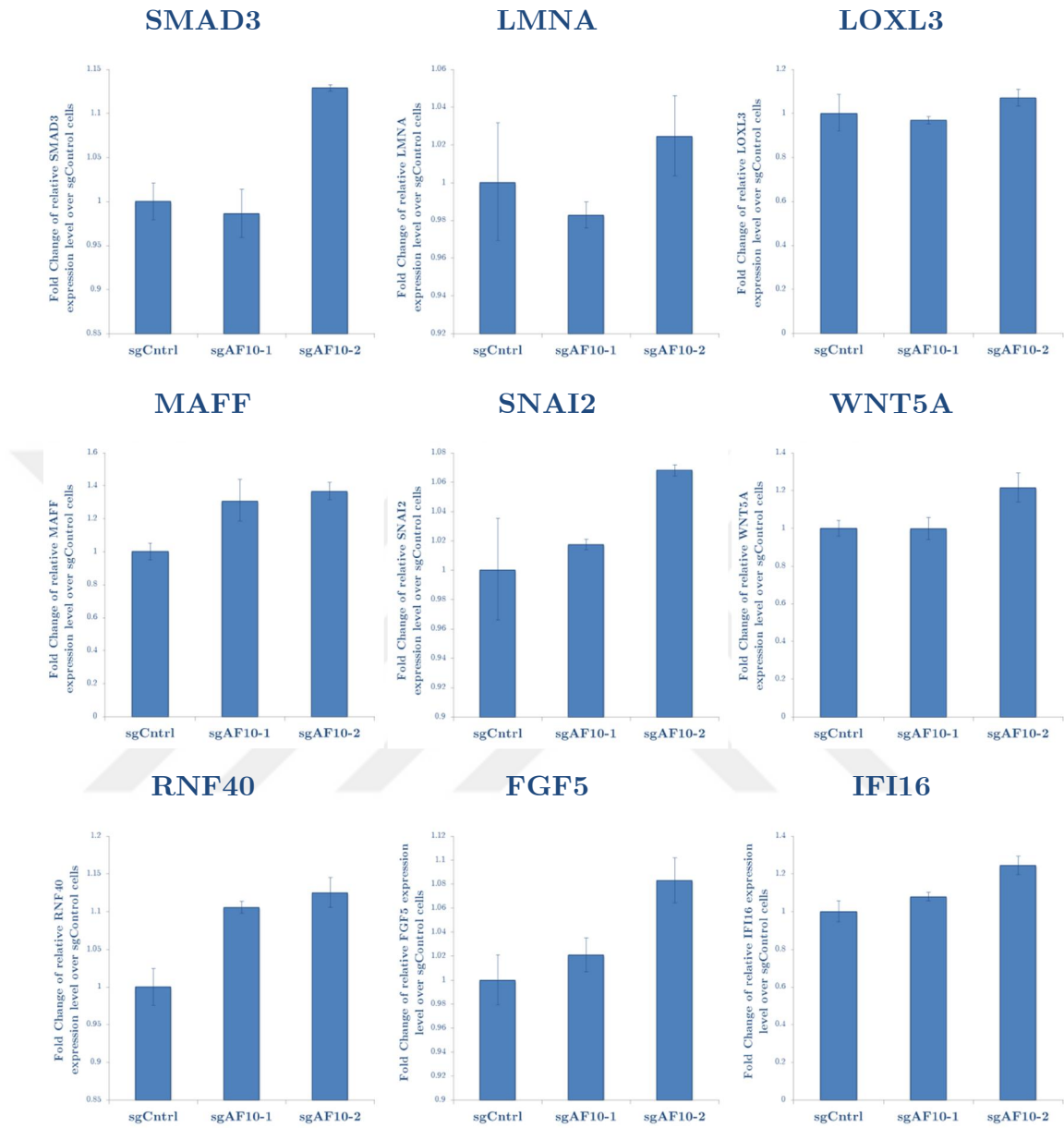


**Figure 40. Fold change in reprogramming efficiency as a result of combination of shRNA-mediated silencing of ENL, AF10 and AF17**

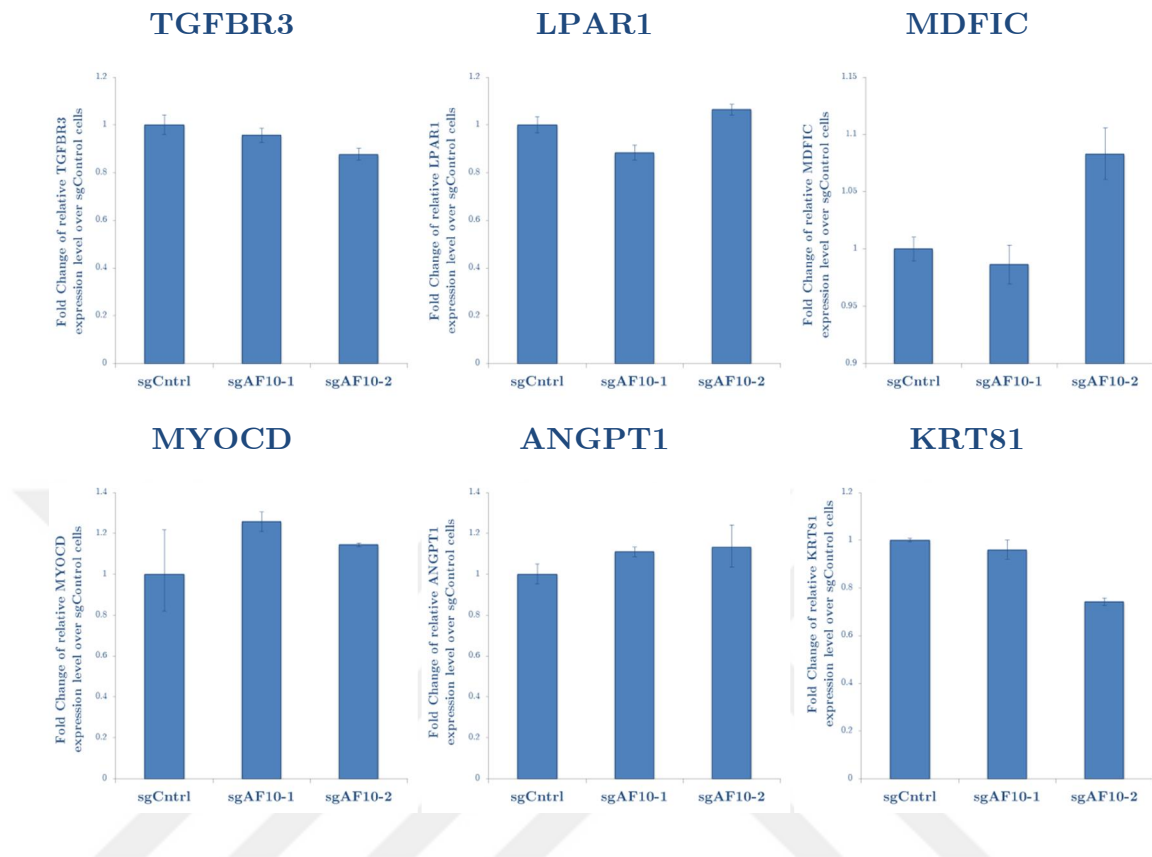
Reprogramming experiment for shRNA infected dH1fs were performed. Average colony number of each experiment was calculated and normalized to shControl to calculate fold change. Average of fold change from different experiments was calculated and standard error is depicted in error bars.



*continues...*

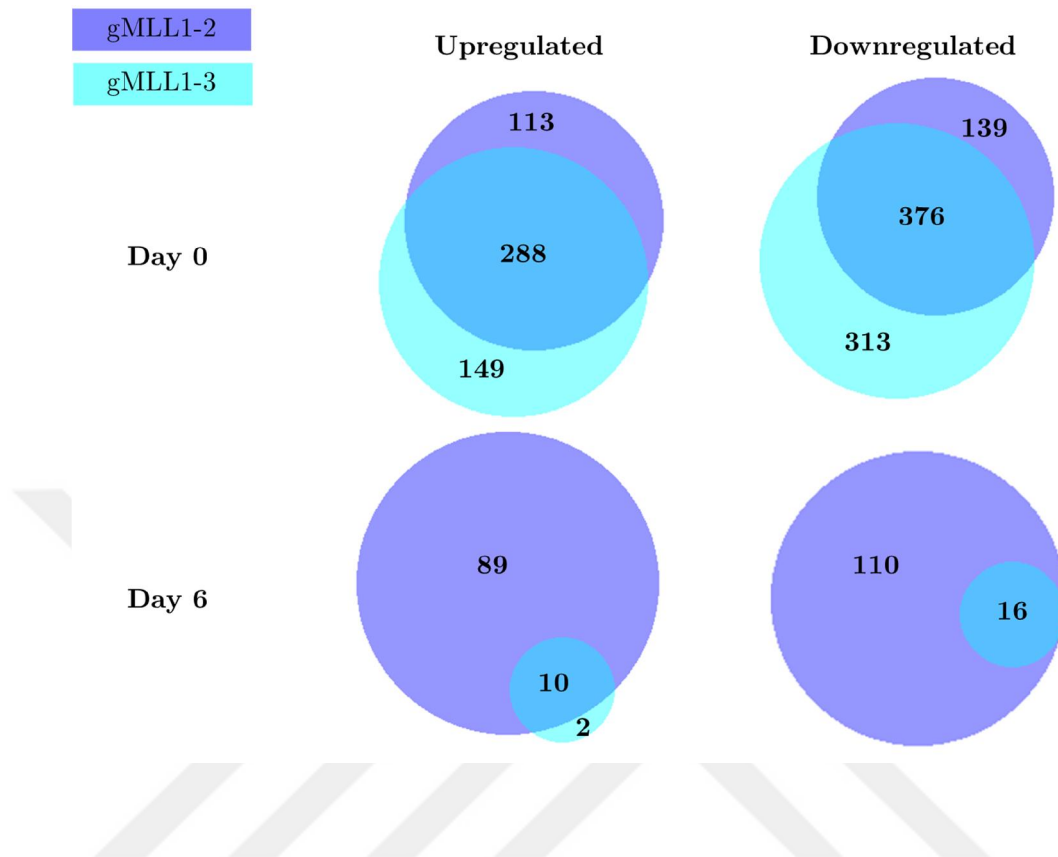


*continues...*



**Figure 41. mRNA expression levels of AF10, AF17, DOT1L and DOT1L-target genes**

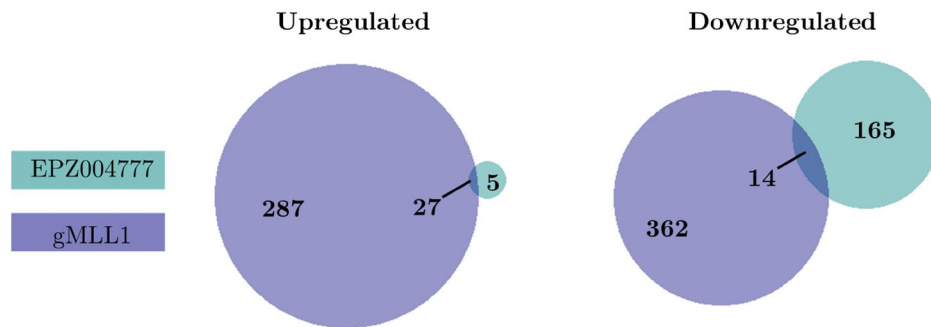
qRT-PCR was performed as duplicate samples and  *$\beta$ -actin* was used as an internal control gene. Every genes expression level is normalized to sgCtrl cells. (n=2; error bars represents s.d.)



**Figure 42. Commonly upregulated and downregulated gene numbers among gMLL1-2 and gMLL1-3 samples at Day 0 and Day 6 samples**

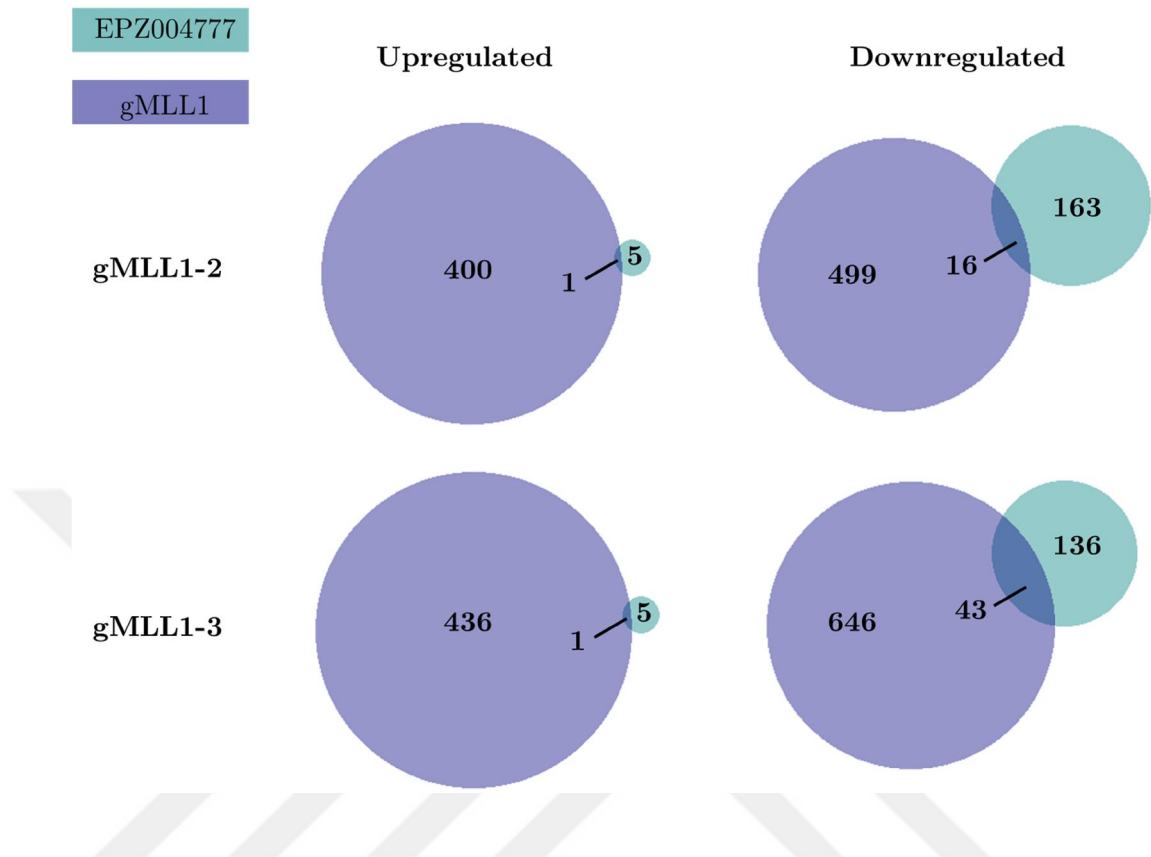
Differentially expressed genes were selected if their p-adjusted-value < 0.05. Upregulated genes were selected if their log2fold\_change > 0 and downregulated genes were selected if log2fold\_change < 0





**Figure 43. Commonly upregulated and downregulated gene numbers among gMLL1 and EPZ004777 samples at Day 0**

Genes that are commonly upregulated and downregulated in gMLL1-2 and gMLL1-3 samples were determined and compared with EPZ004777 sample. Differentially expressed genes were selected if their p-adjusted-value < 0.05. Upregulated genes were selected if their log2fold\_change > 0 and downregulated genes were selected if log2fold\_change < 0.



**Figure 44. Commonly upregulated and downregulated gene numbers among gMML1-2&3 and EPZ004777 samples at Day 0**

Differentially expressed genes were selected if their p-adjusted-value $<0.05$ . Upregulated genes were selected if their  $\log_2\text{fold\_change} >0$  and downregulated genes were selected if  $\log_2\text{fold\_change} <0$ .

## APPENDIX-II

Table 15. Raw data of BioID experiment: HEK293T control cells, 1<sup>st</sup> reading

Accession	Description	ΣCoverage	Σ# PSMs
P62913	60S ribosomal protein L11 OS=Homo sapiens GN=RPL11 PE=1 SV=2 - [RL11_HUMAN]	56.74	28
P46781	40S ribosomal protein S9 OS=Homo sapiens GN=RPS9 PE=1 SV=3 - [RS9_HUMAN]	39.18	28
P04264	Keratin, type II cytoskeletal 1 OS=Homo sapiens GN=KRT1 PE=1 SV=6 - [K2C1_HUMAN]	38.35	51
P26373	60S ribosomal protein L13 OS=Homo sapiens GN=RPL13 PE=1 SV=4 - [RL13_HUMAN]	34.60	20
P18621	60S ribosomal protein L17 OS=Homo sapiens GN=RPL17 PE=1 SV=3 - [RL17_HUMAN]	33.15	14
P63173	60S ribosomal protein L38 OS=Homo sapiens GN=RPL38 PE=1 SV=2 - [RL38_HUMAN]	32.86	4
P49458	Signal recognition particle 9 kDa protein OS=Homo sapiens GN=SRP9 PE=1 SV=2 - [SRP09_HUMAN]	32.56	6
P62241	40S ribosomal protein S8 OS=Homo sapiens GN=RPS8 PE=1 SV=2 - [RS8_HUMAN]	32.21	12
P62701	40S ribosomal protein S4, X isoform OS=Homo sapiens GN=RPS4X PE=1 SV=2 - [RS4X_HUMAN]	31.94	17
P62280	40S ribosomal protein S11 OS=Homo sapiens GN=RPS11 PE=1 SV=3 - [RS11_HUMAN]	31.65	15
P62937	Peptidyl-prolyl cis-trans isomerase A OS=Homo sapiens GN=PP1A PE=1 SV=2 - [PP1A_HUMAN]	31.52	11
P15880	40S ribosomal protein S2 OS=Homo sapiens GN=RPS2 PE=1 SV=2 - [RS2_HUMAN]	30.72	17
P62753	40S ribosomal protein S6 OS=Homo sapiens GN=RPS6 PE=1 SV=1 - [RS6_HUMAN]	29.72	18
P62805	Histone H4 OS=Homo sapiens GN=HIST1H4A PE=1 SV=2 - [H4_HUMAN]	29.13	8
Q06830	Peroxiredoxin-1 OS=Homo sapiens GN=PRDX1 PE=1 SV=1 - [PRDX1_HUMAN]	28.64	11
P83881	60S ribosomal protein L36a OS=Homo sapiens GN=RPL36A PE=1 SV=2 - [RL36A_HUMAN]	28.30	10
O60814	Histone H2B type 1-K OS=Homo sapiens GN=HIST1H2BK PE=1 SV=3 - [H2B1K_HUMAN]	27.78	9
P62888	60S ribosomal protein L30 OS=Homo sapiens GN=RPL30 PE=1 SV=2 - [RL30_HUMAN]	26.96	4
P29372	DNA-3-methyladenine glycosylase OS=Homo sapiens GN=MPG PE=1 SV=3 - [3MG_HUMAN]	26.85	10
P61247	40S ribosomal protein S3a OS=Homo sapiens GN=RPS3A PE=1 SV=2 - [RS3A_HUMAN]	26.52	13
P17844	Probable ATP-dependent RNA helicase DDX5 OS=Homo sapiens GN=DDX5 PE=1 SV=1 - [DDX5_HUMAN]	26.38	41
P27635	60S ribosomal protein L10 OS=Homo sapiens GN=RPL10 PE=1 SV=4 - [RL10_HUMAN]	25.23	12
P46779	60S ribosomal protein L28 OS=Homo sapiens GN=RPL28 PE=1 SV=3 - [RL28_HUMAN]	24.82	9
P83731	60S ribosomal protein L24 OS=Homo sapiens GN=RPL24 PE=1 SV=1 - [RL24_HUMAN]	23.57	7
P60866	40S ribosomal protein S20 OS=Homo sapiens GN=RPS20 PE=1 SV=1 - [RS20_HUMAN]	23.53	9
P46776	60S ribosomal protein L27a OS=Homo sapiens GN=RPL27A PE=1 SV=2 - [RL27A_HUMAN]	22.97	7
P35527	Keratin, type I cytoskeletal 9 OS=Homo sapiens GN=KRT9 PE=1 SV=3 - [K1C9_HUMAN]	22.79	24
P62633	Cellular nucleic acid-binding protein OS=Homo sapiens GN=CNBP PE=1 SV=1 - [CNBP_HUMAN]	22.60	5
P68104	Elongation factor 1-alpha 1 OS=Homo sapiens GN=EEF1A1 PE=1 SV=1 - [EF1A1_HUMAN]	22.51	20
Q9Y3C1	Nucleolar protein 16 OS=Homo sapiens GN=NOP16 PE=1 SV=2 - [NOP16_HUMAN]	22.47	7
P84098	60S ribosomal protein L19 OS=Homo sapiens GN=RPL19 PE=1 SV=1 - [RL19_HUMAN]	22.45	9
P62266	40S ribosomal protein S23 OS=Homo sapiens GN=RPS23 PE=1 SV=3 - [RS23_HUMAN]	22.38	6
P05165	Propionyl-CoA carboxylase alpha chain, mitochondrial OS=Homo sapiens GN=PCCA PE=1 SV=4 - [PCCA_HUMAN]	21.98	24
P62829	60S ribosomal protein L23 OS=Homo sapiens GN=RPL23 PE=1 SV=1 - [RL23_HUMAN]	21.43	3
P42766	60S ribosomal protein L35 OS=Homo sapiens GN=RPL35 PE=1 SV=2 - [RL35_HUMAN]	21.14	5
Q13085	Acetyl-CoA carboxylase 1 OS=Homo sapiens GN=ACACA PE=1 SV=2 - [ACACA_HUMAN]	20.42	74
P11498	Pyruvate carboxylase, mitochondrial OS=Homo sapiens GN=PC PE=1 SV=2 - [PYC_HUMAN]	20.12	43
P62891	60S ribosomal protein L39 OS=Homo sapiens GN=RPL39 PE=1 SV=2 - [RL39_HUMAN]	19.61	1
P62269	40S ribosomal protein S18 OS=Homo sapiens GN=RPS18 PE=1 SV=3 - [RS18_HUMAN]	18.42	4
P06733	Alpha-enolase OS=Homo sapiens GN=ENO1 PE=1 SV=2 - [ENOA_HUMAN]	18.20	14
O95232	Luc7-like protein 3 OS=Homo sapiens GN=LUC7L3 PE=1 SV=2 - [LC7L3_HUMAN]	17.82	16
Q9BQE3	Tubulin alpha-1C chain OS=Homo sapiens GN=TUBA1C PE=1 SV=1 - [TBA1C_HUMAN]	17.82	14
P07437	Tubulin beta chain OS=Homo sapiens GN=TUBB PE=1 SV=2 - [TBB5_HUMAN]	17.57	15
P62857	40S ribosomal protein S28 OS=Homo sapiens GN=RPS28 PE=1 SV=1 - [RS28_HUMAN]	17.39	2
P62861	40S ribosomal protein S30 OS=Homo sapiens GN=FAU PE=1 SV=1 - [RS30_HUMAN]	16.95	2
O00571	ATP-dependent RNA helicase DDX3X OS=Homo sapiens GN=DDX3X PE=1 SV=3 - [DDX3X_HUMAN]	16.16	25
P12236	ADP/ATP translocase 3 OS=Homo sapiens GN=SLC25A6 PE=1 SV=4 - [ADT3_HUMAN]	15.77	12
P46782	40S ribosomal protein S5 OS=Homo sapiens GN=RPS5 PE=1 SV=4 - [RS5_HUMAN]	15.69	8

Accession	Description	ΣCoverage	Σ# PSMs
Q99878	Histone H2A type 1-J OS=Homo sapiens GN=HIST1H2AJ PE=1 SV=3 - [H2A1J_HUMAN]	14.84	1
P62910	60S ribosomal protein L32 OS=Homo sapiens GN=RPL32 PE=1 SV=2 - [RL32_HUMAN]	14.81	6
P35908	Keratin, type II cytoskeletal 2 epidermal OS=Homo sapiens GN=KRT2 PE=1 SV=2 - [K22E_HUMAN]	14.71	15
Q9P258	Protein RCC2 OS=Homo sapiens GN=RCC2 PE=1 SV=2 - [RCC2_HUMAN]	14.37	12
Q96Y4	Translation machinery-associated protein 16 OS=Homo sapiens GN=TMA16 PE=1 SV=2 - [TMA16_HUMAN]	14.29	5
Q9BVP2	Guanine nucleotide-binding protein-like 3 OS=Homo sapiens GN=GNL3 PE=1 SV=2 - [GNL3_HUMAN]	13.84	16
P16403	Histone H1.2 OS=Homo sapiens GN=HIST1H1C PE=1 SV=2 - [H12_HUMAN]	13.62	6
P61978	Heterogeneous nuclear ribonucleoprotein K OS=Homo sapiens GN=HNRNPK PE=1 SV=1 - [HNRPK_HUMAN]	13.61	7
P39019	40S ribosomal protein S19 OS=Homo sapiens GN=RPS19 PE=1 SV=2 - [RS19_HUMAN]	13.10	4
P62273	40S ribosomal protein S29 OS=Homo sapiens GN=RPS29 PE=1 SV=2 - [RS29_HUMAN]	12.50	2
P23396	40S ribosomal protein S3 OS=Homo sapiens GN=RPS3 PE=1 SV=2 - [RS3_HUMAN]	12.35	3
P04080	Cystatin-B OS=Homo sapiens GN=CSTB PE=1 SV=2 - [CYTB_HUMAN]	12.24	2
Q5JTH9	RRP12-like protein OS=Homo sapiens GN=RRP12 PE=1 SV=2 - [RRP12_HUMAN]	12.03	24
Q6NXT2	Histone H3.3C OS=Homo sapiens GN=H3F3C PE=1 SV=3 - [H3C_HUMAN]	11.85	3
E9PL08	Volume-regulated anion channel subunit LRRC8D (Fragment) OS=Homo sapiens GN=LRRC8D PE=4 SV=3 - [E9PL08_HUMAN]	11.76	1
P62249	40S ribosomal protein S16 OS=Homo sapiens GN=RPS16 PE=1 SV=2 - [RS16_HUMAN]	11.64	6
P56270	Myc-associated zinc finger protein OS=Homo sapiens GN=MAZ PE=1 SV=1 - [MAZ_HUMAN]	11.32	10
P62847	40S ribosomal protein S24 OS=Homo sapiens GN=RPS24 PE=1 SV=1 - [RS24_HUMAN]	11.28	3
Q92841	Probable ATP-dependent RNA helicase DDX17 OS=Homo sapiens GN=DDX17 PE=1 SV=2 - [DDX17_HUMAN]	11.25	20
P62917	60S ribosomal protein L8 OS=Homo sapiens GN=RPL8 PE=1 SV=2 - [RL8_HUMAN]	10.51	4
Q15366	Poly(rC)-binding protein 2 OS=Homo sapiens GN=PCBP2 PE=1 SV=1 - [PCBP2_HUMAN]	10.41	5
P48730	Casein kinase I isoform delta OS=Homo sapiens GN=CSNK1D PE=1 SV=2 - [KC1D_HUMAN]	9.88	6
P60709	Actin, cytoplasmic 1 OS=Homo sapiens GN=ACTB PE=1 SV=1 - [ACTB_HUMAN]	9.87	6
P61513	60S ribosomal protein L37a OS=Homo sapiens GN=RPL37A PE=1 SV=2 - [RL37A_HUMAN]	9.78	2
P31943	Heterogeneous nuclear ribonucleoprotein H OS=Homo sapiens GN=HNRNPH1 PE=1 SV=4 - [HNRH1_HUMAN]	9.58	6
H47914	60S ribosomal protein L29 OS=Homo sapiens GN=RPL29 PE=1 SV=2 - [RL29_HUMAN]	9.43	2
H0Y199	28S ribosomal protein S11, mitochondrial OS=Homo sapiens GN=MRPS11 PE=1 SV=1 - [H0Y199_HUMAN]	9.32	1
O15235	28S ribosomal protein S12, mitochondrial OS=Homo sapiens GN=MRPS12 PE=1 SV=1 - [RT12_HUMAN]	8.70	1
P06748	Nucleophosmin OS=Homo sapiens GN=NPM1 PE=1 SV=2 - [NPM_HUMAN]	8.50	5
P09874	Poly [ADP-ribose] polymerase 1 OS=Homo sapiens GN=PARP1 PE=1 SV=4 - [PARP1_HUMAN]	8.28	11
Q9GZV4	Eukaryotic translation initiation factor 5A-2 OS=Homo sapiens GN=EIF5A2 PE=1 SV=3 - [IF5A2_HUMAN]	7.84	2
P62987	Ubiquitin-60S ribosomal protein L40 OS=Homo sapiens GN=UBA52 PE=1 SV=2 - [RL40_HUMAN]	7.81	1
Q9H5H4	Zinc finger protein 768 OS=Homo sapiens GN=ZNF768 PE=1 SV=2 - [ZN768_HUMAN]	7.78	8
P49207	60S ribosomal protein L34 OS=Homo sapiens GN=RPL34 PE=1 SV=3 - [RL34_HUMAN]	7.69	1
G3V5X6	Heterogeneous nuclear ribonucleoproteins C1/C2 (Fragment) OS=Homo sapiens GN=HNRNPC PE=1 SV=1 - [G3V5X6_HUMAN]	7.69	1
P09651	Heterogeneous nuclear ribonucleoprotein A1 OS=Homo sapiens GN=HNRNPA1 PE=1 SV=5 - [ROA1_HUMAN]	7.53	3
P35659	Protein DEK OS=Homo sapiens GN=DEK PE=1 SV=1 - [DEK_HUMAN]	7.47	5
Q02543	60S ribosomal protein L18a OS=Homo sapiens GN=RPL18A PE=1 SV=2 - [RL18A_HUMAN]	7.39	2
Q00839	Heterogeneous nuclear ribonucleoprotein U OS=Homo sapiens GN=HNRNPU PE=1 SV=6 - [HNRPU_HUMAN]	7.39	12
P07477	Trypsin-1 OS=Homo sapiens GN=PRSS1 PE=1 SV=1 - [TRY1_HUMAN]	7.29	13
P62263	40S ribosomal protein S14 OS=Homo sapiens GN=RPS14 PE=1 SV=3 - [RS14_HUMAN]	7.28	1
P61927	60S ribosomal protein L37 OS=Homo sapiens GN=RPL37 PE=1 SV=2 - [RL37_HUMAN]	7.22	1
P62851	40S ribosomal protein S25 OS=Homo sapiens GN=RPS25 PE=1 SV=1 - [RS25_HUMAN]	7.20	2
P62899	60S ribosomal protein L31 OS=Homo sapiens GN=RPL31 PE=1 SV=1 - [RL31_HUMAN]	7.20	1
Q8NC51	Plasminogen activator inhibitor 1 RNA-binding protein OS=Homo sapiens GN=SERBP1 PE=1 SV=2 - [PAIRB_HUMAN]	7.11	4
Q15365	Poly(rC)-binding protein 1 OS=Homo sapiens GN=PCBP1 PE=1 SV=2 - [PCBP1_HUMAN]	7.02	4
Q72317	Zinc finger protein 572 OS=Homo sapiens GN=ZNF572 PE=2 SV=1 - [ZN572_HUMAN]	6.81	1
P61254	60S ribosomal protein L26 OS=Homo sapiens GN=RPL26 PE=1 SV=1 - [RL26_HUMAN]	6.21	2
Q5T280	Putative methyltransferase C9orf114 OS=Homo sapiens GN=C9orf114 PE=1 SV=3 - [CI114_HUMAN]	6.12	3
P62277	40S ribosomal protein S13 OS=Homo sapiens GN=RPS13 PE=1 SV=2 - [RS13_HUMAN]	5.96	1

Accession	Description	ΣCoverage	Σ# PSMs
Q96RQ3	Methylcrotonoyl-CoA carboxylase subunit alpha, mitochondrial OS=Homo sapiens GN=MCCC1 PE=1 SV=3 - [MCCA_HUMAN]	5.93	7
P02533	Keratin, type I cytoskeletal 14 OS=Homo sapiens GN=KRT14 PE=1 SV=4 - [K1C14_HUMAN]	5.93	5
Q16629	Serine/arginine-rich splicing factor 7 OS=Homo sapiens GN=SRSF7 PE=1 SV=1 - [SRSF7_HUMAN]	5.88	2
P50914	60S ribosomal protein L14 OS=Homo sapiens GN=RPL14 PE=1 SV=4 - [RL14_HUMAN]	5.58	2
Q9Y324	rRNA-processing protein FCF1 homolog OS=Homo sapiens GN=FCF1 PE=2 SV=1 - [FCF1_HUMAN]	5.56	2
P78549	Endonuclease III-like protein 1 OS=Homo sapiens GN=NTHL1 PE=1 SV=2 - [NTH_HUMAN]	5.45	2
Q9Y698	Voltage-dependent calcium channel gamma-2 subunit OS=Homo sapiens GN=CACNG2 PE=1 SV=1 - [CCG2_HUMAN]	4.95	1
Q8IWS0	PHD finger protein 6 OS=Homo sapiens GN=PHF6 PE=1 SV=1 - [PHF6_HUMAN]	4.93	3
Q9Y2R4	Probable ATP-dependent RNA helicase DDX52 OS=Homo sapiens GN=DDX52 PE=1 SV=3 - [DDX52_HUMAN]	4.34	3
Q5T440	Putative transferase CAF17, mitochondrial OS=Homo sapiens GN=IBA57 PE=1 SV=1 - [CAF17_HUMAN]	4.21	1
P62424	60S ribosomal protein L7a OS=Homo sapiens GN=RPL7A PE=1 SV=2 - [RL7A_HUMAN]	4.14	2
H0YKS4	Annexin (Fragment) OS=Homo sapiens GN=ANXA2 PE=1 SV=1 - [H0YKS4_HUMAN]	3.98	1
Q9BZE4	Nucleolar GTP-binding protein 1 OS=Homo sapiens GN=GTPBP4 PE=1 SV=3 - [NOG1_HUMAN]	3.94	5
Q5H913-2	Isoform 2 of ADP-ribosylation factor-like protein 13A OS=Homo sapiens GN=ARL13A - [AR13A_HUMAN]	3.91	1
C9JK49	Protein LOC400891 (Fragment) OS=Homo sapiens GN=LOC400891 PE=4 SV=1 - [C9JK49_HUMAN]	3.83	1
P22087	rRNA 2'-O-methyltransferase fibrillar OS=Homo sapiens GN=FBL PE=1 SV=2 - [FBRL_HUMAN]	3.43	2
Q9Y5J1	U3 small nucleolar RNA-associated protein 18 homolog OS=Homo sapiens GN=UTP18 PE=1 SV=3 - [UTP18_HUMAN]	3.24	1
P26599	Polypyrimidine tract-binding protein 1 OS=Homo sapiens GN=PTBP1 PE=1 SV=1 - [PTBP1_HUMAN]	3.20	3
Q02878	60S ribosomal protein L6 OS=Homo sapiens GN=RPL6 PE=1 SV=3 - [RL6_HUMAN]	3.13	2
P02768	Serum albumin OS=Homo sapiens GN=ALB PE=1 SV=2 - [ALBU_HUMAN]	3.12	3
Q9P016	Thymocyte nuclear protein 1 OS=Homo sapiens GN=THYN1 PE=1 SV=1 - [THYN1_HUMAN]	3.11	2
I3L3U9	Ribosomal L1 domain-containing protein 1 (Fragment) OS=Homo sapiens GN=RSL1D1 PE=1 SV=1 - [I3L3U9_HUMAN]	3.04	1
Q96EK4	THAP domain-containing protein 11 OS=Homo sapiens GN=THAP11 PE=1 SV=2 - [THA11_HUMAN]	2.87	2
Q5T6C4	Ataxin-7-like protein 2 OS=Homo sapiens GN=ATXN7L2 PE=4 SV=1 - [Q5T6C4_HUMAN]	2.87	7
MQQY70	Uncharacterized protein (Fragment) OS=Homo sapiens PE=4 SV=1 - [MQQY70_HUMAN]	2.80	1
Q7L354	Zinc finger protein 771 OS=Homo sapiens GN=ZNF771 PE=1 SV=1 - [ZN771_HUMAN]	2.52	2
O75475	PC4 and SFRS1-interacting protein OS=Homo sapiens GN=PSIP1 PE=1 SV=1 - [PSIP1_HUMAN]	2.45	2
G3V1A6	Gasdermin domain containing 1, isoform CRA_d OS=Homo sapiens GN=GSDMD PE=1 SV=1 - [G3V1A6_HUMAN]	2.44	1
Q9H0S4	Probable ATP-dependent RNA helicase DDX47 OS=Homo sapiens GN=DDX47 PE=1 SV=1 - [DDX47_HUMAN]	2.42	2
Q6P087	RNA pseudouridylylase synthase domain-containing protein 3 OS=Homo sapiens GN=RPUSD3 PE=1 SV=3 - [RUSD3_HUMAN]	2.28	2
O00409-2	Isoform 2 of Forkhead box protein N3 OS=Homo sapiens GN=FOXN3 - [FOXN3_HUMAN]	2.14	1
P25205	DNA replication licensing factor MCM3 OS=Homo sapiens GN=MCM3 PE=1 SV=3 - [MCM3_HUMAN]	2.10	3
O43175	D-3-phosphoglycerate dehydrogenase OS=Homo sapiens GN=PHGDH PE=1 SV=4 - [SERA_HUMAN]	2.06	1
O43390	Heterogeneous nuclear ribonucleoprotein R OS=Homo sapiens GN=HNRNPR PE=1 SV=1 - [HNRPR_HUMAN]	2.05	2
Q00325	Phosphate carrier protein, mitochondrial OS=Homo sapiens GN=SLC25A3 PE=1 SV=2 - [MPCP_HUMAN]	1.93	1
Q96P11	Probable 28S rRNA (cytosine-C(5))-methyltransferase OS=Homo sapiens GN=NSUN5 PE=1 SV=2 - [NSUN5_HUMAN]	1.86	2
Q6ZN08	Putative zinc finger protein 66 OS=Homo sapiens GN=ZNF66 PE=5 SV=3 - [ZNF66_HUMAN]	1.57	2
Q8WYQ5	Microprocessor complex subunit DGCR8 OS=Homo sapiens GN=DGCR8 PE=1 SV=1 - [DGCR8_HUMAN]	1.55	1
Q14684	Ribosomal RNA processing protein 1 homolog B OS=Homo sapiens GN=RRP1B PE=1 SV=3 - [RRP1B_HUMAN]	1.45	2
Q9NVP1	ATP-dependent RNA helicase DDX18 OS=Homo sapiens GN=DDX18 PE=1 SV=2 - [DDX18_HUMAN]	1.34	2
A0A087WU6	DNA polymerase alpha catalytic subunit OS=Homo sapiens GN=POLA1 PE=4 SV=1 - [A0A087WU64_HUMAN]	1.30	4
P49711	Transcriptional repressor CTCF OS=Homo sapiens GN=CTCF PE=1 SV=1 - [CTCF_HUMAN]	1.24	2
P55265	Double-stranded RNA-specific adenosine deaminase OS=Homo sapiens GN=ADAR PE=1 SV=4 - [DSRAD_HUMAN]	0.73	2
Q8N2Y8	Iporin OS=Homo sapiens GN=RUSC2 PE=1 SV=3 - [RUSC2_HUMAN]	0.73	1
O60241	Brain-specific angiogenesis inhibitor 2 OS=Homo sapiens GN=BAI2 PE=2 SV=2 - [BAI2_HUMAN]	0.69	1
O00411	DNA-directed RNA polymerase, mitochondrial OS=Homo sapiens GN=POLRMT PE=1 SV=2 - [RPOM_HUMAN]	0.65	2
E9PPJ1	Protein CASC5 OS=Homo sapiens GN=CASC5 PE=1 SV=2 - [E9PPJ1_HUMAN]	0.57	1
H0Y7L2	Dedicator of cytokinesis protein 7 (Fragment) OS=Homo sapiens GN=DOCK7 PE=1 SV=1 - [H0Y7L2_HUMAN]	0.54	1
P78527	DNA-dependent protein kinase catalytic subunit OS=Homo sapiens GN=PRKDC PE=1 SV=3 - [PRKDC_HUMAN]	0.44	3
H3BLS7	Vacuolar protein sorting-associated protein 13D (Fragment) OS=Homo sapiens GN=VPS13D PE=1 SV=1 - [H3BLS7_HUMAN]	0.44	1

Table 16. Raw data of BioID experiment: HEK293T control cells, 2<sup>nd</sup> reading

Accession	Description	ΣCoverage	Σ# PSMS
P49458	Signal recognition particle 9 kDa protein OS=Homo sapiens GN=SRP9 PE=1 SV=2 - [SRP09_HUMAN]	53.49	10
P46781	40S ribosomal protein S9 OS=Homo sapiens GN=RPS9 PE=1 SV=3 - [RS9_HUMAN]	50.52	35
P18621	60S ribosomal protein L17 OS=Homo sapiens GN=RPL17 PE=1 SV=3 - [RL17_HUMAN]	41.85	19
P62280	40S ribosomal protein S11 OS=Homo sapiens GN=RPS11 PE=1 SV=3 - [RS11_HUMAN]	41.77	19
O60814	Histone H2B type 1-K OS=Homo sapiens GN=HIST1H2BK PE=1 SV=3 - [H2B1K_HUMAN]	38.89	10
P26373	60S ribosomal protein L13 OS=Homo sapiens GN=RPL13 PE=1 SV=4 - [RL13_HUMAN]	38.39	26
P62937	Peptidyl-prolyl cis-trans isomerase A OS=Homo sapiens GN=PPIA PE=1 SV=2 - [PPIA_HUMAN]	38.18	10
P62913	60S ribosomal protein L11 OS=Homo sapiens GN=RPL11 PE=1 SV=2 - [RL11_HUMAN]	37.64	22
P62701	40S ribosomal protein S4, X isoform OS=Homo sapiens GN=RPS4X PE=1 SV=2 - [RS4X_HUMAN]	36.50	18
P04264	Keratin, type II cytoskeletal 1 OS=Homo sapiens GN=KRT1 PE=1 SV=6 - [K2C1_HUMAN]	36.18	51
P29372	DNA-3-methyladenine glycosylase OS=Homo sapiens GN=MPG PE=1 SV=3 - [3MG_HUMAN]	34.23	17
P62753	40S ribosomal protein S6 OS=Homo sapiens GN=RPS6 PE=1 SV=1 - [RS6_HUMAN]	33.73	24
P17844	Probable ATP-dependent RNA helicase DDX5 OS=Homo sapiens GN=DDX5 PE=1 SV=1 - [DDX5_HUMAN]	33.06	49
P62249	40S ribosomal protein S16 OS=Homo sapiens GN=RPS16 PE=1 SV=2 - [RS16_HUMAN]	30.14	14
P46779	60S ribosomal protein L28 OS=Homo sapiens GN=RPL28 PE=1 SV=3 - [RL28_HUMAN]	29.93	11
P60866	40S ribosomal protein S20 OS=Homo sapiens GN=RPS20 PE=1 SV=1 - [RS20_HUMAN]	29.41	14
P62805	Histone H4 OS=Homo sapiens GN=HIST1H4A PE=1 SV=2 - [H4_HUMAN]	29.13	6
P61247	40S ribosomal protein S3a OS=Homo sapiens GN=RPS3A PE=1 SV=2 - [RS3A_HUMAN]	28.79	16
P83881	60S ribosomal protein L36a OS=Homo sapiens GN=RPL36A PE=1 SV=2 - [RL36A_HUMAN]	28.30	14
P15880	40S ribosomal protein S2 OS=Homo sapiens GN=RPS2 PE=1 SV=2 - [RS2_HUMAN]	27.65	18
P05165	Propionyl-CoA carboxylase alpha chain, mitochondrial OS=Homo sapiens GN=PCCA PE=1 SV=4 - [PCCA_HUMAN]	27.47	27
P42766	60S ribosomal protein L35 OS=Homo sapiens GN=RPL35 PE=1 SV=2 - [RL35_HUMAN]	26.02	6
P62241	40S ribosomal protein S8 OS=Homo sapiens GN=RPS8 PE=1 SV=2 - [RS8_HUMAN]	25.00	12
P61927	60S ribosomal protein L37 OS=Homo sapiens GN=RPL37 PE=1 SV=2 - [RL37_HUMAN]	24.74	3
Q06830	Peroxioredoxin-1 OS=Homo sapiens GN=PRDX1 PE=1 SV=1 - [PRDX1_HUMAN]	23.12	9
P46776	60S ribosomal protein L27a OS=Homo sapiens GN=RPL27A PE=1 SV=2 - [RL27A_HUMAN]	22.97	9
P62266	40S ribosomal protein S23 OS=Homo sapiens GN=RPS23 PE=1 SV=3 - [RS23_HUMAN]	22.38	8
P68104	Elongation factor 1-alpha 1 OS=Homo sapiens GN=EEF1A1 PE=1 SV=1 - [EF1A1_HUMAN]	22.08	22
Q13085	Acetyl-CoA carboxylase 1 OS=Homo sapiens GN=ACACA PE=1 SV=2 - [ACACA_HUMAN]	21.48	88
P07437	Tubulin beta chain OS=Homo sapiens GN=TUBB PE=1 SV=2 - [TBB5_HUMAN]	21.17	17
P11498	Pyruvate carboxylase, mitochondrial OS=Homo sapiens GN=PC PE=1 SV=2 - [PYC_HUMAN]	21.14	50
P62633	Cellular nucleic acid-binding protein OS=Homo sapiens GN=CNBP PE=1 SV=1 - [CNBP_HUMAN]	20.90	7
Q9BVP2	Guanine nucleotide-binding protein-like 3 OS=Homo sapiens GN=GNL3 PE=1 SV=2 - [GNL3_HUMAN]	20.77	21
Q9P258	Protein RCC2 OS=Homo sapiens GN=RCC2 PE=1 SV=2 - [RCC2_HUMAN]	20.69	15
Q9Y3C1	Nucleolar protein 16 OS=Homo sapiens GN=NOPI6 PE=1 SV=2 - [NOPI6_HUMAN]	20.22	7
Q96EY4	Translation machinery-associated protein 16 OS=Homo sapiens GN=TMA16 PE=1 SV=2 - [TMA16_HUMAN]	20.20	8
P27635	60S ribosomal protein L10 OS=Homo sapiens GN=RPL10 PE=1 SV=4 - [RL10_HUMAN]	20.09	11
P35527	Keratin, type I cytoskeletal 9 OS=Homo sapiens GN=KRT9 PE=1 SV=3 - [K1C9_HUMAN]	20.06	24
P39019	40S ribosomal protein S19 OS=Homo sapiens GN=RPS19 PE=1 SV=2 - [RS19_HUMAN]	20.00	5
P16403	Histone H1.2 OS=Homo sapiens GN=HIST1H1C PE=1 SV=2 - [H12_HUMAN]	19.72	10
P61978	Heterogeneous nuclear ribonucleoprotein K OS=Homo sapiens GN=HNRNPK PE=1 SV=1 - [HNRPK_HUMAN]	19.65	12
P62891	60S ribosomal protein L39 OS=Homo sapiens GN=RPL39 PE=1 SV=2 - [RL39_HUMAN]	19.61	2
P46782	40S ribosomal protein S5 OS=Homo sapiens GN=RPS5 PE=1 SV=4 - [RS5_HUMAN]	19.12	8
P62269	40S ribosomal protein S18 OS=Homo sapiens GN=RPS18 PE=1 SV=3 - [RS18_HUMAN]	18.42	6
P05141	ADP/ATP translocase 2 OS=Homo sapiens GN=SLC25A5 PE=1 SV=7 - [ADT2_HUMAN]	18.12	12
P12236	ADP/ATP translocase 3 OS=Homo sapiens GN=SLC25A6 PE=1 SV=4 - [ADT3_HUMAN]	18.12	12
P84098	60S ribosomal protein L19 OS=Homo sapiens GN=RPL19 PE=1 SV=1 - [RL19_HUMAN]	17.86	10

Accession	Description	ΣCoverage	Σ# PSMs
Q02543	60S ribosomal protein L18a OS=Homo sapiens GN=RPL18A PE=1 SV=2 - [RL18A_HUMAN]	17.61	6
P62888	60S ribosomal protein L30 OS=Homo sapiens GN=RPL30 PE=1 SV=2 - [RL30_HUMAN]	17.39	4
P62857	40S ribosomal protein S28 OS=Homo sapiens GN=RPS28 PE=1 SV=1 - [RS28_HUMAN]	17.39	2
P23396	40S ribosomal protein S3 OS=Homo sapiens GN=RPS3 PE=1 SV=2 - [RS3_HUMAN]	17.28	10
P32969	60S ribosomal protein L9 OS=Homo sapiens GN=RPL9 PE=1 SV=1 - [RL9_HUMAN]	17.19	3
P62861	40S ribosomal protein S30 OS=Homo sapiens GN=FAU PE=1 SV=1 - [RS30_HUMAN]	16.95	2
P22090	40S ribosomal protein S4, Y isoform 1 OS=Homo sapiens GN=RPS4Y1 PE=1 SV=2 - [RS4Y1_HUMAN]	16.35	9
O15235	28S ribosomal protein S12, mitochondrial OS=Homo sapiens GN=MRPS12 PE=1 SV=1 - [RT12_HUMAN]	15.94	4
Q9H5H4	Zinc finger protein 768 OS=Homo sapiens GN=ZNF768 PE=1 SV=2 - [ZN768_HUMAN]	15.93	15
P49207	60S ribosomal protein L34 OS=Homo sapiens GN=RPL34 PE=1 SV=3 - [RL34_HUMAN]	15.38	3
P62910	60S ribosomal protein L32 OS=Homo sapiens GN=RPL32 PE=1 SV=2 - [RL32_HUMAN]	14.81	6
Q92841	Probable ATP-dependent RNA helicase DDX17 OS=Homo sapiens GN=DDX17 PE=1 SV=2 - [DDX17_HUMAN]	14.81	21
P13645	Keratin, type I cytoskeletal 10 OS=Homo sapiens GN=KRT10 PE=1 SV=6 - [K1C10_HUMAN]	14.55	16
P62829	60S ribosomal protein L23 OS=Homo sapiens GN=RPL23 PE=1 SV=1 - [RL23_HUMAN]	14.29	2
P63173	60S ribosomal protein L38 OS=Homo sapiens GN=RPL38 PE=1 SV=2 - [RL38_HUMAN]	14.29	1
P35908	Keratin, type II cytoskeletal 2 epidermal OS=Homo sapiens GN=KRT2 PE=1 SV=2 - [K2E2_HUMAN]	13.93	18
Q96RQ3	Methylcrotonoyl-CoA carboxylase subunit alpha, mitochondrial OS=Homo sapiens GN=MCCC1 PE=1 SV=3 - [MCCA_HUMAN]	13.79	13
Q15365	Poly(rC)-binding protein 1 OS=Homo sapiens GN=PCBP1 PE=1 SV=2 - [PCBP1_HUMAN]	13.76	8
P06748	Nucleophosmin OS=Homo sapiens GN=NPM1 PE=1 SV=2 - [NPM_HUMAN]	13.61	9
Q9BQE3	Tubulin alpha-1C chain OS=Homo sapiens GN=TUBA1C PE=1 SV=1 - [TBA1C_HUMAN]	13.59	10
P68363	Tubulin alpha-1B chain OS=Homo sapiens GN=TUBA1B PE=1 SV=1 - [TBA1B_HUMAN]	13.53	10
Q15366	Poly(rC)-binding protein 2 OS=Homo sapiens GN=PCBP2 PE=1 SV=1 - [PCBP2_HUMAN]	13.42	8
Q9GZV4	Eukaryotic translation initiation factor 5A-2 OS=Homo sapiens GN=EIF5A2 PE=1 SV=3 - [IF5A2_HUMAN]	13.07	3
Q5JTH9	RRP12-like protein OS=Homo sapiens GN=RRP12 PE=1 SV=2 - [RRP12_HUMAN]	12.57	28
P62273	40S ribosomal protein S29 OS=Homo sapiens GN=RPS29 PE=1 SV=2 - [RS29_HUMAN]	12.50	3
P04080	Cystatin-B OS=Homo sapiens GN=CSTB PE=1 SV=2 - [CYTB_HUMAN]	12.24	2
O00571	ATP-dependent RNA helicase DDX3X OS=Homo sapiens GN=DDX3X PE=1 SV=3 - [DDX3X_HUMAN]	11.78	19
P78549	Endonuclease III-like protein 1 OS=Homo sapiens GN=NTHL1 PE=1 SV=2 - [NTH_HUMAN]	11.54	4
P06733	Alpha-enolase OS=Homo sapiens GN=ENO1 PE=1 SV=2 - [ENOA_HUMAN]	11.52	8
Q81YL3	UPF0688 protein C1orf174 OS=Homo sapiens GN=C1orf174 PE=1 SV=2 - [CA174_HUMAN]	11.52	4
P48730	Casein kinase I isoform delta OS=Homo sapiens GN=CSNK1D PE=1 SV=2 - [KC1D_HUMAN]	11.33	9
P62847	40S ribosomal protein S24 OS=Homo sapiens GN=RPS24 PE=1 SV=1 - [RS24_HUMAN]	11.28	4
O95232	Luc7-like protein 3 OS=Homo sapiens GN=LUC7L3 PE=1 SV=2 - [LC7L3_HUMAN]	10.88	15
P09874	Poly [ADP-ribose] polymerase 1 OS=Homo sapiens GN=PARP1 PE=1 SV=4 - [PARP1_HUMAN]	10.55	15
Q8N339	Zinc finger protein 664 OS=Homo sapiens GN=ZNF664 PE=2 SV=1 - [ZN664_HUMAN]	10.34	3
P68431	Histone H3.1 OS=Homo sapiens GN=HIST1H3A PE=1 SV=2 - [H31_HUMAN]	10.29	4
Q9BZE4	Nucleolar GTP-binding protein 1 OS=Homo sapiens GN=GTPBP4 PE=1 SV=3 - [NOG1_HUMAN]	10.09	12
Q7L354	Zinc finger protein 771 OS=Homo sapiens GN=ZNF771 PE=1 SV=1 - [ZN771_HUMAN]	10.09	3
P13647	Keratin, type II cytoskeletal 5 OS=Homo sapiens GN=KRT5 PE=1 SV=3 - [K2C5_HUMAN]	9.83	13
P61513	60S ribosomal protein L37a OS=Homo sapiens GN=RPL37A PE=1 SV=2 - [RL37A_HUMAN]	9.78	2
P56270-2	Isoform 2 of Myc-associated zinc finger protein OS=Homo sapiens GN=MAZ - [MAZ_HUMAN]	9.74	7
P42677	40S ribosomal protein S27 OS=Homo sapiens GN=RPS27 PE=1 SV=3 - [RS27_HUMAN]	9.52	2
P47914	60S ribosomal protein L29 OS=Homo sapiens GN=RPL29 PE=1 SV=2 - [RL29_HUMAN]	9.43	2
Q00839	Heterogeneous nuclear ribonucleoprotein U OS=Homo sapiens GN=HNRNPU PE=1 SV=6 - [HNRPU_HUMAN]	8.97	16
Q96P11-2	Isoform 2 of Probable 28S rRNA (cytosine-C(5))-methyltransferase OS=Homo sapiens GN=NSUN5 - [NSUN5_HUMAN]	8.58	4
P14866	Heterogeneous nuclear ribonucleoprotein L OS=Homo sapiens GN=HNRNPL PE=1 SV=2 - [HNRPL_HUMAN]	8.15	10
F6QUH3	Actin filament-associated protein 1-like 2 (Fragment) OS=Homo sapiens GN=AFAP1L2 PE=1 SV=1 - [F6QUH3_HUMAN]	8.01	1
MOR2S1	Ubiquitin-60S ribosomal protein L40 (Fragment) OS=Homo sapiens GN=UBA52 PE=1 SV=1 - [MOR2S1_HUMAN]	7.97	2
P62854	40S ribosomal protein S26 OS=Homo sapiens GN=RPS26 PE=1 SV=3 - [RS26_HUMAN]	7.83	2
P20719	Homeobox protein Hox-A5 OS=Homo sapiens GN=HOXA5 PE=1 SV=2 - [HXA5_HUMAN]	7.78	2

Accession	Description	ΣCoverage	Σ# PSMs
Q81WS0	PHD finger protein 6 OS=Homo sapiens GN=PHF6 PE=1 SV=1 - [PHF6_HUMAN]	7.40	5
P07477	Trypsin-1 OS=Homo sapiens GN=PRSS1 PE=1 SV=1 - [TRY1_HUMAN]	7.29	16
P62263	40S ribosomal protein S14 OS=Homo sapiens GN=RPS14 PE=1 SV=3 - [RS14_HUMAN]	7.28	2
P62851	40S ribosomal protein S25 OS=Homo sapiens GN=RPS25 PE=1 SV=1 - [RS25_HUMAN]	7.20	2
P62899	60S ribosomal protein L31 OS=Homo sapiens GN=RPL31 PE=1 SV=1 - [RL31_HUMAN]	7.20	1
Q6P087	RNA pseudouridylate synthase domain-containing protein 3 OS=Homo sapiens GN=RPUSD3 PE=1 SV=3 - [RUSD3_HUMAN]	7.12	6
Q8NC51	Plasminogen activator inhibitor 1 RNA-binding protein OS=Homo sapiens GN=SERBP1 PE=1 SV=2 - [PAIRB_HUMAN]	7.11	4
P0C055	Histone H2A.Z OS=Homo sapiens GN=H2AFZ PE=1 SV=2 - [H2AZ_HUMAN]	7.03	2
O43390	Heterogeneous nuclear ribonucleoprotein R OS=Homo sapiens GN=HNRNPR PE=1 SV=1 - [HNRPR_HUMAN]	6.95	7
P60709	Actin, cytoplasmic 1 OS=Homo sapiens GN=ACTB PE=1 SV=1 - [ACTB_HUMAN]	6.93	4
E9PB28	Transcription factor HES-4 OS=Homo sapiens GN=HES4 PE=4 SV=1 - [E9PB28_HUMAN]	6.88	1
Q9Y3B4	Splicing factor 3B subunit 6 OS=Homo sapiens GN=SF3B6 PE=1 SV=1 - [SF3B6_HUMAN]	6.40	1
P62917	60S ribosomal protein L8 OS=Homo sapiens GN=RPL8 PE=1 SV=2 - [RL8_HUMAN]	6.23	2
P61254	60S ribosomal protein L26 OS=Homo sapiens GN=RPL26 PE=1 SV=1 - [RL26_HUMAN]	6.21	2
Q96127	Zinc finger protein 625 OS=Homo sapiens GN=ZNF625 PE=2 SV=1 - [ZNF625_HUMAN]	6.21	2
Q5T280	Putative methyltransferase C9orf114 OS=Homo sapiens GN=C9orf114 PE=1 SV=3 - [C1114_HUMAN]	6.12	4
P31943	Heterogeneous nuclear ribonucleoprotein H OS=Homo sapiens GN=HNRNPH1 PE=1 SV=4 - [HNRH1_HUMAN]	6.01	4
P62277	40S ribosomal protein S13 OS=Homo sapiens GN=RPS13 PE=1 SV=2 - [RS13_HUMAN]	5.96	4
Q14119	Vascular endothelial zinc finger 1 OS=Homo sapiens GN=VEZF1 PE=1 SV=2 - [VEZF1_HUMAN]	5.95	4
P61313	60S ribosomal protein L15 OS=Homo sapiens GN=RPL15 PE=1 SV=2 - [RL15_HUMAN]	5.88	2
P83731	60S ribosomal protein L24 OS=Homo sapiens GN=RPL24 PE=1 SV=1 - [RL24_HUMAN]	5.73	2
Q9Y2R4	Probable ATP-dependent RNA helicase DDX52 OS=Homo sapiens GN=DDX52 PE=1 SV=3 - [DDX52_HUMAN]	5.68	4
P22626	Heterogeneous nuclear ribonucleoproteins A2/B1 OS=Homo sapiens GN=HNRNPA2B1 PE=1 SV=2 - [ROA2_HUMAN]	5.67	4
P50914	60S ribosomal protein L14 OS=Homo sapiens GN=RPL14 PE=1 SV=4 - [RL14_HUMAN]	5.58	2
Q9Y324	rRNA-processing protein FCF1 homolog OS=Homo sapiens GN=FCF1 PE=2 SV=1 - [FCF1_HUMAN]	5.56	2
Q96EK4	THAP domain-containing protein 11 OS=Homo sapiens GN=THAP11 PE=1 SV=2 - [THA11_HUMAN]	5.41	2
Q9P016	Thymocyte nuclear protein 1 OS=Homo sapiens GN=THYN1 PE=1 SV=1 - [THYN1_HUMAN]	5.33	2
P35659	Protein DEK OS=Homo sapiens GN=DEK PE=1 SV=1 - [DEK_HUMAN]	5.07	4
Q16629	Serine/arginine-rich splicing factor 7 OS=Homo sapiens GN=SRSF7 PE=1 SV=1 - [SRSF7_HUMAN]	5.04	2
Q07666	KH domain-containing, RNA-binding, signal transduction-associated protein 1 OS=Homo sapiens GN=KHDRBS1 PE=1 SV=1 - [KHDR1_HUMAN]	4.97	3
Q00059	Transcription factor A, mitochondrial OS=Homo sapiens GN=TFAM PE=1 SV=1 - [TFAM_HUMAN]	4.88	1
P27695	DNA-(apurinic or apyrimidinic site) lyase OS=Homo sapiens GN=APEX1 PE=1 SV=2 - [APEX1_HUMAN]	4.72	2
P46778	60S ribosomal protein L21 OS=Homo sapiens GN=RPL21 PE=1 SV=2 - [RL21_HUMAN]	4.38	1
Q9BXK1	Kruppel-like factor 16 OS=Homo sapiens GN=KLF16 PE=1 SV=1 - [KLF16_HUMAN]	4.37	2
P50454	Serpin H1 OS=Homo sapiens GN=SERPINH1 PE=1 SV=2 - [SERPH_HUMAN]	4.31	4
P02768	Serum albumin OS=Homo sapiens GN=ALB PE=1 SV=2 - [ALBU_HUMAN]	4.27	4
F6TRA5	Spliceosome RNA helicase DDX39B (Fragment) OS=Homo sapiens GN=DDX39B PE=1 SV=1 - [F6TRA5_HUMAN]	4.20	1
Q5VV52	Zinc finger protein 691 OS=Homo sapiens GN=ZNF691 PE=1 SV=1 - [ZNF691_HUMAN]	4.17	2
Q9Y383	Putative RNA-binding protein Luc7-like 2 OS=Homo sapiens GN=LUC7L2 PE=1 SV=2 - [LC7L2_HUMAN]	4.08	2
O75475	PC4 and SFRS1-interacting protein OS=Homo sapiens GN=PSIP1 PE=1 SV=1 - [PSIP1_HUMAN]	3.77	5
P26599	Polypyrimidine tract-binding protein 1 OS=Homo sapiens GN=PTBP1 PE=1 SV=1 - [PTBP1_HUMAN]	3.58	4
P01891	HLA class I histocompatibility antigen, A-68 alpha chain OS=Homo sapiens GN=HLA-A PE=1 SV=4 - [1A68_HUMAN]	3.56	1
Q6ZNL6-2	Isoform 2 of FYVE, RhoGEF and PH domain-containing protein 5 OS=Homo sapiens GN=FGD5 - [FGD5_HUMAN]	3.52	1
O76021	Ribosomal L1 domain-containing protein 1 OS=Homo sapiens GN=RSL1D1 PE=1 SV=3 - [RL1D1_HUMAN]	3.47	4
O95478	Ribosome biogenesis protein NSA2 homolog OS=Homo sapiens GN=NSA2 PE=1 SV=1 - [NSA2_HUMAN]	3.46	1
Q9Y5J1	U3 small nucleolar RNA-associated protein 18 homolog OS=Homo sapiens GN=UTP18 PE=1 SV=3 - [UTP18_HUMAN]	3.24	2
P63244	Guanine nucleotide-binding protein subunit beta-2-like 1 OS=Homo sapiens GN=GNB2L1 PE=1 SV=3 - [GNBP_HUMAN]	3.15	1



Accession	Description	ΣCoverage	Σ# PSMs
Q9Y4X4	Kruppel-like factor 12 OS=Homo sapiens GN=KLF12 PE=1 SV=2 - [KLF12_HUMAN]	2.99	2
P04406	Glyceraldehyde-3-phosphate dehydrogenase OS=Homo sapiens GN=GAPDH PE=1 SV=3 - [G3P_HUMAN]	2.99	2
P09651	Heterogeneous nuclear ribonucleoprotein A1 OS=Homo sapiens GN=HNRNPA1 PE=1 SV=5 - [ROA1_HUMAN]	2.96	2
P25705	ATP synthase subunit alpha, mitochondrial OS=Homo sapiens GN=ATP5A1 PE=1 SV=1 - [ATPA_HUMAN]	2.89	2
Q15717	ELAV-like protein 1 OS=Homo sapiens GN=ELAVL1 PE=1 SV=2 - [ELAV1_HUMAN]	2.76	1
P25205	DNA replication licensing factor MCM3 OS=Homo sapiens GN=MCM3 PE=1 SV=3 - [MCM3_HUMAN]	2.72	5
Q9HC62	Sentrin-specific protease 2 OS=Homo sapiens GN=SEN2 PE=1 SV=3 - [SEN2_HUMAN]	2.72	2
P19012	Keratin, type I cytoskeletal 15 OS=Homo sapiens GN=KRT15 PE=1 SV=3 - [K1C15_HUMAN]	2.63	2
P33992	DNA replication licensing factor MCM5 OS=Homo sapiens GN=MCM5 PE=1 SV=5 - [MCM5_HUMAN]	2.45	3
Q9H054	Probable ATP-dependent RNA helicase DDX47 OS=Homo sapiens GN=DDX47 PE=1 SV=1 - [DDX47_HUMAN]	2.42	2
Q12834	Cell division cycle protein 20 homolog OS=Homo sapiens GN=CDC20 PE=1 SV=2 - [CDC20_HUMAN]	2.40	1
P13797-3	Isoform 3 of Plastin-3 OS=Homo sapiens GN=PLS3 - [PLST_HUMAN]	2.39	1
P18754	Regulator of chromosome condensation OS=Homo sapiens GN=RCC1 PE=1 SV=1 - [RCC1_HUMAN]	2.38	2
P55265	Double-stranded RNA-specific adenosine deaminase OS=Homo sapiens GN=ADAR PE=1 SV=4 - [DSRAD_HUMAN]	2.37	6
P42167	Lamina-associated polypeptide 2, isoforms beta/gamma OS=Homo sapiens GN=TMPO PE=1 SV=2 - [LAP2B_HUMAN]	2.20	2
Q14498	RNA-binding protein 39 OS=Homo sapiens GN=RBM39 PE=1 SV=2 - [RBM39_HUMAN]	2.08	1
Q14667-3	Isoform 3 of UPF0378 protein KIAA0100 OS=Homo sapiens GN=KIAA0100 - [K0100_HUMAN]	2.06	1
Q00325	Phosphate carrier protein, mitochondrial OS=Homo sapiens GN=SLC25A3 PE=1 SV=2 - [MPCP_HUMAN]	1.93	2
Q5C924	Nucleolar MIF4G domain-containing protein 1 OS=Homo sapiens GN=NOM1 PE=1 SV=1 - [NOM1_HUMAN]	1.86	2
Q9BRX2	Protein pelota homolog OS=Homo sapiens GN=PELO PE=1 SV=2 - [PELO_HUMAN]	1.82	1
P14618	Pyruvate kinase PKM OS=Homo sapiens GN=PKM PE=1 SV=4 - [KPYM_HUMAN]	1.69	2
Q86Y23	Hornerin OS=Homo sapiens GN=HRNR PE=1 SV=2 - [HORN_HUMAN]	1.68	2
Q8NFW8	N-acetylneuraminyl cytidyltransferase OS=Homo sapiens GN=CMAS PE=1 SV=2 - [NEUA_HUMAN]	1.61	1
Q6ZN08	Putative zinc finger protein 66 OS=Homo sapiens GN=ZNF66 PE=5 SV=3 - [ZNF66_HUMAN]	1.57	2
Q9UJP1	Histone lysine demethylase PHF8 OS=Homo sapiens GN=PHF8 PE=1 SV=3 - [PHF8_HUMAN]	1.42	2
C9J185	Eukaryotic translation initiation factor 2-alpha kinase 3 (Fragment) OS=Homo sapiens GN=EIF2AK3 PE=1 SV=1 - [C9J185_HUMAN]	1.36	2
Q8NE71	ATP-binding cassette sub-family F member 1 OS=Homo sapiens GN=ABCF1 PE=1 SV=2 - [ABCF1_HUMAN]	1.30	1
Q96T88	E3 ubiquitin-protein ligase UHRF1 OS=Homo sapiens GN=UHRF1 PE=1 SV=1 - [UHRF1_HUMAN]	1.26	2
P49711	Transcriptional repressor CTCF OS=Homo sapiens GN=CTCF PE=1 SV=1 - [CTCF_HUMAN]	1.24	2
Q96ME7	Zinc finger protein 512 OS=Homo sapiens GN=ZNF512 PE=1 SV=2 - [ZNF512_HUMAN]	1.23	2
P52272	Heterogeneous nuclear ribonucleoprotein M OS=Homo sapiens GN=HNRNPM PE=1 SV=3 - [HNRPM_HUMAN]	1.23	2
O00411	DNA-directed RNA polymerase, mitochondrial OS=Homo sapiens GN=POLRMT PE=1 SV=2 - [RPOM_HUMAN]	1.14	2
Q9BQGO	Myb-binding protein 1A OS=Homo sapiens GN=MYBBP1A PE=1 SV=2 - [MBB1A_HUMAN]	1.13	4
P11387	DNA topoisomerase 1 OS=Homo sapiens GN=TOP1 PE=1 SV=2 - [TOP1_HUMAN]	1.05	2
Q53TS8-4	Isoform 4 of Amyotrophic lateral sclerosis 2 chromosomal region candidate gene 11 protein OS=Homo sapiens GN=ALS2CR11 - [AL2SA_HUMAN]	0.99	1
Q9NR82-5	Isoform 5 of Potassium voltage-gated channel subfamily KQT member 5 OS=Homo sapiens GN=KCNQ5 - [KCNQ5_HUMAN]	0.97	1
P40939	Trifunctional enzyme subunit alpha, mitochondrial OS=Homo sapiens GN=HADHA PE=1 SV=2 - [ECHA_HUMAN]	0.92	1
P13639	Elongation factor 2 OS=Homo sapiens GN=EEF2 PE=1 SV=4 - [EF2_HUMAN]	0.82	1
Q15746	Myosin light chain kinase, smooth muscle OS=Homo sapiens GN=MYLK PE=1 SV=4 - [MYLK_HUMAN]	0.78	1
Q9NQ78	Kinesin-like protein KIF13B OS=Homo sapiens GN=KIF13B PE=1 SV=2 - [K13B_HUMAN]	0.71	1
O60241	Brain-specific angiogenesis inhibitor 2 OS=Homo sapiens GN=BAI2 PE=2 SV=2 - [BAI2_HUMAN]	0.69	3
Q8WXE0	Caskin-2 OS=Homo sapiens GN=CASKIN2 PE=1 SV=2 - [CSK12_HUMAN]	0.58	1
Q5T200	Zinc finger CCH domain-containing protein 13 OS=Homo sapiens GN=ZC3H13 PE=1 SV=1 - [ZC3HD_HUMAN]	0.42	1
Q5VYK3	Proteasome-associated protein ECM29 homolog OS=Homo sapiens GN=ECM29 PE=1 SV=2 - [ECM29_HUMAN]	0.38	1
P46013	Antigen KI-67 OS=Homo sapiens GN=MKI67 PE=1 SV=2 - [KI67_HUMAN]	0.28	2
Q5UIP0	Telomere-associated protein RIF1 OS=Homo sapiens GN=RIF1 PE=1 SV=2 - [RIF1_HUMAN]	0.28	1
P78527	DNA-dependent protein kinase catalytic subunit OS=Homo sapiens GN=PRKDC PE=1 SV=3 - [PRKDC_HUMAN]	0.24	2

**Table 17. Raw data of BioID experiment: BioID-DOT1L-wt infected HEK293T cells, 1<sup>st</sup> reading**

Accession	Description	ΣCoverage	Σ# PSMS
P49458	Signal recognition particle 9 kDa protein OS=Homo sapiens GN=SRP9 PE=1 SV=2 - [SRP09_HUMAN]	41.86	7
P62913	60S ribosomal protein L11 OS=Homo sapiens GN=RPL11 PE=1 SV=2 - [RL11_HUMAN]	39.33	15
P62805	Histone H4 OS=Homo sapiens GN=HIST1H4A PE=1 SV=2 - [H4_HUMAN]	38.83	10
P62249	40S ribosomal protein S16 OS=Homo sapiens GN=RPS16 PE=1 SV=2 - [RS16_HUMAN]	36.30	8
P62888	60S ribosomal protein L30 OS=Homo sapiens GN=RPL30 PE=1 SV=2 - [RL30_HUMAN]	34.78	8
P26373	60S ribosomal protein L13 OS=Homo sapiens GN=RPL13 PE=1 SV=4 - [RL13_HUMAN]	34.60	17
P46781	40S ribosomal protein S9 OS=Homo sapiens GN=RPS9 PE=1 SV=3 - [RS9_HUMAN]	34.02	18
P62241	40S ribosomal protein S8 OS=Homo sapiens GN=RPS8 PE=1 SV=2 - [RS8_HUMAN]	32.21	12
P62880	40S ribosomal protein S11 OS=Homo sapiens GN=RPS11 PE=1 SV=3 - [RS11_HUMAN]	31.01	13
P29372	DNA-3-methyladenine glycosylase OS=Homo sapiens GN=MPG PE=1 SV=3 - [3MG_HUMAN]	30.87	13
P23396	40S ribosomal protein S3 OS=Homo sapiens GN=RPS3 PE=1 SV=2 - [RS3_HUMAN]	28.81	12
Q8TEK3	Histone-lysine N-methyltransferase, H3 lysine-79 specific OS=Homo sapiens GN=DOT1L PE=1 SV=2 - [DOT1L_HUMAN]	28.12	101
P62753	40S ribosomal protein S6 OS=Homo sapiens GN=RPS6 PE=1 SV=1 - [RS6_HUMAN]	28.11	23
P17844	Probable ATP-dependent RNA helicase DDX5 OS=Homo sapiens GN=DDX5 PE=1 SV=1 - [DDX5_HUMAN]	27.85	36
P62701	40S ribosomal protein S4, X isoform OS=Homo sapiens GN=RPS4X PE=1 SV=2 - [RS4X_HUMAN]	26.62	11
Q8N257	Histone H2B type 3-B OS=Homo sapiens GN=HIST3H2BB PE=1 SV=3 - [H2B3B_HUMAN]	26.19	7
O60814	Histone H2B type 1-K OS=Homo sapiens GN=HIST1H2BK PE=1 SV=3 - [H2B1K_HUMAN]	26.19	8
P27635	60S ribosomal protein L10 OS=Homo sapiens GN=RPL10 PE=1 SV=4 - [RL10_HUMAN]	25.23	12
P11498	Pyruvate carboxylase, mitochondrial OS=Homo sapiens GN=PC PE=1 SV=2 - [PYC_HUMAN]	24.62	45
P83881	60S ribosomal protein L36a OS=Homo sapiens GN=RPL36A PE=1 SV=2 - [RL36A_HUMAN]	24.53	6
P18621	60S ribosomal protein L17 OS=Homo sapiens GN=RPL17 PE=1 SV=3 - [RL17_HUMAN]	23.37	7
P46776	60S ribosomal protein L27a OS=Homo sapiens GN=RPL27A PE=1 SV=2 - [RL27A_HUMAN]	22.97	6
P60866	40S ribosomal protein S20 OS=Homo sapiens GN=RPS20 PE=1 SV=1 - [RS20_HUMAN]	22.69	8
P62937	Peptidyl-prolyl cis-trans isomerase A OS=Homo sapiens GN=PPIA PE=1 SV=2 - [PPIA_HUMAN]	20.61	4
Q06830	Peroxioredoxin-1 OS=Homo sapiens GN=PRDX1 PE=1 SV=1 - [PRDX1_HUMAN]	20.10	6
P62891	60S ribosomal protein L39 OS=Homo sapiens GN=RPL39 PE=1 SV=2 - [RL39_HUMAN]	19.61	1
P07437	Tubulin beta chain OS=Homo sapiens GN=TUBB PE=1 SV=2 - [TBB5_HUMAN]	19.59	11
Q15365	Poly(rC)-binding protein 1 OS=Homo sapiens GN=PCBP1 PE=1 SV=2 - [PCBP1_HUMAN]	19.38	9
P46779	60S ribosomal protein L28 OS=Homo sapiens GN=RPL28 PE=1 SV=3 - [RL28_HUMAN]	18.98	5
P15880	40S ribosomal protein S2 OS=Homo sapiens GN=RPS2 PE=1 SV=2 - [RS2_HUMAN]	18.77	8
Q9P258	Protein RCC2 OS=Homo sapiens GN=RCC2 PE=1 SV=2 - [RCC2_HUMAN]	18.77	14
P68104	Elongation factor 1-alpha 1 OS=Homo sapiens GN=EEF1A1 PE=1 SV=1 - [EF1A1_HUMAN]	18.61	16
P63173	60S ribosomal protein L38 OS=Homo sapiens GN=RPL38 PE=1 SV=2 - [RL38_HUMAN]	18.57	2
P62269	40S ribosomal protein S18 OS=Homo sapiens GN=RPS18 PE=1 SV=3 - [RS18_HUMAN]	18.42	4
Q13085	Acetyl-CoA carboxylase 1 OS=Homo sapiens GN=ACACA PE=1 SV=2 - [ACACA_HUMAN]	18.24	57
P61978	Heterogeneous nuclear ribonucleoprotein K OS=Homo sapiens GN=HNRNPK PE=1 SV=1 - [HNRPK_HUMAN]	17.71	11
Q9BQE3	Tubulin alpha-1C chain OS=Homo sapiens GN=TUBA1C PE=1 SV=1 - [TBA1C_HUMAN]	17.59	13
P68363	Tubulin alpha-1B chain OS=Homo sapiens GN=TUBA1B PE=1 SV=1 - [TBA1B_HUMAN]	17.52	13
Q9BVP2	Guanine nucleotide-binding protein-like 3 OS=Homo sapiens GN=GNL3 PE=1 SV=2 - [GNL3_HUMAN]	17.49	12
F8W8C9	Immunoglobulin iota chain OS=Homo sapiens GN=VPREB1 PE=1 SV=1 - [F8W8C9_HUMAN]	17.36	1
P62861	40S ribosomal protein S30 OS=Homo sapiens GN=FAU PE=1 SV=1 - [RS30_HUMAN]	16.95	1
O00571	ATP-dependent RNA helicase DDX3X OS=Homo sapiens GN=DDX3X PE=1 SV=3 - [DDX3X_HUMAN]	16.92	19
Q9Y3C1	Nucleolar protein 16 OS=Homo sapiens GN=NOP16 PE=1 SV=2 - [NOP16_HUMAN]	16.29	4
Q15366	Poly(rC)-binding protein 2 OS=Homo sapiens GN=PCBP2 PE=1 SV=1 - [PCBP2_HUMAN]	15.89	6
P05165	Propionyl-CoA carboxylase alpha chain, mitochondrial OS=Homo sapiens GN=PCCA PE=1 SV=4 - [PCCA_HUMAN]	15.66	19
P42766	60S ribosomal protein L35 OS=Homo sapiens GN=RPL35 PE=1 SV=2 - [RL35_HUMAN]	15.45	3
P62851	40S ribosomal protein S25 OS=Homo sapiens GN=RPS25 PE=1 SV=1 - [RS25_HUMAN]	15.20	4
P62910	60S ribosomal protein L32 OS=Homo sapiens GN=RPL32 PE=1 SV=2 - [RL32_HUMAN]	14.81	5

Accession	Description	ΣCoverage	Σ# PSMs
P62266	40S ribosomal protein S23 OS=Homo sapiens GN=RPS23 PE=1 SV=3 - [RS23_HUMAN]	14.69	4
P16403	Histone H1.2 OS=Homo sapiens GN=H1S1H1C PE=1 SV=2 - [H12_HUMAN]	14.08	8
P06733	Alpha-enolase OS=Homo sapiens GN=ENO1 PE=1 SV=2 - [ENO1_HUMAN]	14.06	10
O95232	Luc7-like protein 3 OS=Homo sapiens GN=LUC7L3 PE=1 SV=2 - [LC7L3_HUMAN]	13.89	11
P84098	60S ribosomal protein L19 OS=Homo sapiens GN=RPL19 PE=1 SV=1 - [RL19_HUMAN]	13.78	4
P04247	40S ribosomal protein S3a OS=Homo sapiens GN=RPS3A PE=1 SV=2 - [RS3A_HUMAN]	13.64	7
P62633	Cellular nucleic acid-binding protein OS=Homo sapiens GN=CNBP PE=1 SV=1 - [CNBP_HUMAN]	13.56	3
P83731	60S ribosomal protein L24 OS=Homo sapiens GN=RPL24 PE=1 SV=1 - [RL24_HUMAN]	13.38	5
P05141	ADP/ATP translocase 2 OS=Homo sapiens GN=SLC25A5 PE=1 SV=7 - [ADT2_HUMAN]	13.09	8
P62273	40S ribosomal protein S29 OS=Homo sapiens GN=RPS29 PE=1 SV=2 - [RS29_HUMAN]	12.50	2
P04264	Keratin, type II cytoskeletal 1 OS=Homo sapiens GN=KRT1 PE=1 SV=6 - [K2C1_HUMAN]	12.11	13
Q6NXT2	Histone H3.3C OS=Homo sapiens GN=H3F3C PE=1 SV=3 - [H3C_HUMAN]	11.85	2
P46782	40S ribosomal protein S5 OS=Homo sapiens GN=RPS5 PE=1 SV=4 - [RS5_HUMAN]	11.76	6
P62847	40S ribosomal protein S24 OS=Homo sapiens GN=RPS24 PE=1 SV=1 - [RS24_HUMAN]	11.28	4
Q96EY4	Translation machinery-associated protein 16 OS=Homo sapiens GN=TMA16 PE=1 SV=2 - [TMA16_HUMAN]	10.84	3
Q92841	Probable ATP-dependent RNA helicase DDX17 OS=Homo sapiens GN=DDX17 PE=1 SV=2 - [DDX17_HUMAN]	10.70	15
P62917	60S ribosomal protein L8 OS=Homo sapiens GN=RPL8 PE=1 SV=2 - [RL8_HUMAN]	10.51	4
P61927	60S ribosomal protein L37 OS=Homo sapiens GN=RPL37 PE=1 SV=2 - [RL37_HUMAN]	10.31	1
P12236	ADP/ATP translocase 3 OS=Homo sapiens GN=SLC25A6 PE=1 SV=4 - [ADT3_HUMAN]	9.73	6
P31943	Heterogeneous nuclear ribonucleoprotein H OS=Homo sapiens GN=HNRNPH1 PE=1 SV=4 - [HNRH1_HUMAN]	9.58	6
P09874	Poly [ADP-ribose] polymerase 1 OS=Homo sapiens GN=PARP1 PE=1 SV=4 - [PARP1_HUMAN]	9.47	10
P47914	60S ribosomal protein L29 OS=Homo sapiens GN=RPL29 PE=1 SV=2 - [RL29_HUMAN]	9.43	2
O15235	28S ribosomal protein S12, mitochondrial OS=Homo sapiens GN=MRPS12 PE=1 SV=1 - [RT12_HUMAN]	8.70	1
P78549	Endonuclease III-like protein 1 OS=Homo sapiens GN=NTHL1 PE=1 SV=2 - [NTH_HUMAN]	8.33	3
Q9BQ61	Uncharacterized protein C19orf43 OS=Homo sapiens GN=C19orf43 PE=1 SV=1 - [CS043_HUMAN]	7.95	1
Q9GZV4	Eukaryotic translation initiation factor 5A-2 OS=Homo sapiens GN=EIF5A2 PE=1 SV=3 - [IF5A2_HUMAN]	7.84	2
P62987	Ubiquitin-60S ribosomal protein L40 OS=Homo sapiens GN=UBA52 PE=1 SV=2 - [RL40_HUMAN]	7.81	4
P35527	Keratin, type I cytoskeletal 9 OS=Homo sapiens GN=KRT9 PE=1 SV=3 - [K1C9_HUMAN]	7.70	7
P49207	60S ribosomal protein L34 OS=Homo sapiens GN=RPL34 PE=1 SV=3 - [RL34_HUMAN]	7.69	2
C9JPD0	Neurexophilin-1 (Fragment) OS=Homo sapiens GN=NXP1 PE=4 SV=1 - [C9JPD0_HUMAN]	7.69	1
Q02878	60S ribosomal protein L6 OS=Homo sapiens GN=RPL6 PE=1 SV=3 - [RL6_HUMAN]	7.64	4
Q9H5H4	Zinc finger protein 768 OS=Homo sapiens GN=ZNF768 PE=1 SV=2 - [ZN768_HUMAN]	7.59	7
P22087	rRNA 2'-O-methyltransferase fibrillarin OS=Homo sapiens GN=FBL PE=1 SV=2 - [FBRL_HUMAN]	7.48	2
P48730	Casein kinase 1 isoform delta OS=Homo sapiens GN=CSNK1D PE=1 SV=2 - [KC1D_HUMAN]	7.47	5
P60709	Actin, cytoplasmic 1 OS=Homo sapiens GN=ACTB PE=1 SV=1 - [ACTB_HUMAN]	7.47	4
Q86T24	Transcriptional regulator Kaiso OS=Homo sapiens GN=ZBTB33 PE=1 SV=2 - [KAISO_HUMAN]	7.44	8
Q02543	60S ribosomal protein L18a OS=Homo sapiens GN=RPL18A PE=1 SV=2 - [RL18A_HUMAN]	7.39	2
P07477	Trypsin-1 OS=Homo sapiens GN=PRSS1 PE=1 SV=1 - [TRY1_HUMAN]	7.29	9
P62263	40S ribosomal protein S14 OS=Homo sapiens GN=RPS14 PE=1 SV=3 - [RS14_HUMAN]	7.28	2
Q8NC51	Plasminogen activator inhibitor 1 RNA-binding protein OS=Homo sapiens GN=SERBP1 PE=1 SV=2 - [PAIRB_HUMAN]	7.11	4
Q86UA1-2	Isoform 2 of Pre-mRNA-processing factor 39 OS=Homo sapiens GN=PRPF39 - [PRP39_HUMAN]	6.99	1
P39019	40S ribosomal protein S19 OS=Homo sapiens GN=RPS19 PE=1 SV=2 - [RS19_HUMAN]	6.90	1
Q9NSQ0	Putative ribosomal RNA-processing protein 7 homolog B OS=Homo sapiens GN=RRP7B PE=5 SV=1 - [RRP7B_HUMAN]	6.80	1
P26641	Elongation factor 1-gamma OS=Homo sapiens GN=EEF1G PE=1 SV=3 - [EF1G_HUMAN]	6.41	3
Q03111	Protein ENL OS=Homo sapiens GN=MLLT1 PE=1 SV=2 - [ENL_HUMAN]	6.26	4
P61254	60S ribosomal protein L26 OS=Homo sapiens GN=RPL26 PE=1 SV=1 - [RL26_HUMAN]	6.21	2
Q96RQ3	Methylocrotonoyl-CoA carboxylase subunit alpha, mitochondrial OS=Homo sapiens GN=MCCC1 PE=1 SV=3 - [MCCA_HUMAN]	5.79	7
P06748	Nucleophosmin OS=Homo sapiens GN=NPM1 PE=1 SV=2 - [NPM_HUMAN]	5.78	3
Q00839	Heterogeneous nuclear ribonucleoprotein U OS=Homo sapiens GN=HNRNPU PE=1 SV=6 - [HNRPU_HUMAN]	5.70	10
P84103-2	Isoform 2 of Serine/arginine-rich splicing factor 3 OS=Homo sapiens GN=SRSF3 - [SRSF3_HUMAN]	5.65	1
Q9Y324	rRNA-processing protein FCF1 homolog OS=Homo sapiens GN=FCF1 PE=2 SV=1 - [FCF1_HUMAN]	5.56	2
Q5JTH9	RRP12-like protein OS=Homo sapiens GN=RRP12 PE=1 SV=2 - [RRP12_HUMAN]	5.55	10
P09651	Heterogeneous nuclear ribonucleoprotein A1 OS=Homo sapiens GN=HNRNPA1 PE=1 SV=5 - [ROA1_HUMAN]	5.38	3

Accession	Description	ΣCoverage	Σ# PSMs
P22626	Heterogeneous nuclear ribonucleoproteins A2/B1 OS=Homo sapiens GN=HNRNPA2B1 PE=1 SV=2 - [ROA2_HUMAN]	5.38	3
Q7L354	Zinc finger protein 771 OS=Homo sapiens GN=ZNF771 PE=1 SV=1 - [ZN771_HUMAN]	5.36	3
Q9NP85	Podocin OS=Homo sapiens GN=NPHS2 PE=1 SV=1 - [PODO_HUMAN]	5.22	1
P07305	Histone H1.0 OS=Homo sapiens GN=H1F0 PE=1 SV=3 - [H10_HUMAN]	5.15	2
P56270	Myc-associated zinc finger protein OS=Homo sapiens GN=MAZ PE=1 SV=1 - [MAZ_HUMAN]	4.82	4
P25705	ATP synthase subunit alpha, mitochondrial OS=Homo sapiens GN=ATP5A1 PE=1 SV=1 - [ATPA_HUMAN]	4.70	4
P62277	40S ribosomal protein S13 OS=Homo sapiens GN=RPS13 PE=1 SV=2 - [RS13_HUMAN]	4.64	1
Q8IZD6	Solute carrier family 22 member 15 OS=Homo sapiens GN=SLC22A15 PE=2 SV=1 - [S22AF_HUMAN]	4.20	1
Q98ZE4	Nucleolar GTP-binding protein 1 OS=Homo sapiens GN=GTPBP4 PE=1 SV=3 - [NOG1_HUMAN]	3.94	4
Q75475	PC4 and SFRS1-interacting protein OS=Homo sapiens GN=PSIP1 PE=1 SV=1 - [PSIP1_HUMAN]	3.77	4
Q5T280	Putative methyltransferase C9orf114 OS=Homo sapiens GN=C9orf114 PE=1 SV=3 - [C1114_HUMAN]	3.72	2
P42167	Lamina-associated polypeptide 2, isoforms beta/gamma OS=Homo sapiens GN=TMPO PE=1 SV=2 - [LAP2B_HUMAN]	3.52	2
Q9NVV4	Poly(A) RNA polymerase, mitochondrial OS=Homo sapiens GN=MTPAP PE=1 SV=1 - [PAPD1_HUMAN]	3.44	1
Q9Y5J1	U3 small nucleolar RNA-associated protein 18 homolog OS=Homo sapiens GN=UTP18 PE=1 SV=3 - [UTP18_HUMAN]	3.24	1
P14866	Heterogeneous nuclear ribonucleoprotein L OS=Homo sapiens GN=HNRNPL PE=1 SV=2 - [HNRPL_HUMAN]	3.23	2
Q16352	Alpha-interxerin OS=Homo sapiens GN=INA PE=1 SV=2 - [AINX_HUMAN]	3.21	2
P55198	Protein AF-17 OS=Homo sapiens GN=MLLT6 PE=1 SV=2 - [AF17_HUMAN]	3.20	4
Q15233	Non-POU domain-containing octamer-binding protein OS=Homo sapiens GN=NONO PE=1 SV=4 - [NONO_HUMAN]	3.18	3
Q9P016	Thymocyte nuclear protein 1 OS=Homo sapiens GN=THYN1 PE=1 SV=1 - [THYN1_HUMAN]	3.11	1
Q8IWS0	PHD finger protein 6 OS=Homo sapiens GN=PHF6 PE=1 SV=1 - [PHF6_HUMAN]	3.01	1
P04406	Glyceraldehyde-3-phosphate dehydrogenase OS=Homo sapiens GN=GAPDH PE=1 SV=3 - [G3P_HUMAN]	2.99	2
Q5T6C4	Ataxin-7-like protein 2 OS=Homo sapiens GN=ATXN7L2 PE=4 SV=1 - [Q5T6C4_HUMAN]	2.87	1
Q15717	ELAV-like protein 1 OS=Homo sapiens GN=ELAVL1 PE=1 SV=2 - [ELAV1_HUMAN]	2.76	1
Q9H054	Probable ATP-dependent RNA helicase DDX47 OS=Homo sapiens GN=DDX47 PE=1 SV=1 - [DDX47_HUMAN]	2.42	2
P35659	Protein DEK OS=Homo sapiens GN=DEK PE=1 SV=1 - [DEK_HUMAN]	2.40	1
Q9Y2R4	Probable ATP-dependent RNA helicase DDX52 OS=Homo sapiens GN=DDX52 PE=1 SV=3 - [DDX52_HUMAN]	2.34	2
Q6P087	RNA pseudouridylylase synthase domain-containing protein 3 OS=Homo sapiens GN=RPUSD3 PE=1 SV=3 - [RUSD3_HUMAN]	2.28	1
Q178G1	BTB/POZ domain-containing protein KCTD19 OS=Homo sapiens GN=KCTD19 PE=2 SV=1 - [KCD19_HUMAN]	2.27	1
P12270	Nucleoprotein TPR OS=Homo sapiens GN=TPR PE=1 SV=3 - [TPR_HUMAN]	2.07	6
Q8TF46	DIS3-like exonuclease 1 OS=Homo sapiens GN=DIS3L PE=1 SV=2 - [D13L1_HUMAN]	1.99	1
Q9NR30	Nucleolar RNA helicase 2 OS=Homo sapiens GN=DDX21 PE=1 SV=5 - [DDX21_HUMAN]	1.92	1
P26599	Polypyrimidine tract-binding protein 1 OS=Homo sapiens GN=PTBP1 PE=1 SV=1 - [PTBP1_HUMAN]	1.88	1
P49959	Double-strand break repair protein MRE11A OS=Homo sapiens GN=MRE11A PE=1 SV=3 - [MRE11_HUMAN]	1.84	2
Q75182	Paired amphipathic helix protein Sin3b OS=Homo sapiens GN=Sin3B PE=1 SV=2 - [SIN3B_HUMAN]	1.81	2
Q8NFW8	N-acetylneuraminyl transferase OS=Homo sapiens GN=CMAS PE=1 SV=2 - [NEUA_HUMAN]	1.61	1
Q8WYQ5	Microprocessor complex subunit DGCR8 OS=Homo sapiens GN=DGCR8 PE=1 SV=1 - [DGCR8_HUMAN]	1.55	2
P55786-2	Isoform 2 of Puromycin-sensitive aminopeptidase OS=Homo sapiens GN=NPEPPS - [PSA_HUMAN]	1.55	1
Q14684	Ribosomal RNA processing protein 1 homolog B OS=Homo sapiens GN=RRP1B PE=1 SV=3 - [RRP1B_HUMAN]	1.45	1
P55197	Protein AF-10 OS=Homo sapiens GN=MLLT10 PE=1 SV=2 - [AF10_HUMAN]	1.31	2
O60506	Heterogeneous nuclear ribonucleoprotein Q OS=Homo sapiens GN=SYNCRIP PE=1 SV=2 - [HNRPQ_HUMAN]	1.28	2
P49711	Transcriptional repressor CTCF OS=Homo sapiens GN=CTCF PE=1 SV=1 - [CTCF_HUMAN]	1.24	1
P02768	Serum albumin OS=Homo sapiens GN=ALB PE=1 SV=2 - [ALBU_HUMAN]	1.15	1
Q86VF7	Nebulin-related-anchoring protein OS=Homo sapiens GN=NRAP PE=2 SV=2 - [NRAP_HUMAN]	1.10	1
Q12879	Glutamate receptor ionotropic, NMDA 2A OS=Homo sapiens GN=GRIN2A PE=1 SV=1 - [NMDE1_HUMAN]	1.09	1
Q9NVP1	ATP-dependent RNA helicase DDX18 OS=Homo sapiens GN=DDX18 PE=1 SV=2 - [DDX18_HUMAN]	1.04	1
P25205	DNA replication licensing factor MCM3 OS=Homo sapiens GN=MCM3 PE=1 SV=3 - [MCM3_HUMAN]	0.87	1
P13639	Elongation factor 2 OS=Homo sapiens GN=EEF2 PE=1 SV=4 - [EF2_HUMAN]	0.82	1
P55265	Double-stranded RNA-specific adenosine deaminase OS=Homo sapiens GN=ADAR PE=1 SV=4 - [DSRAD_HUMAN]	0.73	1
P46940	Ras GTPase-activating-like protein IQGAP1 OS=Homo sapiens GN=IQGAP1 PE=1 SV=1 - [IQGA1_HUMAN]	0.66	1
P11388	DNA topoisomerase 2-alpha OS=Homo sapiens GN=TOP2A PE=1 SV=3 - [TOP2A_HUMAN]	0.65	1
O00411	DNA-directed RNA polymerase, mitochondrial OS=Homo sapiens GN=POLRMT PE=1 SV=2 - [RPOH_HUMAN]	0.65	1
Q14980	Nuclear mitotic apparatus protein 1 OS=Homo sapiens GN=NUMA1 PE=1 SV=2 - [NUMA1_HUMAN]	0.57	2
Q9BQG0	Myb-binding protein 1A OS=Homo sapiens GN=MYBBP1A PE=1 SV=2 - [MBB1A_HUMAN]	0.53	1
Q8TD26	Chromodomain-helicase-DNA-binding protein 6 OS=Homo sapiens GN=CHD6 PE=1 SV=4 - [CHD6_HUMAN]	0.41	1
Q5VYK3	Proteasome-associated protein ECM29 homolog OS=Homo sapiens GN=ECM29 PE=1 SV=2 - [ECM29_HUMAN]	0.38	1
Q8IVF2	Protein AHNAK2 OS=Homo sapiens GN=AHNAK2 PE=1 SV=2 - [AHNK2_HUMAN]	0.35	1
Q4LDE5	Sushi, von Willebrand factor type A, EGF and pentraxin domain-containing protein 1 OS=Homo sapiens GN=SVEP1 PE=1 SV=3 - [SVEP1_HUMAN]	0.31	2
P46013	Antigen KI-67 OS=Homo sapiens GN=MKI67 PE=1 SV=2 - [KI67_HUMAN]	0.28	1

**Table 18. Raw data of BioID experiment: BioID-DOT1L-wt infected HEK293T cells, 2<sup>nd</sup> reading**

Accession	Description	ΣCoverage	Σ# PSMs
P62888	60S ribosomal protein L30 OS=Homo sapiens GN=RPL30 PE=1 SV=2 - [RL30_HUMAN]	47.83	8
P46781	40S ribosomal protein S9 OS=Homo sapiens GN=RPS9 PE=1 SV=3 - [RS9_HUMAN]	46.91	33
P49458	Signal recognition particle 9 kDa protein OS=Homo sapiens GN=SRP9 PE=1 SV=2 - [SRP9_HUMAN]	45.35	8
P62913	60S ribosomal protein L11 OS=Homo sapiens GN=RPL11 PE=1 SV=2 - [RL11_HUMAN]	43.26	30
P18621	60S ribosomal protein L17 OS=Homo sapiens GN=RPL17 PE=1 SV=3 - [RL17_HUMAN]	41.85	19
P26373	60S ribosomal protein L13 OS=Homo sapiens GN=RPL13 PE=1 SV=4 - [RL13_HUMAN]	39.81	27
P62241	40S ribosomal protein S8 OS=Homo sapiens GN=RPS8 PE=1 SV=2 - [RS8_HUMAN]	37.98	12
Q8TEK3	Histone-lysine N-methyltransferase, H3 lysine-79 specific OS=Homo sapiens GN=DOT1L PE=1 SV=2 - [DOT1L_HUMAN]	35.02	145
P62249	40S ribosomal protein S16 OS=Homo sapiens GN=RPS16 PE=1 SV=2 - [RS16_HUMAN]	34.93	16
P62280	40S ribosomal protein S11 OS=Homo sapiens GN=RPS11 PE=1 SV=3 - [RS11_HUMAN]	34.81	19
O60814	Histone H2B type 1-K OS=Homo sapiens GN=HIST1H2BK PE=1 SV=3 - [H2BK_HUMAN]	33.33	9
P39019	40S ribosomal protein S19 OS=Homo sapiens GN=RPS19 PE=1 SV=2 - [RS19_HUMAN]	32.41	8
Q06830	Peroxisomal protein OS=Homo sapiens GN=PRDX1 PE=1 SV=1 - [PRDX1_HUMAN]	32.16	10
P62937	Peptidyl-prolyl cis-trans isomerase A OS=Homo sapiens GN=PPIA PE=1 SV=2 - [PPIA_HUMAN]	31.52	13
P15880	40S ribosomal protein S2 OS=Homo sapiens GN=RPS2 PE=1 SV=2 - [RS2_HUMAN]	31.40	18
P68431	Histone H3.1 OS=Homo sapiens GN=HIST1H3A PE=1 SV=2 - [H31_HUMAN]	30.88	7
P29372	DNA-3-methyladenine glycosylase OS=Homo sapiens GN=MPG PE=1 SV=3 - [3MG_HUMAN]	30.87	15
P23396	40S ribosomal protein S3 OS=Homo sapiens GN=RPS3 PE=1 SV=2 - [RS3_HUMAN]	30.86	14
P62753	40S ribosomal protein S6 OS=Homo sapiens GN=RPS6 PE=1 SV=1 - [RS6_HUMAN]	30.52	20
P62857	40S ribosomal protein S28 OS=Homo sapiens GN=RPS28 PE=1 SV=1 - [RS28_HUMAN]	30.43	4
Q9Y3C1	Nucleolar protein 16 OS=Homo sapiens GN=NOP16 PE=1 SV=2 - [NOP16_HUMAN]	30.34	11
P46779	60S ribosomal protein L28 OS=Homo sapiens GN=RPL28 PE=1 SV=3 - [RL28_HUMAN]	29.93	12
F22ZW6	Non-histone chromosomal protein HMG-14 OS=Homo sapiens GN=HMGN1 PE=1 SV=1 - [F22ZW6_HUMAN]	29.55	1
P60866	60S ribosomal protein S20 OS=Homo sapiens GN=RPS20 PE=1 SV=1 - [RS20_HUMAN]	29.41	15
P62805	Histone H4 OS=Homo sapiens GN=HIST1H4A PE=1 SV=2 - [H4_HUMAN]	29.13	6
P17844	Probable ATP-dependent RNA helicase DDX5 OS=Homo sapiens GN=DDX5 PE=1 SV=1 - [DDX5_HUMAN]	28.50	43
P62633	Cellular nucleic acid-binding protein OS=Homo sapiens GN=CNBP PE=1 SV=1 - [CNBP_HUMAN]	27.68	7
P61247	40S ribosomal protein S3a OS=Homo sapiens GN=RPS3A PE=1 SV=2 - [RS3A_HUMAN]	27.27	14
Q5TEC6	Histone H3 OS=Homo sapiens GN=HIST2H3P52 PE=1 SV=1 - [Q5TEC6_HUMAN]	27.21	4
P42766	60S ribosomal protein L35 OS=Homo sapiens GN=RPL35 PE=1 SV=2 - [RL35_HUMAN]	26.83	6
Q13885	Tubulin beta-2A chain OS=Homo sapiens GN=TUBB2A PE=1 SV=1 - [TBB2A_HUMAN]	26.74	25
P83881	60S ribosomal protein L36a OS=Homo sapiens GN=RPL36A PE=1 SV=2 - [RL36A_HUMAN]	26.42	10
P84098	60S ribosomal protein L19 OS=Homo sapiens GN=RPL19 PE=1 SV=1 - [RL19_HUMAN]	26.02	12
Q9BQE3	Tubulin alpha-1C chain OS=Homo sapiens GN=TUBA1C PE=1 SV=1 - [TBA1C_HUMAN]	25.61	16
P68363	Tubulin alpha-1B chain OS=Homo sapiens GN=TUBA1B PE=1 SV=1 - [TBA1B_HUMAN]	25.50	16
P27635	60S ribosomal protein L10 OS=Homo sapiens GN=RPL10 PE=1 SV=4 - [RL10_HUMAN]	25.23	13
P62910	60S ribosomal protein L32 OS=Homo sapiens GN=RPL32 PE=1 SV=2 - [RL32_HUMAN]	25.19	8
P11498	Pyruvate carboxylase, mitochondrial OS=Homo sapiens GN=PC PE=1 SV=2 - [PYC_HUMAN]	24.36	53
P63173	60S ribosomal protein L38 OS=Homo sapiens GN=RPL38 PE=1 SV=2 - [RL38_HUMAN]	24.29	3
P05141	ADP/ATP translocase 2 OS=Homo sapiens GN=SLC25A5 PE=1 SV=7 - [ADT2_HUMAN]	24.16	16
Q13085	Acetyl-CoA carboxylase 1 OS=Homo sapiens GN=ACACA PE=1 SV=2 - [ACACA_HUMAN]	23.36	88
Q15366	Poly(rC)-binding protein 2 OS=Homo sapiens GN=PCBP2 PE=1 SV=1 - [PCBP2_HUMAN]	23.29	10
P61978	Heterogeneous nuclear ribonucleoprotein K OS=Homo sapiens GN=HNRNPK PE=1 SV=1 - [HNRPK_HUMAN]	23.11	20
P46776	60S ribosomal protein L27a OS=Homo sapiens GN=RPL27A PE=1 SV=2 - [RL27A_HUMAN]	22.97	7
P62701	40S ribosomal protein S4, X isoform OS=Homo sapiens GN=RPS4X PE=1 SV=2 - [RS4X_HUMAN]	22.81	13
P05165	Propionyl-CoA carboxylase alpha chain, mitochondrial OS=Homo sapiens GN=PCCA PE=1 SV=4 - [PCCA_HUMAN]	22.66	27
P62266	40S ribosomal protein S23 OS=Homo sapiens GN=RPS23 PE=1 SV=3 - [RS23_HUMAN]	22.38	10
Q9BVP2	Guanine nucleotide-binding protein-like 3 OS=Homo sapiens GN=GNL3 PE=1 SV=2 - [GNL3_HUMAN]	22.22	20
Q92841	Probable ATP-dependent RNA helicase DDX17 OS=Homo sapiens GN=DDX17 PE=1 SV=2 - [DDX17_HUMAN]	21.12	32
P68104	Elongation factor 1-alpha 1 OS=Homo sapiens GN=EEF1A1 PE=1 SV=1 - [EF1A1_HUMAN]	20.56	17
P12236	ADP/ATP translocase 3 OS=Homo sapiens GN=SLC25A6 PE=1 SV=4 - [ADT3_HUMAN]	20.47	14
P62987	Ubiquitin-60S ribosomal protein L40 OS=Homo sapiens GN=UBA52 PE=1 SV=2 - [RL40_HUMAN]	20.31	4
P16403	Histone H1.2 OS=Homo sapiens GN=HIST1H1C PE=1 SV=2 - [H12_HUMAN]	19.72	11
P62891	60S ribosomal protein L39 OS=Homo sapiens GN=RPL39 PE=1 SV=2 - [RL39_HUMAN]	19.61	2
P62847	40S ribosomal protein S24 OS=Homo sapiens GN=RPS24 PE=1 SV=1 - [RS24_HUMAN]	19.55	2
Q15365	Poly(rC)-binding protein 1 OS=Homo sapiens GN=PCBP1 PE=1 SV=2 - [PCBP1_HUMAN]	19.38	10

## Appendix-II

Accession	Description	ΣCoverage	Σ# PSMs
O00571	ATP-dependent RNA helicase DDX3X OS=Homo sapiens GN=DDX3X PE=1 SV=3 - [DDX3X_HUMAN]	18.88	29
Q9H5H4	Zinc finger protein 768 OS=Homo sapiens GN=ZNF768 PE=1 SV=2 - [ZN768_HUMAN]	17.78	15
Q9P258	Protein RCC2 OS=Homo sapiens GN=RCC2 PE=1 SV=2 - [RCC2_HUMAN]	17.05	13
P62861	40S ribosomal protein S30 OS=Homo sapiens GN=FAU PE=1 SV=1 - [RS30_HUMAN]	16.95	2
P60709	Actin, cytoplasmic 1 OS=Homo sapiens GN=ACTB PE=1 SV=1 - [ACTB_HUMAN]	16.80	9
P62263	40S ribosomal protein S14 OS=Homo sapiens GN=RPS14 PE=1 SV=3 - [RS14_HUMAN]	16.56	4
O15235	28S ribosomal protein S12, mitochondrial OS=Homo sapiens GN=MRPS12 PE=1 SV=1 - [RT12_HUMAN]	15.94	2
P04264	Keratin, type II cytoskeletal 1 OS=Homo sapiens GN=KRT1 PE=1 SV=6 - [K2C1_HUMAN]	15.84	15
P05204	Non-histone chromosomal protein HMG-17 OS=Homo sapiens GN=HMGN2 PE=1 SV=3 - [HMGN2_HUMAN]	15.56	2
P49207	60S ribosomal protein L34 OS=Homo sapiens GN=RPL34 PE=1 SV=3 - [RL34_HUMAN]	15.38	2
P62851	40S ribosomal protein S25 OS=Homo sapiens GN=RPS25 PE=1 SV=1 - [RS25_HUMAN]	15.20	3
P09874	Poly [ADP-ribose] polymerase 1 OS=Homo sapiens GN=PARP1 PE=1 SV=4 - [PARP1_HUMAN]	15.09	27
P78549	Endonuclease III-like protein 1 OS=Homo sapiens GN=NTHL1 PE=1 SV=2 - [NTH_HUMAN]	14.42	7
ESRHH7	ADP-ribosylation factor GTPase-activating protein 1 (Fragment) OS=Homo sapiens GN=ARFGAP1 PE=1 SV=1 - [ESRHH7_HUMAN]	14.38	1
P62829	60S ribosomal protein L23 OS=Homo sapiens GN=RPL23 PE=1 SV=1 - [RL23_HUMAN]	14.29	2
Q96EY4	Translation machinery-associated protein 16 OS=Homo sapiens GN=TMA16 PE=1 SV=2 - [TMA16_HUMAN]	14.29	7
P26641	Elongation factor 1-gamma OS=Homo sapiens GN=EEF1G PE=1 SV=3 - [EF1G_HUMAN]	13.50	8
O95232	Luc7-like protein 3 OS=Homo sapiens GN=LUC7L3 PE=1 SV=2 - [LC7L3_HUMAN]	13.43	11
Q96RQ3	Methylcrotonoyl-CoA carboxylase subunit alpha, mitochondrial OS=Homo sapiens GN=MCCC1 PE=1 SV=3 - [MCCA_HUMAN]	13.38	14
Q9QZV4	Eukaryotic translation initiation factor 5A-2 OS=Homo sapiens GN=EIF5A2 PE=1 SV=3 - [IF5A2_HUMAN]	13.07	3
Q02543	60S ribosomal protein L18a OS=Homo sapiens GN=RPL18A PE=1 SV=2 - [RL18A_HUMAN]	13.07	4
C9JD17	U2 snRNP-associated SURP motif-containing protein (Fragment) OS=Homo sapiens GN=U2SURP PE=1 SV=3 - [C9JD17_HUMAN]	12.98	1
Q5JTH9	RRP12-like protein OS=Homo sapiens GN=RRP12 PE=1 SV=2 - [RRP12_HUMAN]	12.95	29
Q86T24	Transcriptional regulator Kaiso OS=Homo sapiens GN=ZBTB33 PE=1 SV=2 - [KAISO_HUMAN]	12.65	16
P42167	Lamina-associated polypeptide 2, isoforms beta/gamma OS=Homo sapiens GN=TMPO PE=1 SV=2 - [LAP2B_HUMAN]	12.56	6
P62273	60S ribosomal protein S29 OS=Homo sapiens GN=RPS29 PE=1 SV=2 - [RS29_HUMAN]	12.50	3
P0C055	Histone H2A. Z OS=Homo sapiens GN=H2AFZ PE=1 SV=2 - [H2AZ_HUMAN]	12.50	3
O00483	Cytochrome c oxidase subunit NDUF4A OS=Homo sapiens GN=NDUFA4 PE=1 SV=1 - [NDUA4_HUMAN]	12.35	1
P04080	Cystatin-B OS=Homo sapiens GN=CSTB PE=1 SV=2 - [CYTB_HUMAN]	12.24	2
Q5T280	Putative methyltransferase C9orf114 OS=Homo sapiens GN=C9orf114 PE=1 SV=3 - [C1114_HUMAN]	11.97	7
P06733	Alpha-enolase OS=Homo sapiens GN=ENO1 PE=1 SV=2 - [ENOA_HUMAN]	11.52	7
P50454	Serpin H1 OS=Homo sapiens GN=SERPINH1 PE=1 SV=2 - [SERPH_HUMAN]	11.48	6
P62269	40S ribosomal protein S18 OS=Homo sapiens GN=RPS18 PE=1 SV=3 - [RS18_HUMAN]	11.18	2
Q16629	Serine/arginine-rich splicing factor 7 OS=Homo sapiens GN=SRSF7 PE=1 SV=1 - [SRSF7_HUMAN]	10.92	2
P06748	Nucleophosmin OS=Homo sapiens GN=NPM1 PE=1 SV=2 - [NPM_HUMAN]	10.88	8
P10599-2	Isoform 2 of Thioredoxin OS=Homo sapiens GN=TXN - [THIO_HUMAN]	10.59	1
Q02878	60S ribosomal protein L6 OS=Homo sapiens GN=RPL6 PE=1 SV=3 - [RL6_HUMAN]	10.42	5
P50990	T-complex protein 1 subunit theta OS=Homo sapiens GN=CCT8 PE=1 SV=4 - [TCPO_HUMAN]	9.85	9
Q9P016	Thymocyte nuclear protein 1 OS=Homo sapiens GN=THYN1 PE=1 SV=1 - [THYN1_HUMAN]	9.78	3
P61513	60S ribosomal protein L37a OS=Homo sapiens GN=RPL37A PE=1 SV=2 - [RL37A_HUMAN]	9.78	2
P04406	Glyceraldehyde-3-phosphate dehydrogenase OS=Homo sapiens GN=GAPDH PE=1 SV=3 - [G3P_HUMAN]	9.55	4
P42677	40S ribosomal protein S27 OS=Homo sapiens GN=RPS27 PE=1 SV=3 - [RS27_HUMAN]	9.52	2
P47914	60S ribosomal protein L29 OS=Homo sapiens GN=RPL29 PE=1 SV=2 - [RL29_HUMAN]	9.43	2
P61927	60S ribosomal protein L37 OS=Homo sapiens GN=RPL37 PE=1 SV=2 - [RL37_HUMAN]	9.28	2
P62314	Small nuclear ribonucleoprotein Sm D1 OS=Homo sapiens GN=SNRPD1 PE=1 SV=1 - [SMD1_HUMAN]	9.24	1
P62917	60S ribosomal protein L8 OS=Homo sapiens GN=RPL8 PE=1 SV=2 - [RL8_HUMAN]	8.95	3
Q19T08	Endothelial cell-specific chemotaxis regulator OS=Homo sapiens GN=ECSCR PE=1 SV=1 - [ECSCR_HUMAN]	8.78	1
P48730	Casein kinase I isoform delta OS=Homo sapiens GN=CSNK1D PE=1 SV=2 - [KC1D_HUMAN]	8.67	7
P35527	Keratin, type I cytoskeletal 9 OS=Homo sapiens GN=KRT9 PE=1 SV=3 - [KIC9_HUMAN]	8.51	7
P56270	Myc-associated zinc finger protein OS=Homo sapiens GN=MAZ PE=1 SV=1 - [MAZ_HUMAN]	8.39	9
P62424	60S ribosomal protein L7a OS=Homo sapiens GN=RPL7A PE=1 SV=2 - [RL7A_HUMAN]	8.27	3
P63167	Dynein light chain 1, cytoplasmic OS=Homo sapiens GN=DYNLL1 PE=1 SV=1 - [DYLL1_HUMAN]	7.87	1
P62854	40S ribosomal protein S26 OS=Homo sapiens GN=RPS26 PE=1 SV=3 - [RS26_HUMAN]	7.83	1
C9JPD0	Neurexophilin-1 (Fragment) OS=Homo sapiens GN=NXPH1 PE=4 SV=1 - [C9JPD0_HUMAN]	7.69	1
Q15233	Non-POU domain-containing octamer-binding protein OS=Homo sapiens GN=NONO PE=1 SV=4 - [NONO_HUMAN]	7.64	4
P22087	rRNA 2'-O-methyltransferase fibrillarin OS=Homo sapiens GN=FBL PE=1 SV=2 - [FBL_HUMAN]	7.48	4
Q8IYL3	UPF0688 protein C1orf174 OS=Homo sapiens GN=C1orf174 PE=1 SV=2 - [CA174_HUMAN]	7.41	2
Q8IWS0	PHD finger protein 6 OS=Homo sapiens GN=PHF6 PE=1 SV=1 - [PHF6_HUMAN]	7.40	5
P55197	Protein AF-10 OS=Homo sapiens GN=MLLT10 PE=1 SV=2 - [AF10_HUMAN]	7.30	10

Accession	Description	ΣCoverage	Σ# PSMs
P32969	60S ribosomal protein L9 OS=Homo sapiens GN=RPL9 PE=1 SV=1 - [RL9_HUMAN]	7.29	2
P62899	60S ribosomal protein L31 OS=Homo sapiens GN=RPL31 PE=1 SV=1 - [RL31_HUMAN]	7.20	2
Q8NC51	Plasminogen activator inhibitor 1 RNA-binding protein OS=Homo sapiens GN=SERBP1 PE=1 SV=2 - [PAIRB_HUMAN]	7.11	4
Q9P2Y5-2	Isoform 2 of UV radiation resistance-associated gene protein OS=Homo sapiens GN=UVRAG - [UVRAG_HUMAN]	6.73	1
P07305	Histone H1.0 OS=Homo sapiens GN=H1F0 PE=1 SV=3 - [H10_HUMAN]	6.70	1
P26599	Polypyrimidine tract-binding protein 1 OS=Homo sapiens GN=PTBP1 PE=1 SV=1 - [PTBP1_HUMAN]	6.59	9
P07910	Heterogeneous nuclear ribonucleoproteins C1/C2 OS=Homo sapiens GN=HNRNPC PE=1 SV=4 - [HNRPC_HUMAN]	6.54	3
Q9Y3B4	Splicing factor 3B subunit 6 OS=Homo sapiens GN=SF3B6 PE=1 SV=1 - [SF3B6_HUMAN]	6.40	1
P61254	60S ribosomal protein L26 OS=Homo sapiens GN=RPL26 PE=1 SV=1 - [RL26_HUMAN]	6.21	2
Q96127	Zinc finger protein 625 OS=Homo sapiens GN=ZNF625 PE=2 SV=1 - [ZNF625_HUMAN]	6.21	2
P09651	Heterogeneous nuclear ribonucleoprotein A1 OS=Homo sapiens GN=HNRNPA1 PE=1 SV=5 - [ROA1_HUMAN]	6.18	3
P62244	40S ribosomal protein S15a OS=Homo sapiens GN=RPS15A PE=1 SV=2 - [RS15A_HUMAN]	6.15	1
P31943	Heterogeneous nuclear ribonucleoprotein H OS=Homo sapiens GN=HNRNPH1 PE=1 SV=4 - [HNRHL_HUMAN]	6.01	5
Q00839	Heterogeneous nuclear ribonucleoprotein U OS=Homo sapiens GN=HNRNPU PE=1 SV=6 - [HNRPU_HUMAN]	5.94	10
P61313	60S ribosomal protein L15 OS=Homo sapiens GN=RPL15 PE=1 SV=2 - [RL15_HUMAN]	5.88	1
Q98Z64	Nucleolar GTP-binding protein 1 OS=Homo sapiens GN=GTPBP4 PE=1 SV=3 - [NOG1_HUMAN]	5.84	8
P83731	60S ribosomal protein L24 OS=Homo sapiens GN=RPL24 PE=1 SV=1 - [RL24_HUMAN]	5.73	3
C9JGC1	Zinc finger protein neuro-d4 (Fragment) OS=Homo sapiens GN=DPF1 PE=4 SV=1 - [C9JGC1_HUMAN]	5.71	1
P22626	Heterogeneous nuclear ribonucleoproteins A2/B1 OS=Homo sapiens GN=HNRNPA2B1 PE=1 SV=2 - [ROA2_HUMAN]	5.67	3
Q9Y324	rRNA-processing protein FCF1 homolog OS=Homo sapiens GN=FCF1 PE=2 SV=1 - [FCF1_HUMAN]	5.56	2
P49959	Double-strand break repair protein MRE11A OS=Homo sapiens GN=MRE11A PE=1 SV=3 - [MRE11_HUMAN]	5.51	6
Q95EK4	THAP domain-containing protein 11 OS=Homo sapiens GN=THAP11 PE=1 SV=2 - [THA11_HUMAN]	5.41	3
O43390	Heterogeneous nuclear ribonucleoprotein R OS=Homo sapiens GN=HNRNPR PE=1 SV=1 - [HNRPR_HUMAN]	5.37	4
Q07020	60S ribosomal protein L18 OS=Homo sapiens GN=RPL18 PE=1 SV=2 - [RL18_HUMAN]	5.32	2
P07355	Annexin A2 OS=Homo sapiens GN=ANXA2 PE=1 SV=2 - [ANXA2_HUMAN]	5.31	3
P12270	Nucleoprotein TPR OS=Homo sapiens GN=TPR PE=1 SV=3 - [TPR_HUMAN]	5.25	21
Q9Y4X4	Kruppel-like factor 12 OS=Homo sapiens GN=KLF12 PE=1 SV=2 - [KLF12_HUMAN]	5.22	3
Q07666	KH domain-containing, RNA-binding, signal transduction-associated protein 1 OS=Homo sapiens GN=KHDRB51 PE=1 SV=1 - [KHDR1_HUMAN]	5.19	3
P20719	Homeobox protein Hox-A5 OS=Homo sapiens GN=HOXA5 PE=1 SV=2 - [HXA5_HUMAN]	5.19	2
Q16527	Cysteine and glycine-rich protein 2 OS=Homo sapiens GN=CSR2 PE=1 SV=3 - [CSR2_HUMAN]	5.18	1
O76021	Ribosomal L1 domain-containing protein 1 OS=Homo sapiens GN=RSL1D1 PE=1 SV=3 - [RL1D1_HUMAN]	5.10	4
Q6HA08	Astacin-like metalloendopeptidase OS=Homo sapiens GN=ASTL PE=1 SV=4 - [ASTL_HUMAN]	5.10	1
P14866	Heterogeneous nuclear ribonucleoprotein L OS=Homo sapiens GN=HNRNPL PE=1 SV=2 - [HNRPL_HUMAN]	5.09	4
P35659	Protein DEK OS=Homo sapiens GN=DEK PE=1 SV=1 - [DEK_HUMAN]	5.07	3
P08107	Heat shock 70 kDa protein 1A/1B OS=Homo sapiens GN=HSPA1A PE=1 SV=5 - [HSP71_HUMAN]	4.99	6
P35908	Keratin, type II cytoskeletal 2 epidermal OS=Homo sapiens GN=KRT2 PE=1 SV=2 - [K2E2_HUMAN]	4.85	6
P27695	DNA-(apurinic or apyrimidinic site) lyase OS=Homo sapiens GN=APEX1 PE=1 SV=2 - [APEX1_HUMAN]	4.72	2
P25705	ATP synthase subunit alpha, mitochondrial OS=Homo sapiens GN=ATP5A1 PE=1 SV=1 - [ATPA_HUMAN]	4.70	3
Q9H0C8	Integrin-linked kinase-associated serine/threonine phosphatase 2C OS=Homo sapiens GN=ILKAP PE=1 SV=1 - [ILKAP_HUMAN]	4.59	2
P35637	RNA-binding protein FUS OS=Homo sapiens GN=FUS PE=1 SV=1 - [FUS_HUMAN]	4.56	1
Q96P11	Probable 28S rRNA (cytosine-C(5))-methyltransferase OS=Homo sapiens GN=NSUN5 PE=1 SV=2 - [NSUN5_HUMAN]	4.43	3
Q14119	Vascular endothelial zinc finger 1 OS=Homo sapiens GN=VEZF1 PE=1 SV=2 - [VEZF1_HUMAN]	4.41	2
Q9ULJ8-2	Isoform 2 of Neurabin-1 OS=Homo sapiens GN=PPP1R9A - [NEB1_HUMAN]	4.39	1
P23246	Splicing factor, proline- and glutamine-rich OS=Homo sapiens GN=SFPO PE=1 SV=2 - [SFPO_HUMAN]	4.38	3
Q9Y2R4	Probable ATP-dependent RNA helicase DDX52 OS=Homo sapiens GN=DDX52 PE=1 SV=3 - [DDX52_HUMAN]	4.34	3
P08670	Vimentin OS=Homo sapiens GN=VIM PE=1 SV=4 - [VIME_HUMAN]	4.29	4
P02768	Serum albumin OS=Homo sapiens GN=ALB PE=1 SV=2 - [ALBU_HUMAN]	4.27	4
P46783	40S ribosomal protein S10 OS=Homo sapiens GN=RPS10 PE=1 SV=1 - [RS10_HUMAN]	4.24	1
P55265	Double-stranded RNA-specific adenosine deaminase OS=Homo sapiens GN=ADAR PE=1 SV=4 - [DSRAD_HUMAN]	4.16	10
Q00325	Phosphate carrier protein, mitochondrial OS=Homo sapiens GN=SLC25A3 PE=1 SV=2 - [MPCP_HUMAN]	4.14	3
P46109	Crk-like protein OS=Homo sapiens GN=CRKL PE=1 SV=1 - [CRKL_HUMAN]	3.96	2
Q9BQ75	Protein CMS1 OS=Homo sapiens GN=CMS1 PE=1 SV=2 - [CMS1_HUMAN]	3.94	1
P25205	DNA replication licensing factor MCM3 OS=Homo sapiens GN=MCM3 PE=1 SV=3 - [MCM3_HUMAN]	3.84	7
Q6PK04	Coiled-coil domain-containing protein 137 OS=Homo sapiens GN=CCDC137 PE=1 SV=1 - [CC137_HUMAN]	3.81	2
Q9BQ04	Myb-binding protein 1A OS=Homo sapiens GN=MYBBP1A PE=1 SV=2 - [MBB1A_HUMAN]	3.77	9
O75475	PC4 and SFRS1-interacting protein OS=Homo sapiens GN=PSIP1 PE=1 SV=1 - [PSIP1_HUMAN]	3.77	5
Q9UQ80	Sex comb on midleg-like protein 2 OS=Homo sapiens GN=SCML2 PE=1 SV=1 - [SCML2_HUMAN]	3.71	5
P62826	GTP-binding nuclear protein Ran OS=Homo sapiens GN=RAN PE=1 SV=3 - [RAN_HUMAN]	3.70	2
P33993	DNA replication licensing factor MCM7 OS=Homo sapiens GN=MCM7 PE=1 SV=4 - [MCM7_HUMAN]	3.62	3

Accession	Description	ΣCoverage	Σ# PSMs
P33992	DNA replication licensing factor MCM5 OS=Homo sapiens GN=MCM5 PE=1 SV=5 - [MCM5_HUMAN]	3.54	4
Q9UNQ2	Probable dimethyladenosine transferase OS=Homo sapiens GN=DIMT1 PE=1 SV=1 - [DIM1_HUMAN]	3.51	1
O95478	Ribosome biogenesis protein NSA2 homolog OS=Homo sapiens GN=NSA2 PE=1 SV=1 - [NSA2_HUMAN]	3.46	1
Q8N3J9	Zinc finger protein 664 OS=Homo sapiens GN=ZNF664 PE=2 SV=1 - [ZNF664_HUMAN]	3.45	2
Q9Y511	U3 small nucleolar RNA-associated protein 18 homolog OS=Homo sapiens GN=LTP18 PE=1 SV=3 - [LTP18_HUMAN]	3.24	2
P55198	Protein AF-17 OS=Homo sapiens GN=MLL16 PE=1 SV=2 - [AF17_HUMAN]	3.20	7
Q9Y265	RuvB-like 1 OS=Homo sapiens GN=RUVBL1 PE=1 SV=1 - [RUVB1_HUMAN]	3.07	2
Q8I2J6	Inactive L-threonine 3-dehydrogenase, mitochondrial OS=Homo sapiens GN=TDH PE=2 SV=1 - [TDH_HUMAN]	3.04	2
P39023	60S ribosomal protein L3 OS=Homo sapiens GN=RPL3 PE=1 SV=2 - [RL3_HUMAN]	2.98	2
P52655	Transcription initiation factor IIA subunit 1 OS=Homo sapiens GN=GTF2A1 PE=1 SV=1 - [TF2AA_HUMAN]	2.93	1
Q14684	Ribosomal RNA processing protein 1 homolog B OS=Homo sapiens GN=RRP1B PE=1 SV=3 - [RRP1B_HUMAN]	2.90	3
Q13148	TAR DNA-binding protein 43 OS=Homo sapiens GN=TARDBP PE=1 SV=1 - [TARDBP_HUMAN]	2.90	2
Q14980	Nuclear mitotic apparatus protein 1 OS=Homo sapiens GN=NUMA1 PE=1 SV=2 - [NUMA1_HUMAN]	2.88	7
P27348	14-3-3 protein theta OS=Homo sapiens GN=YWHAQ PE=1 SV=1 - [1433T_HUMAN]	2.86	1
Q5T440	Putative transferase CAF17, mitochondrial OS=Homo sapiens GN=IBAS7 PE=1 SV=1 - [CAF17_HUMAN]	2.81	1
P17482	Homeobox protein Hox-B9 OS=Homo sapiens GN=HOXB9 PE=1 SV=2 - [HXB9_HUMAN]	2.80	2
Q9HC62	Sentrin-specific protease 2 OS=Homo sapiens GN=SEN2 PE=1 SV=3 - [SEN2_HUMAN]	2.72	1
Q00712	Nuclear factor 1 B-type OS=Homo sapiens GN=NFIB PE=1 SV=2 - [NFIB_HUMAN]	2.62	1
Q7L354	Zinc finger protein 771 OS=Homo sapiens GN=ZNF771 PE=1 SV=1 - [ZNF771_HUMAN]	2.52	2
O75182	Paired amphipathic helix protein Sin3b OS=Homo sapiens GN=SIN3B PE=1 SV=2 - [SIN3B_HUMAN]	2.50	2
O94900	Thymocyte selection-associated high mobility group box protein TOX OS=Homo sapiens GN=TOX PE=2 SV=3 - [TOX_HUMAN]	2.47	2
P18754	Regulator of chromosome condensation OS=Homo sapiens GN=RCC1 PE=1 SV=1 - [RCC1_HUMAN]	2.38	2
O00148	ATP-dependent RNA helicase DDX39A OS=Homo sapiens GN=DDX39A PE=1 SV=2 - [DX39A_HUMAN]	2.34	1
Q03111	Protein ENL OS=Homo sapiens GN=MLLT1 PE=1 SV=2 - [ENL_HUMAN]	2.33	2
Q8WVMD	Dimethyladenosine transferase 1, mitochondrial OS=Homo sapiens GN=TFB1M PE=1 SV=1 - [TFB1M_HUMAN]	2.31	1
P42568	Protein AF-9 OS=Homo sapiens GN=MLL3 PE=1 SV=2 - [AF9_HUMAN]	2.29	2
Q6P087	RNA pseudouridylylase synthase domain-containing protein 3 OS=Homo sapiens GN=RPUSD3 PE=1 SV=3 - [RUSD3_HUMAN]	2.28	1
P04075	Fructose-bisphosphate aldolase A OS=Homo sapiens GN=ALDOA PE=1 SV=2 - [ALDOA_HUMAN]	2.20	2
Q8HWS3	DNA-binding protein RFX6 OS=Homo sapiens GN=RFX6 PE=1 SV=2 - [RFX6_HUMAN]	2.16	2
Q43246	DNA mismatch repair protein Msh2 OS=Homo sapiens GN=MSH2 PE=1 SV=1 - [MSH2_HUMAN]	2.14	3
P26368	Splicing factor U2AF 65 kDa subunit OS=Homo sapiens GN=U2AF2 PE=1 SV=4 - [U2AF2_HUMAN]	2.11	1
Q14498	RNA-binding protein 39 OS=Homo sapiens GN=RBM39 PE=1 SV=2 - [RBM39_HUMAN]	2.08	2
P21333	Filamin-A OS=Homo sapiens GN=FLNA PE=1 SV=4 - [FLNA_HUMAN]	2.08	6
Q9NPA5-2	Isoform 2 of Zinc finger protein 64 homolog, isoforms 1 and 2 OS=Homo sapiens GN=ZFP64 - [ZFP64_HUMAN]	1.91	1
Q8WVY3	U4/U6 small nuclear ribonucleoprotein Prp31 OS=Homo sapiens GN=PRPF31 PE=1 SV=2 - [PRP31_HUMAN]	1.80	1
O00541	Pescadillo homolog OS=Homo sapiens GN=PES1 PE=1 SV=1 - [PESC_HUMAN]	1.70	2
Q9NR30	Nucleolar RNA helicase 2 OS=Homo sapiens GN=DDX21 PE=1 SV=5 - [DDX21_HUMAN]	1.66	1
Q8WVQ5-3	Isoform 3 of Microprocessor complex subunit DGC8 OS=Homo sapiens GN=DGC8 - [DGC8_HUMAN]	1.62	1
O00264	SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A member 5 OS=Homo sapiens GN=SMARCA5 PE=1 SV=1 - [SMCA5_HUMAN]	1.62	2
Q8NFV8	N-acylneuraminate cytidyllyltransferase OS=Homo sapiens GN=CMAS PE=1 SV=2 - [NEUA_HUMAN]	1.61	2
Q86V48	Leucine zipper protein 1 OS=Homo sapiens GN=LUZP1 PE=1 SV=2 - [LUZP1_HUMAN]	1.58	2
Q6ZND8	Putative zinc finger protein 66 OS=Homo sapiens GN=ZNF66 PE=5 SV=3 - [ZNF66_HUMAN]	1.57	2
Q92945	Far upstream element-binding protein 2 OS=Homo sapiens GN=KHSRP PE=1 SV=4 - [FUBP2_HUMAN]	1.55	1
Q78527	DNA-dependent protein kinase catalytic subunit OS=Homo sapiens GN=PRKDC PE=1 SV=3 - [PRKDC_HUMAN]	1.53	11
P49792	E3 SUMO-protein ligase RanBP2 OS=Homo sapiens GN=RANBP2 PE=1 SV=2 - [RBP2_HUMAN]	1.52	7
Q8WVW6	High affinity immunoglobulin alpha and immunoglobulin mu Fc receptor OS=Homo sapiens GN=FCAMR PE=1 SV=1 - [FCAMR_HUMAN]	1.50	1
Q99959	Plakophilin-2 OS=Homo sapiens GN=PKP2 PE=1 SV=2 - [PKP2_HUMAN]	1.48	2
O00567	Nucleolar protein 56 OS=Homo sapiens GN=NOPS6 PE=1 SV=4 - [NOPS6_HUMAN]	1.35	2
A0A087WU6	DNA polymerase alpha catalytic subunit OS=Homo sapiens GN=POLA1 PE=4 SV=1 - [A0A087WU64_HUMAN]	1.30	3
O75400	Pre-mRNA-processing factor 40 homolog A OS=Homo sapiens GN=PRPF40A PE=1 SV=2 - [PR40A_HUMAN]	1.25	1
P49711	Transcriptional repressor CTCF OS=Homo sapiens GN=CTCF PE=1 SV=1 - [CTCF_HUMAN]	1.24	2
P52272	Heterogeneous nuclear ribonucleoprotein M OS=Homo sapiens GN=HNRNPM PE=1 SV=3 - [HNRPM_HUMAN]	1.23	1
Q7Z2T5	TRMT1-like protein OS=Homo sapiens GN=TRMT1L PE=1 SV=2 - [TRMT1L_HUMAN]	1.23	1
Q14839	Chromodomain-helicase-DNA-binding protein 4 OS=Homo sapiens GN=CHD4 PE=1 SV=2 - [CHD4_HUMAN]	1.20	3
P05107	Integrin beta-2 OS=Homo sapiens GN=ITGB2 PE=1 SV=2 - [ITB2_HUMAN]	1.17	1
O00411	DNA-directed RNA polymerase, mitochondrial OS=Homo sapiens GN=POLRMT PE=1 SV=2 - [RPOM_HUMAN]	1.14	2
P11387	DNA topoisomerase 1 OS=Homo sapiens GN=TOP1 PE=1 SV=2 - [TOP1_HUMAN]	1.05	2
Q9NVP1	ATP-dependent RNA helicase DDX18 OS=Homo sapiens GN=DDX18 PE=1 SV=2 - [DDX18_HUMAN]	1.04	1
Q96T88	E3 ubiquitin-protein ligase UHRF1 OS=Homo sapiens GN=UHRF1 PE=1 SV=1 - [UHRF1_HUMAN]	1.01	1
E7EP10	Inhibitor of Bruton tyrosine kinase OS=Homo sapiens GN=IBTK PE=1 SV=1 - [E7EP10_HUMAN]	0.97	1
Q08211	ATP-dependent RNA helicase A OS=Homo sapiens GN=DHX9 PE=1 SV=4 - [DHX9_HUMAN]	0.94	2
Q8I2X4	Transcription initiation factor TFIID subunit 1-like OS=Homo sapiens GN=TAF1L PE=1 SV=1 - [TAF1L_HUMAN]	0.93	2
Q8NEB9	Phosphatidylinositol 3-kinase catalytic subunit type 3 OS=Homo sapiens GN=PIK3C3 PE=1 SV=1 - [PK3C3_HUMAN]	0.90	1
Q9NZM1	Myoferlin OS=Homo sapiens GN=MYOF PE=1 SV=1 - [MYOF_HUMAN]	0.87	1
Q3KQU3	MAP7 domain-containing protein 1 OS=Homo sapiens GN=MAP7D1 PE=1 SV=1 - [MA7D1_HUMAN]	0.83	2
P13639	Elongation factor 2 OS=Homo sapiens GN=EEF2 PE=1 SV=4 - [EF2_HUMAN]	0.82	2
O60885	Bromodomain-containing protein 4 OS=Homo sapiens GN=BRD4 PE=1 SV=2 - [BRD4_HUMAN]	0.73	2
Q9ULM3	YEATS domain-containing protein 2 OS=Homo sapiens GN=YEATS2 PE=1 SV=2 - [YETS2_HUMAN]	0.70	2
O60241	Brain-specific angiogenesis inhibitor 2 OS=Homo sapiens GN=BAI2 PE=2 SV=2 - [BAI2_HUMAN]	0.69	2
Q5SW79	Centrosomal protein of 170 kDa OS=Homo sapiens GN=CEP170 PE=1 SV=1 - [CE170_HUMAN]	0.57	1
Q5VYK3	Proteasome-associated protein ECM29 homolog OS=Homo sapiens GN=ECM29 PE=1 SV=2 - [ECM29_HUMAN]	0.38	1
P46013	Antigen KI-67 OS=Homo sapiens GN=MKI67 PE=1 SV=2 - [KI67_HUMAN]	0.28	2
Q5UIP0	Telomere-associated protein RIF1 OS=Homo sapiens GN=RIF1 PE=1 SV=2 - [RIF1_HUMAN]	0.28	1



**Table 19. Raw data of BioID experiment: BioID-DOT1L-mut infected HEK293T cells, 1<sup>st</sup> reading**

Accession	Description	ΣCoverage	Σ# PSMs
P62888	60S ribosomal protein L30 OS=Homo sapiens GN=RPL30 PE=1 SV=2 - [RL30_HUMAN]	55.65	9
P08670	Vimentin OS=Homo sapiens GN=VIM PE=1 SV=4 - [VIME_HUMAN]	44.64	44
P46781	40S ribosomal protein S9 OS=Homo sapiens GN=RPS9 PE=1 SV=3 - [RS9_HUMAN]	42.27	30
P49458	Signal recognition particle 9 kDa protein OS=Homo sapiens GN=SRP9 PE=1 SV=2 - [SRP09_HUMAN]	41.86	9
P18621	60S ribosomal protein L17 OS=Homo sapiens GN=RPL17 PE=1 SV=3 - [RL17_HUMAN]	41.85	18
P62913	60S ribosomal protein L11 OS=Homo sapiens GN=RPL11 PE=1 SV=2 - [RL11_HUMAN]	41.57	26
P62280	40S ribosomal protein S11 OS=Homo sapiens GN=RPS11 PE=1 SV=3 - [RS11_HUMAN]	41.14	21
Q8N257	Histone H2B type 3-B OS=Homo sapiens GN=HIST3H2BB PE=1 SV=3 - [H2B3B_HUMAN]	39.68	18
O60814	Histone H2B type 1-K OS=Homo sapiens GN=HIST1H2BK PE=1 SV=3 - [H2B1K_HUMAN]	39.68	18
P62805	Histone H4 OS=Homo sapiens GN=HIST1H4A PE=1 SV=2 - [H4_HUMAN]	38.83	10
P22626	Heterogeneous nuclear ribonucleoproteins A2/B1 OS=Homo sapiens GN=HNRNPA2B1 PE=1 SV=2 - [ROA2_HUMAN]	38.24	37
P62987	Ubiquitin-60S ribosomal protein L40 OS=Homo sapiens GN=UBA52 PE=1 SV=2 - [RL40_HUMAN]	35.16	7
P62937	Peptidyl-prolyl cis-trans isomerase A OS=Homo sapiens GN=PPIA PE=1 SV=2 - [PPIA_HUMAN]	35.15	13
P26373	60S ribosomal protein L13 OS=Homo sapiens GN=RPL13 PE=1 SV=4 - [RL13_HUMAN]	35.07	26
P04406	Glyceraldehyde-3-phosphate dehydrogenase OS=Homo sapiens GN=GAPDH PE=1 SV=3 - [G3P_HUMAN]	34.93	22
P15880	40S ribosomal protein S2 OS=Homo sapiens GN=RPS2 PE=1 SV=2 - [RS2_HUMAN]	33.45	18
P62701	40S ribosomal protein S4, X isoform OS=Homo sapiens GN=RPS4X PE=1 SV=2 - [RS4X_HUMAN]	33.08	21
Q9B0E3	Tubulin alpha-1C chain OS=Homo sapiens GN=TUBA1C PE=1 SV=1 - [TBA1C_HUMAN]	32.96	25
P68363	Tubulin alpha-1B chain OS=Homo sapiens GN=TUBA1B PE=1 SV=1 - [TBA1B_HUMAN]	32.82	25
P09651	Heterogeneous nuclear ribonucleoprotein A1 OS=Homo sapiens GN=HNRNPA1 PE=1 SV=5 - [ROA1_HUMAN]	32.53	30
P46779	60S ribosomal protein L28 OS=Homo sapiens GN=RPL28 PE=1 SV=3 - [RL28_HUMAN]	32.12	8
Q9Y3C1	Nucleolar protein 16 OS=Homo sapiens GN=NOP16 PE=1 SV=2 - [NOP16_HUMAN]	30.90	13
P62241	40S ribosomal protein S8 OS=Homo sapiens GN=RPS8 PE=1 SV=2 - [RS8_HUMAN]	30.29	10
P15531	Nucleoside diphosphate kinase A OS=Homo sapiens GN=NME1 PE=1 SV=1 - [NDKA_HUMAN]	30.26	7
P61978	Heterogeneous nuclear ribonucleoprotein K OS=Homo sapiens GN=HNRNPK PE=1 SV=1 - [HNRPK_HUMAN]	30.24	25
P09874	Poly [ADP-ribose] polymerase 1 OS=Homo sapiens GN=PARP1 PE=1 SV=4 - [PARP1_HUMAN]	29.49	56
Q05680	Peroxiredoxin-1 OS=Homo sapiens GN=PRDX1 PE=1 SV=1 - [PRDX1_HUMAN]	28.64	13
P60866	40S ribosomal protein S20 OS=Homo sapiens GN=RPS20 PE=1 SV=1 - [RS20_HUMAN]	28.57	10
P62753	40S ribosomal protein S6 OS=Homo sapiens GN=RPS6 PE=1 SV=1 - [RS6_HUMAN]	28.11	21
P17844	Probable ATP-dependent RNA helicase DDX5 OS=Homo sapiens GN=DDX5 PE=1 SV=1 - [DDX5_HUMAN]	27.85	41
P29372	DNA-3-methyladenine glycosylase OS=Homo sapiens GN=MPG PE=1 SV=3 - [3MG_HUMAN]	27.85	11
P60709	Actin, cytoplasmic 1 OS=Homo sapiens GN=ACTB PE=1 SV=1 - [ACTB_HUMAN]	27.73	22
P61247	40S ribosomal protein S3a OS=Homo sapiens GN=RPS3A PE=1 SV=2 - [RS3A_HUMAN]	27.65	15
P39019	40S ribosomal protein S19 OS=Homo sapiens GN=RPS19 PE=1 SV=2 - [RS19_HUMAN]	27.59	7
P62249	40S ribosomal protein S16 OS=Homo sapiens GN=RPS16 PE=1 SV=2 - [RS16_HUMAN]	27.40	12
Q96K55	Histone H2A type 1-H OS=Homo sapiens GN=HIST1H2AH PE=1 SV=3 - [H2A1H_HUMAN]	27.34	7
P62829	60S ribosomal protein L23 OS=Homo sapiens GN=RPL23 PE=1 SV=1 - [RL23_HUMAN]	27.14	6
P63267	Actin, gamma-enteric smooth muscle OS=Homo sapiens GN=ACTG2 PE=1 SV=1 - [ACTH_HUMAN]	27.13	22
P07437	Tubulin beta chain OS=Homo sapiens GN=TUBB PE=1 SV=2 - [TBB5_HUMAN]	26.35	19
P04075	Fructose-bisphosphate aldolase A OS=Homo sapiens GN=ALDOA PE=1 SV=2 - [ALDOA_HUMAN]	25.27	17
Q8TEK3	Histone-lysine N-methyltransferase, H3 lysine-79 specific OS=Homo sapiens GN=DOT1L PE=1 SV=2 - [DOT1L_HUMAN]	25.13	87
P83881	60S ribosomal protein L36a OS=Homo sapiens GN=RPL36A PE=1 SV=2 - [RL36A_HUMAN]	24.53	9
P63173	60S ribosomal protein L38 OS=Homo sapiens GN=RPL38 PE=1 SV=2 - [RL38_HUMAN]	24.29	4
P68104	Elongation factor 1-alpha 1 OS=Homo sapiens GN=EEF1A1 PE=1 SV=1 - [EF1A1_HUMAN]	24.03	29
P05165	Propionyl-CoA carboxylase alpha chain, mitochondrial OS=Homo sapiens GN=PCCA PE=1 SV=4 - [PCCA_HUMAN]	23.90	26
P62318	Small nuclear ribonucleoprotein Sm D3 OS=Homo sapiens GN=SNRPD3 PE=1 SV=1 - [SMD3_HUMAN]	23.81	4
P26641	Elongation factor 1-gamma OS=Homo sapiens GN=EEF1G PE=1 SV=3 - [EF1G_HUMAN]	23.57	17
P16402	Histone H1.3 OS=Homo sapiens GN=HIST1H1D PE=1 SV=2 - [H13_HUMAN]	23.08	12
P46776	60S ribosomal protein L27a OS=Homo sapiens GN=RPL27A PE=1 SV=2 - [RL27A_HUMAN]	22.97	6
P07910	Heterogeneous nuclear ribonucleoproteins C1/C2 OS=Homo sapiens GN=HNRNPC PE=1 SV=4 - [HNRPC_HUMAN]	22.88	11
P68431	Histone H3.1 OS=Homo sapiens GN=HIST1H3A PE=1 SV=2 - [H31_HUMAN]	22.79	5
P08238	Heat shock protein HSP 90-beta OS=Homo sapiens GN=HSP90AB1 PE=1 SV=4 - [HS90B_HUMAN]	22.65	26
Q98V92	Guanine nucleotide-binding protein-like 3 OS=Homo sapiens GN=GNL3 PE=1 SV=2 - [GNL3_HUMAN]	22.59	20
P06733	Alpha-enolase OS=Homo sapiens GN=ENO1 PE=1 SV=2 - [ENOA_HUMAN]	22.58	15
P52272	Heterogeneous nuclear ribonucleoprotein M OS=Homo sapiens GN=HNRNPM PE=1 SV=3 - [HNRPM_HUMAN]	22.47	34
P62266	40S ribosomal protein S23 OS=Homo sapiens GN=RPS23 PE=1 SV=3 - [RS23_HUMAN]	22.38	9
Q02543	60S ribosomal protein L18a OS=Homo sapiens GN=RPL18A PE=1 SV=2 - [RL18A_HUMAN]	22.16	8
Q99623	Inhibitor-2 OS=Homo sapiens GN=PHB2 PE=1 SV=2 - [PHB2_HUMAN]	22.07	11

Accession	Description	ΣCoverage	Σ# PSMs
Q9P258	Protein RCC2 OS=Homo sapiens GN=RCC2 PE=1 SV=2 - [RCC2_HUMAN]	21.84	14
P11142	Heat shock cognate 71 kDa protein OS=Homo sapiens GN=HSPA8 PE=1 SV=1 - [HSP7C_HUMAN]	21.83	21
P05141	ADP/ATP translocase 2 OS=Homo sapiens GN=SLC25A5 PE=1 SV=7 - [ADT2_HUMAN]	21.48	16
P51991	Heterogeneous nuclear ribonucleoprotein A3 OS=Homo sapiens GN=HNRNPA3 PE=1 SV=2 - [ROA3_HUMAN]	20.11	11
P27635	60S ribosomal protein L10 OS=Homo sapiens GN=RPL10 PE=1 SV=4 - [RL10_HUMAN]	20.09	9
P62910	60S ribosomal protein L32 OS=Homo sapiens GN=RPL32 PE=1 SV=2 - [RL32_HUMAN]	20.00	7
P08107	Heat shock 70 kDa protein 1A/1B OS=Homo sapiens GN=HSPA1A PE=1 SV=5 - [HSP71_HUMAN]	19.81	25
P11498	Pyruvate carboxylase, mitochondrial OS=Homo sapiens GN=PC PE=1 SV=2 - [PYC_HUMAN]	19.69	38
P62891	60S ribosomal protein L39 OS=Homo sapiens GN=RPL39 PE=1 SV=2 - [RL39_HUMAN]	19.61	2
P62847	40S ribosomal protein S24 OS=Homo sapiens GN=RPS24 PE=1 SV=1 - [RS24_HUMAN]	19.55	4
Q15365	Poly(rC)-binding protein 1 OS=Homo sapiens GN=PCBP1 PE=1 SV=2 - [PCBP1_HUMAN]	19.38	10
Q13151	Heterogeneous nuclear ribonucleoprotein A0 OS=Homo sapiens GN=HNRNPA0 PE=1 SV=1 - [ROA0_HUMAN]	19.02	10
P30050	60S ribosomal protein L12 OS=Homo sapiens GN=RPL12 PE=1 SV=1 - [RL12_HUMAN]	18.79	4
P83731	60S ribosomal protein L24 OS=Homo sapiens GN=RPL24 PE=1 SV=1 - [RL24_HUMAN]	18.47	6
P62269	40S ribosomal protein S18 OS=Homo sapiens GN=RPS18 PE=1 SV=3 - [RS18_HUMAN]	18.42	6
Q96EY4	Translation machinery-associated protein 16 OS=Homo sapiens GN=TMA16 PE=1 SV=2 - [TMA16_HUMAN]	18.23	7
P12236	ADP/ATP translocase 3 OS=Homo sapiens GN=SLC25A6 PE=1 SV=4 - [ADT3_HUMAN]	18.12	14
P23396	40S ribosomal protein S3 OS=Homo sapiens GN=RPS3 PE=1 SV=2 - [RS3_HUMAN]	18.11	6
P62244	40S ribosomal protein S15a OS=Homo sapiens GN=RPS15A PE=1 SV=2 - [RS15A_HUMAN]	17.69	4
P62314	Small nuclear ribonucleoprotein Sm D1 OS=Homo sapiens GN=SNRDP1 PE=1 SV=1 - [SMD1_HUMAN]	17.65	2
P61927	60S ribosomal protein L37 OS=Homo sapiens GN=RPL37 PE=1 SV=1 - [RL37_HUMAN]	17.53	4
P29692	Elongation factor 1-delta OS=Homo sapiens GN=EEF1D PE=1 SV=5 - [EF1D_HUMAN]	17.44	3
Q13085	Acetyl-CoA carboxylase 1 OS=Homo sapiens GN=ACACA PE=1 SV=2 - [ACACA_HUMAN]	17.39	67
P62857	40S ribosomal protein S28 OS=Homo sapiens GN=RPS28 PE=1 SV=1 - [RS28_HUMAN]	17.39	1
O95232	Luc7-like protein 3 OS=Homo sapiens GN=LUC7L3 PE=1 SV=2 - [LC7L3_HUMAN]	17.36	13
P84098	60S ribosomal protein L19 OS=Homo sapiens GN=RPL19 PE=1 SV=1 - [RL19_HUMAN]	17.35	5
A8MW9D	Putative small nuclear ribonucleoprotein G-like protein 15 OS=Homo sapiens GN=SNRPG15 PE=5 SV=2 - [RUXGL_HUMAN]	17.11	2
P62861	40S ribosomal protein S30 OS=Homo sapiens GN=FAU PE=1 SV=1 - [RS30_HUMAN]	16.95	2
O00571	ATP-dependent RNA helicase DDX3X OS=Homo sapiens GN=DDX3X PE=1 SV=3 - [DDX3X_HUMAN]	16.47	22
P35527	Keratin, type I cytoskeletal 9 OS=Homo sapiens GN=KRT9 PE=1 SV=3 - [K1C9_HUMAN]	16.37	19
Q02878	60S ribosomal protein L6 OS=Homo sapiens GN=RPL6 PE=1 SV=3 - [RL6_HUMAN]	15.97	8
Q15366	Poly(rC)-binding protein 2 OS=Homo sapiens GN=PCBP2 PE=1 SV=1 - [PCBP2_HUMAN]	15.89	7
P31943	Heterogeneous nuclear ribonucleoprotein H OS=Homo sapiens GN=HNRNPH1 PE=1 SV=4 - [HNRH1_HUMAN]	15.81	12
P46782	40S ribosomal protein S5 OS=Homo sapiens GN=RPS5 PE=1 SV=4 - [RS5_HUMAN]	15.69	8
P42766	60S ribosomal protein L35 OS=Homo sapiens GN=RPL35 PE=1 SV=2 - [RL35_HUMAN]	15.45	4
Q92841	Probable ATP-dependent RNA helicase DDX17 OS=Homo sapiens GN=DDX17 PE=1 SV=2 - [DDX17_HUMAN]	15.36	21
R84103	Serine/arginine-rich splicing factor 3 OS=Homo sapiens GN=SRSF3 PE=1 SV=1 - [SRSF3_HUMAN]	15.24	4
P62851	40S ribosomal protein S25 OS=Homo sapiens GN=RPS25 PE=1 SV=1 - [RS25_HUMAN]	15.20	4
P62306	Small nuclear ribonucleoprotein F OS=Homo sapiens GN=SNRPF PE=1 SV=1 - [RUXF_HUMAN]	15.12	2
P07900	Heat shock protein HSP 90-alpha OS=Homo sapiens GN=HSP90AA1 PE=1 SV=5 - [HS90A_HUMAN]	14.89	17
P0C055	Histone H2A.Z OS=Homo sapiens GN=H2AFZ PE=1 SV=2 - [H2AZ_HUMAN]	14.84	4
P10809	60 kDa heat shock protein, mitochondrial OS=Homo sapiens GN=HSPD1 PE=1 SV=2 - [CH60_HUMAN]	14.66	15
P25705	ATP synthase subunit alpha, mitochondrial OS=Homo sapiens GN=ATP5A1 PE=1 SV=1 - [ATPA_HUMAN]	14.65	13
P23246	Splicing factor, proline- and glutamine-rich OS=Homo sapiens GN=SFPQ PE=1 SV=2 - [SFPQ_HUMAN]	14.43	17
P50914	60S ribosomal protein L14 OS=Homo sapiens GN=RPL14 PE=1 SV=4 - [RL14_HUMAN]	14.42	5
Q9NR30	Nucleolar RNA helicase 2 OS=Homo sapiens GN=DDX21 PE=1 SV=5 - [DDX21_HUMAN]	14.30	18
Q00839	Heterogeneous nuclear ribonucleoprotein U OS=Homo sapiens GN=HNRNPU PE=1 SV=6 - [HNRPU_HUMAN]	14.18	23
Q07020	60S ribosomal protein L18 OS=Homo sapiens GN=RPL18 PE=1 SV=2 - [RL18_HUMAN]	13.83	4
P78549	Endonuclease III-like protein 1 OS=Homo sapiens GN=NTHL1 PE=1 SV=2 - [NTH_HUMAN]	13.78	4
Q96E39	RNA binding motif protein, X-linked-like-1 OS=Homo sapiens GN=RBMXL1 PE=1 SV=1 - [RMXL1_HUMAN]	13.59	9
Q15233	Non-POU domain-containing octamer-binding protein OS=Homo sapiens GN=NONO PE=1 SV=4 - [NONO_HUMAN]	13.59	15
Q8NC51	Plasminogen activator inhibitor 1 RNA-binding protein OS=Homo sapiens GN=SERBP1 PE=1 SV=2 - [PAIRB_HUMAN]	13.48	7
P62917	60S ribosomal protein L8 OS=Homo sapiens GN=RPL8 PE=1 SV=2 - [RL8_HUMAN]	13.23	5
Q92522	Histone H1x OS=Homo sapiens GN=H1FX PE=1 SV=1 - [H1X_HUMAN]	13.15	3
P60842	Eukaryotic initiation factor 4A-1 OS=Homo sapiens GN=EIF4A1 PE=1 SV=1 - [IF4A1_HUMAN]	13.05	9
Q9Y265	RuvB-like 1 OS=Homo sapiens GN=RUVBL1 PE=1 SV=1 - [RUVBL1_HUMAN]	12.94	9
P26599	Poly(pyrimidine tract-binding protein 1 OS=Homo sapiens GN=PTBP1 PE=1 SV=1 - [PTBP1_HUMAN]	12.81	11
P62424	60S ribosomal protein L7a OS=Homo sapiens GN=RPL7A PE=1 SV=2 - [RL7A_HUMAN]	12.78	6
P62273	40S ribosomal protein S29 OS=Homo sapiens GN=RPS29 PE=1 SV=2 - [RS29_HUMAN]	12.50	2
Q9NY12	H/ACA ribonucleoprotein complex subunit 1 OS=Homo sapiens GN=GAR1 PE=1 SV=1 - [GAR1_HUMAN]	12.44	3
P61254	60S ribosomal protein L26 OS=Homo sapiens GN=RPL26 PE=1 SV=1 - [RL26_HUMAN]	12.41	3
HOYAQ1	KH domain-containing, RNA-binding, signal transduction-associated protein 3 (Fragment) OS=Homo sapiens GN=KHDRBS3 PE=1 SV=1 - [HOYAQ1_HUMAN]	12.31	1
P07195	L-lactate dehydrogenase B chain OS=Homo sapiens GN=LDHB PE=1 SV=2 - [LDHB_HUMAN]	12.28	6
Q43390	Heterogeneous nuclear ribonucleoprotein R OS=Homo sapiens GN=HNRNRP PE=1 SV=1 - [HNRPR_HUMAN]	12.16	12
P13639	Elongation factor 2 OS=Homo sapiens GN=EEF2 PE=1 SV=4 - [EF2_HUMAN]	11.89	19
P14618	Pyruvate kinase PKM OS=Homo sapiens GN=PKM PE=1 SV=4 - [KPYM_HUMAN]	11.86	11
O43175	D-3-phosphoglycerate dehydrogenase OS=Homo sapiens GN=PHGDH PE=1 SV=4 - [SERA_HUMAN]	11.63	9
P14866	Heterogeneous nuclear ribonucleoprotein L OS=Homo sapiens GN=HNRNPL PE=1 SV=2 - [HNRPL_HUMAN]	11.38	12
P22087	rRNA 2'-O-methyltransferase fibrillarin OS=Homo sapiens GN=FBL PE=1 SV=2 - [FBRL_HUMAN]	11.21	6

Accession	Description	ΣCoverage	Σ# PSMs
P18077	60S ribosomal protein L35a OS=Homo sapiens GN=RPL35A PE=1 SV=2 - [RL35A_HUMAN]	10.91	2
P06748	Nucleophosmin OS=Homo sapiens GN=NPM1 PE=1 SV=2 - [NPM_HUMAN]	10.88	7
Q6NWX1	Putative 60S ribosomal protein L13a protein RPL13AP3 OS=Homo sapiens GN=RPL13AP3 PE=5 SV=1 - [R13P3_HUMAN]	10.78	2
P32119	Peroxiredoxin-2 OS=Homo sapiens GN=PRDX2 PE=1 SV=5 - [PRDX2_HUMAN]	10.61	4
P62277	40S ribosomal protein S13 OS=Homo sapiens GN=RPS13 PE=1 SV=2 - [RS13_HUMAN]	10.60	3
P05387	60S acidic ribosomal protein P2 OS=Homo sapiens GN=RPL2 PE=1 SV=1 - [RLA2_HUMAN]	10.43	2
P37108	Signal recognition particle 14 kDa protein OS=Homo sapiens GN=SRP14 PE=1 SV=2 - [SRP14_HUMAN]	10.29	1
P61313	60S ribosomal protein L15 OS=Homo sapiens GN=RPL15 PE=1 SV=2 - [RL15_HUMAN]	10.29	3
P62258	14-3-3 protein epsilon OS=Homo sapiens GN=YWHAE PE=1 SV=1 - [1433E_HUMAN]	10.20	6
P62826	GTP-binding nuclear protein Ran OS=Homo sapiens GN=RAN PE=1 SV=3 - [RAN_HUMAN]	10.19	4
P32668	60S ribosomal protein L22 OS=Homo sapiens GN=RPL22 PE=1 SV=2 - [RL22_HUMAN]	10.16	2
P27695	DNA-(apurinic or apyrimidinic site) lyase OS=Homo sapiens GN=APEX1 PE=1 SV=2 - [APEX1_HUMAN]	10.06	4
B2RPK0	Putative high mobility group protein B1-like 1 OS=Homo sapiens GN=HMGB1P1 PE=5 SV=1 - [HGB1A_HUMAN]	9.95	4
F8WVW4	60S acidic ribosomal protein P0 (Fragment) OS=Homo sapiens GN=RPLP0 PE=1 SV=1 - [F8WVW4_HUMAN]	9.91	1
P32969	60S ribosomal protein L9 OS=Homo sapiens GN=RPL9 PE=1 SV=1 - [RL9_HUMAN]	9.90	1
P61513	60S ribosomal protein L37a OS=Homo sapiens GN=RPL37A PE=1 SV=2 - [RL37A_HUMAN]	9.78	2
P42677	40S ribosomal protein S27 OS=Homo sapiens GN=RPS27 PE=1 SV=3 - [RS27_HUMAN]	9.52	2
P47914	60S ribosomal protein L29 OS=Homo sapiens GN=RPL29 PE=1 SV=2 - [RL29_HUMAN]	9.43	2
P55769	NHP2-like protein 1 OS=Homo sapiens GN=NHP2L1 PE=1 SV=3 - [NHP2L1_HUMAN]	9.38	2
O60506	Heterogeneous nuclear ribonucleoprotein Q OS=Homo sapiens GN=SYNCRIP PE=1 SV=2 - [HNRPO_HUMAN]	9.31	10
Q5JTH9	RRP12-like protein OS=Homo sapiens GN=RRP12 PE=1 SV=2 - [RRP12_HUMAN]	9.18	20
P50990	T-complex protein 1 subunit theta OS=Homo sapiens GN=CCT8 PE=1 SV=4 - [TCQP_HUMAN]	9.12	8
Q14103	Heterogeneous nuclear ribonucleoprotein D0 OS=Homo sapiens GN=HNRND0 PE=1 SV=1 - [HNRPD_HUMAN]	9.01	3
P39748	Flap endonuclease 1 OS=Homo sapiens GN=FEN1 PE=1 SV=1 - [FEN1_HUMAN]	8.95	5
P43243	Matrin-3 OS=Homo sapiens GN=MATR3 PE=1 SV=2 - [MATR3_HUMAN]	8.85	13
Q16629	Serine/arginine-rich splicing factor 7 OS=Homo sapiens GN=SRSF7 PE=1 SV=1 - [SRSF7_HUMAN]	8.82	3
Q12905	Interleukin enhancer-binding factor 2 OS=Homo sapiens GN=ILF2 PE=1 SV=2 - [ILF2_HUMAN]	8.72	6
O15235	28S ribosomal protein S12, mitochondrial OS=Homo sapiens GN=MRPS12 PE=1 SV=1 - [RT12_HUMAN]	8.70	2
P56270	Myc-associated zinc finger protein OS=Homo sapiens GN=MAZ PE=1 SV=1 - [MAZ_HUMAN]	8.60	9
P62633	Cellular nucleic acid-binding protein OS=Homo sapiens GN=CNBP PE=1 SV=1 - [CNBP_HUMAN]	8.47	1
P15311	Ezrin OS=Homo sapiens GN=EZR PE=1 SV=4 - [EZRI_HUMAN]	8.36	8
P35908	Keratin, type II cytoskeletal 2 epidermal OS=Homo sapiens GN=KRT2 PE=1 SV=2 - [K22E_HUMAN]	8.29	10
F8WB88	Cullin-associated NEDD8-dissociated protein 2 OS=Homo sapiens GN=CAND2 PE=1 SV=1 - [F8WB88_HUMAN]	8.18	1
P50454	Serpin H1 OS=Homo sapiens GN=SERP1NH1 PE=1 SV=2 - [SERPH_HUMAN]	7.89	3
Q96RQ3	Methylcrotonoyl-CoA carboxylase subunit alpha, mitochondrial OS=Homo sapiens GN=MCCC1 PE=1 SV=3 - [MCCA_HUMAN]	7.86	10
Q9GZV4	Eukaryotic translation initiation factor 5A-2 OS=Homo sapiens GN=EIF5A2 PE=1 SV=3 - [IF5A2_HUMAN]	7.84	4
P14174	Macrophage migration inhibitory factor OS=Homo sapiens GN=MIF PE=1 SV=4 - [MIF_HUMAN]	7.83	1
P62854	40S ribosomal protein S26 OS=Homo sapiens GN=RPS26 PE=1 SV=3 - [RS26_HUMAN]	7.83	2
P50991	T-complex protein 1 subunit delta OS=Homo sapiens GN=CCT4 PE=1 SV=4 - [TCPD_HUMAN]	7.79	8
Q98Z64	Nucleolar GTP-binding protein 1 OS=Homo sapiens GN=GTPBP4 PE=1 SV=3 - [NOGI_HUMAN]	7.73	9
Q13148	TAR DNA-binding protein 43 OS=Homo sapiens GN=TARDBP PE=1 SV=1 - [TADBP_HUMAN]	7.73	6
P49207	60S ribosomal protein L34 OS=Homo sapiens GN=RPL34 PE=1 SV=3 - [RL34_HUMAN]	7.69	2
P18124	60S ribosomal protein L7 OS=Homo sapiens GN=RPL7 PE=1 SV=1 - [RL7_HUMAN]	7.66	2
P35637	RNA-binding protein FUS OS=Homo sapiens GN=FUS PE=1 SV=1 - [FUS_HUMAN]	7.60	4
D6R9P3	Heterogeneous nuclear ribonucleoprotein A/B OS=Homo sapiens GN=HNRNPAB PE=1 SV=1 - [D6R9P3_HUMAN]	7.50	2
P35659	Protein DEK OS=Homo sapiens GN=DEK PE=1 SV=1 - [DEK_HUMAN]	7.47	6
Q15691	Microtubule-associated protein RP/EB family member 1 OS=Homo sapiens GN=MAPRE1 PE=1 SV=3 - [MARE1_HUMAN]	7.46	1
P23284	Peptidyl-prolyl cis-trans isomerase B OS=Homo sapiens GN=PP1B PE=1 SV=2 - [PP1B_HUMAN]	7.41	4
Q9H5H4	Zinc finger protein 768 OS=Homo sapiens GN=ZNF768 PE=1 SV=2 - [ZN768_HUMAN]	7.41	7
P07477	Trypsin-1 OS=Homo sapiens GN=PRSS1 PE=1 SV=1 - [TRY1_HUMAN]	7.29	9
P62263	40S ribosomal protein S14 OS=Homo sapiens GN=RPS14 PE=1 SV=3 - [RS14_HUMAN]	7.28	1
P62899	60S ribosomal protein L31 OS=Homo sapiens GN=RPL31 PE=1 SV=1 - [RL31_HUMAN]	7.20	2
P52292	Importin subunit alpha-1 OS=Homo sapiens GN=KPN2 PE=1 SV=1 - [IMA1_HUMAN]	7.18	3
O00148	ATP-dependent RNA helicase DD39A OS=Homo sapiens GN=DD39A PE=1 SV=2 - [DX39A_HUMAN]	7.03	7
P13645	Keratin, type I cytoskeletal 10 OS=Homo sapiens GN=KRT10 PE=1 SV=6 - [K1C10_HUMAN]	7.02	10
Q08211	ATP-dependent RNA helicase A OS=Homo sapiens GN=DHX9 PE=1 SV=4 - [DHX9_HUMAN]	6.85	16
A6NMY6	Putative annexin A2-like protein OS=Homo sapiens GN=ANXA2P2 PE=5 SV=2 - [AXA2L_HUMAN]	6.78	2
Q12906	Interleukin enhancer-binding factor 3 OS=Homo sapiens GN=ILF3 PE=1 SV=3 - [ILF3_HUMAN]	6.71	10
P49006	MARCKS-related protein OS=Homo sapiens GN=MARCKSL1 PE=1 SV=2 - [MRP_HUMAN]	6.67	1
P61353	60S ribosomal protein L27 OS=Homo sapiens GN=RPL27 PE=1 SV=2 - [RL27_HUMAN]	6.62	2
O76021	Ribosomal L1 domain-containing protein 1 OS=Homo sapiens GN=RSL1D1 PE=1 SV=3 - [RL1D1_HUMAN]	6.53	7
P52597	Heterogeneous nuclear ribonucleoprotein F OS=Homo sapiens GN=HNRNPF PE=1 SV=3 - [HNRPF_HUMAN]	6.27	4
P14678	Small nuclear ribonucleoprotein-associated proteins B and B' OS=Homo sapiens GN=SNRBP PE=1 SV=2 - [RSMB_HUMAN]	6.25	2
Q15717	ELAV-like protein 1 OS=Homo sapiens GN=ELAVL1 PE=1 SV=2 - [ELAVL1_HUMAN]	6.13	6
Q57280	Putative methyltransferase C9orf114 OS=Homo sapiens GN=C9orf114 PE=1 SV=3 - [C1114_HUMAN]	6.12	4
P63244	Guanine nucleotide-binding protein subunit beta-2-like 1 OS=Homo sapiens GN=GNB2L1 PE=1 SV=3 - [GBLP_HUMAN]	5.99	3
P18754	Regulator of chromosome condensation OS=Homo sapiens GN=RCC1 PE=1 SV=1 - [RCC1_HUMAN]	5.94	2
Q9UKM9	RNA-binding protein Raly OS=Homo sapiens GN=RALY PE=1 SV=1 - [RALY_HUMAN]	5.88	4
P06576	ATP synthase subunit beta, mitochondrial OS=Homo sapiens GN=ATP5B PE=1 SV=3 - [ATPB_HUMAN]	5.86	4

Accession	Description	ΣCoverage	Σ# PSMs
P25205	DNA replication licensing factor MCM3 OS=Homo sapiens GN=MCM3 PE=1 SV=3 - [MCM3_HUMAN]	5.69	6
P38919	Eukaryotic initiation factor 4A-III OS=Homo sapiens GN=EIF4A3 PE=1 SV=4 - [EIF4A3_HUMAN]	5.60	4
Q9Y324	rRNA-processing protein FCF1 homolog OS=Homo sapiens GN=FCF1 PE=2 SV=1 - [FCF1_HUMAN]	5.56	1
P31942	Heterogeneous nuclear ribonucleoprotein H3 OS=Homo sapiens GN=HNRNPH3 PE=1 SV=2 - [HNRH3_HUMAN]	5.49	5
Q96AE4	Far upstream element-binding protein 1 OS=Homo sapiens GN=FUBP1 PE=1 SV=3 - [FUBP1_HUMAN]	5.43	9
Q9PQ00	Myb-binding protein 1A OS=Homo sapiens GN=MYBBP1A PE=1 SV=2 - [MBB1A_HUMAN]	5.42	12
Q96EK4	THAP domain-containing protein 11 OS=Homo sapiens GN=THAP11 PE=1 SV=2 - [THA11_HUMAN]	5.41	2
Q9Y4X4	Kruppel-like factor 12 OS=Homo sapiens GN=KLF12 PE=1 SV=2 - [KLF12_HUMAN]	5.22	2
P39023	60S ribosomal protein L3 OS=Homo sapiens GN=RPL3 PE=1 SV=2 - [RL3_HUMAN]	5.21	4
P07305	Histone H1.0 OS=Homo sapiens GN=H1F0 PE=1 SV=3 - [H10_HUMAN]	5.15	1
P20700	Lamin-B1 OS=Homo sapiens GN=LMBN1 PE=1 SV=2 - [LMBN1_HUMAN]	5.12	6
P48730	Casein kinase I isoform delta OS=Homo sapiens GN=CSNK1D PE=1 SV=2 - [KC1D_HUMAN]	5.06	3
O00567	Nucleolar protein 56 OS=Homo sapiens GN=NOP56 PE=1 SV=4 - [NOP56_HUMAN]	5.05	6
P17987	T-complex protein 1 subunit alpha OS=Homo sapiens GN=TCP1 PE=1 SV=1 - [TCPA_HUMAN]	4.86	3
Q6P087	RNA pseudouridylylase synthase domain-containing protein 3 OS=Homo sapiens GN=RPUSD3 PE=1 SV=3 - [RUSD3_HUMAN]	4.84	4
P12277	Creatine kinase B-type OS=Homo sapiens GN=CKB PE=1 SV=1 - [KCRB_HUMAN]	4.72	4
P08621	U1 small nuclear ribonucleoprotein 70 kDa OS=Homo sapiens GN=SNRNP70 PE=1 SV=2 - [RU17_HUMAN]	4.58	4
K7E1I0	Scaffold attachment factor B1 (Fragment) OS=Homo sapiens GN=SAFB PE=1 SV=1 - [K7E1I0_HUMAN]	4.58	1
P36578	60S ribosomal protein L4 OS=Homo sapiens GN=RPL4 PE=1 SV=5 - [RL4_HUMAN]	4.45	4
P61204	ADP-ribosylation factor 3 OS=Homo sapiens GN=ARF3 PE=1 SV=2 - [ARF3_HUMAN]	4.42	1
Q8IWS0	PHD finger protein 6 OS=Homo sapiens GN=PHF6 PE=1 SV=1 - [PHF6_HUMAN]	4.38	2
A6N1Z1	Ras-related protein Rap-1b-like protein OS=Homo sapiens PE=2 SV=1 - [RP1BL_HUMAN]	4.35	1
Q9Y2R4	Probable ATP-dependent RNA helicase DDX52 OS=Homo sapiens GN=DDX52 PE=1 SV=3 - [DDX52_HUMAN]	4.34	4
Q99832	T-complex protein 1 subunit eta OS=Homo sapiens GN=CCT7 PE=1 SV=2 - [TCPH_HUMAN]	4.24	3
Q51440	Putative transferase CAF17, mitochondrial OS=Homo sapiens GN=IBAS7 PE=1 SV=1 - [CAF17_HUMAN]	4.21	2
Q00325	Phosphate carrier protein, mitochondrial OS=Homo sapiens GN=SLC25A3 PE=1 SV=2 - [MPCP_HUMAN]	4.14	5
P38646	Stress-70 protein, mitochondrial OS=Homo sapiens GN=HSPA9 PE=1 SV=2 - [GRP75_HUMAN]	4.12	5
P35232	Prohibitin OS=Homo sapiens GN=PHB PE=1 SV=1 - [PHB_HUMAN]	4.04	1
P78527	DNA-dependent protein kinase catalytic subunit OS=Homo sapiens GN=PRKDC PE=1 SV=3 - [PRKDC_HUMAN]	4.00	32
P46109	Crk-like protein OS=Homo sapiens GN=CRKL PE=1 SV=1 - [CRKL_HUMAN]	3.96	1
P15259	Phosphoglycerate mutase 2 OS=Homo sapiens GN=PGAM2 PE=1 SV=3 - [PGAM2_HUMAN]	3.95	1
O75475	PC4 and SFRS1-interacting protein OS=Homo sapiens GN=PSIP1 PE=1 SV=1 - [PSIP1_HUMAN]	3.77	3
Q13642	Four and a half LIM domains protein 1 OS=Homo sapiens GN=FHL1 PE=1 SV=4 - [FHL1_HUMAN]	3.72	2
P62906	60S ribosomal protein L10a OS=Homo sapiens GN=RPL10A PE=1 SV=2 - [RL10A_HUMAN]	3.69	1
P55265	Double-stranded RNA-specific adenosine deaminase OS=Homo sapiens GN=ADAR PE=1 SV=4 - [DSRAD_HUMAN]	3.67	10
P33993	DNA replication licensing factor MCM7 OS=Homo sapiens GN=MCM7 PE=1 SV=4 - [MCM7_HUMAN]	3.62	3
P01891	HLA class I histocompatibility antigen, A-68 alpha chain OS=Homo sapiens GN=HLA-A PE=1 SV=4 - [IA68_HUMAN]	3.56	1
P78371	T-complex protein 1 subunit beta OS=Homo sapiens GN=CCT2 PE=1 SV=4 - [TCPB_HUMAN]	3.55	3
O75533	Splicing factor 3B subunit 1 OS=Homo sapiens GN=SF3B1 PE=1 SV=3 - [SF3B1_HUMAN]	3.53	5
P21796	Voltage-dependent anion-selective channel protein 1 OS=Homo sapiens GN=VDAC1 PE=1 SV=2 - [VDAC1_HUMAN]	3.53	2
P49368	T-complex protein 1 subunit gamma OS=Homo sapiens GN=CCT3 PE=1 SV=4 - [TCPG_HUMAN]	3.49	4
P00918	Carbonic anhydrase 2 OS=Homo sapiens GN=CA2 PE=1 SV=2 - [CAH2_HUMAN]	3.46	1
HOY4T4	Centrosomal protein of 170 kDa (Fragment) OS=Homo sapiens GN=CEP170 PE=1 SV=1 - [HOY4T4_HUMAN]	3.42	1
P33992	DNA replication licensing factor MCM5 OS=Homo sapiens GN=MCM5 PE=1 SV=5 - [MCM5_HUMAN]	3.41	6
P21333	Filamin-A OS=Homo sapiens GN=FLNA PE=1 SV=4 - [FLNA_HUMAN]	3.36	13
Q01081	Splicing factor U2AF 35 kDa subunit OS=Homo sapiens GN=U2AF1 PE=1 SV=3 - [U2AF1_HUMAN]	3.33	1
P18846	Cyclic AMP-dependent transcription factor ATF-1 OS=Homo sapiens GN=ATF1 PE=1 SV=2 - [ATF1_HUMAN]	3.32	1
A8M236	Enoplakin-like protein OS=Homo sapiens GN=EVPL1 PE=2 SV=1 - [EVPL1_HUMAN]	3.32	1
Q14978	Nucleolar and coiled-body phosphoprotein 1 OS=Homo sapiens GN=NOLC1 PE=1 SV=2 - [NOLC1_HUMAN]	3.29	2
Q9Y5J1	U3 small nucleolar RNA-associated protein 18 homolog OS=Homo sapiens GN=UTP18 PE=1 SV=3 - [UTP18_HUMAN]	3.24	1
Q15084	Protein disulfide-isomerase A6 OS=Homo sapiens GN=PDI6 PE=1 SV=1 - [PDI6_HUMAN]	3.18	2
Q07666	KH domain-containing, RNA-binding, signal transduction-associated protein 1 OS=Homo sapiens GN=KHDRBS1 PE=1 SV=1 - [KHDR1_HUMAN]	3.16	2
P78347	General transcription factor II-1 OS=Homo sapiens GN=GTF2I PE=1 SV=2 - [GTF2I_HUMAN]	3.11	5
Q9P016	Thymocyte nuclear protein 1 OS=Homo sapiens GN=THYN1 PE=1 SV=1 - [THYN1_HUMAN]	3.11	2
Q13263	Transcription intermediary factor 1-beta OS=Homo sapiens GN=TRIM28 PE=1 SV=5 - [TIF1B_HUMAN]	3.11	4
P36873	Serine/threonine-protein phosphatase PP1-gamma catalytic subunit OS=Homo sapiens GN=PPP1CC PE=1 SV=1 - [PP1G_HUMAN]	3.10	2
Q8I2J6	Inactive L-threonine 3-dehydrogenase, mitochondrial OS=Homo sapiens GN=TDH PE=2 SV=1 - [TDH_HUMAN]	3.04	1
Q7L190	Developmental pluripotency-associated protein 4 OS=Homo sapiens GN=DPPA4 PE=1 SV=2 - [DPPA4_HUMAN]	2.96	1
Q15061	WD repeat-containing protein 43 OS=Homo sapiens GN=WDR43 PE=1 SV=3 - [WDR43_HUMAN]	2.95	4
O43615	Mitochondrial import inner membrane translocase subunit TIM44 OS=Homo sapiens GN=TIMM44 PE=1 SV=2 - [TIM44_HUMAN]	2.88	1
Q5T6C4	Ataxin-7-like protein 2 OS=Homo sapiens GN=ATXN7L2 PE=4 SV=1 - [Q5T6C4_HUMAN]	2.87	1
P46087	Probable 28S rRNA (cytosine(4447)-C(5))-methyltransferase OS=Homo sapiens GN=NOP2 PE=1 SV=2 - [NOP2_HUMAN]	2.83	4
P12004	Proliferating cell nuclear antigen OS=Homo sapiens GN=PCNA PE=1 SV=1 - [PCNA_HUMAN]	2.68	1
Q8G124	Transcriptional regulator Kaiso OS=Homo sapiens GN=ZBTB33 PE=1 SV=2 - [KAISO_HUMAN]	2.53	2
Q14684	Ribosomal RNA processing protein 1 homolog B OS=Homo sapiens GN=RRP1B PE=1 SV=3 - [RRP1B_HUMAN]	2.51	2
Q9H054	Probable ATP-dependent RNA helicase DDX47 OS=Homo sapiens GN=DDX47 PE=1 SV=1 - [DDX47_HUMAN]	2.42	2

Accession	Description	ΣCoverage	Σ# PSMs
Q9NVP1	ATP-dependent RNA helicase DDX18 OS=Homo sapiens GN=DDX18 PE=1 SV=2 - [DDX18_HUMAN]	2.39	4
Q9Y2X3	Nucleolar protein 58 OS=Homo sapiens GN=NOPS8 PE=1 SV=1 - [NOPS8_HUMAN]	2.27	2
P15170	Eukaryotic peptide chain release factor GTP-binding subunit ERF3A OS=Homo sapiens GN=GSP11 PE=1 SV=1 - [ERF3A_HUMAN]	2.20	1
P11908	Ribose-phosphate pyrophosphokinase 2 OS=Homo sapiens GN=PRPS2 PE=1 SV=2 - [PRPS2_HUMAN]	2.20	1
Q9NQ29	Putative RNA-binding protein Luc7-like 1 OS=Homo sapiens GN=LUC7L PE=1 SV=1 - [LUC7L_HUMAN]	2.16	1
P12532	Creatine kinase U-type, mitochondrial OS=Homo sapiens GN=CKMT1A PE=1 SV=1 - [KCRU_HUMAN]	2.16	1
Q14204	Cytoplasmic dynein 1 heavy chain 1 OS=Homo sapiens GN=DYNC1H1 PE=1 SV=5 - [DYHC1_HUMAN]	2.13	16
P26368	Splicing factor UZAF 65 kDa subunit OS=Homo sapiens GN=UZAF2 PE=1 SV=4 - [UZAF2_HUMAN]	2.11	2
Q13247	Serine/arginine-rich splicing factor 6 OS=Homo sapiens GN=SRSF6 PE=1 SV=2 - [SRSF6_HUMAN]	2.03	1
O60832	H/ACA ribonucleoprotein complex subunit 4 OS=Homo sapiens GN=DKC1 PE=1 SV=3 - [DKC1_HUMAN]	1.95	2
P13010	X-ray repair cross-complementing protein 5 OS=Homo sapiens GN=XRCC5 PE=1 SV=3 - [XRCC5_HUMAN]	1.91	1
O14979	Heterogeneous nuclear ribonucleoprotein D-like OS=Homo sapiens GN=HNRNPD1 PE=1 SV=3 - [HNRDL_HUMAN]	1.90	1
Q14498	RNA-binding protein 39 OS=Homo sapiens GN=RBM39 PE=1 SV=2 - [RBM39_HUMAN]	1.89	1
P61619	Protein transport protein Sec61 subunit alpha isoform 1 OS=Homo sapiens GN=SEC61A1 PE=1 SV=2 - [S61A1_HUMAN]	1.89	1
P49959	Double-strand break repair protein MRE11A OS=Homo sapiens GN=MRE11A PE=1 SV=3 - [MRE11_HUMAN]	1.84	2
P35520	Cystathionine beta-synthase OS=Homo sapiens GN=CBS PE=1 SV=2 - [CBS_HUMAN]	1.81	1
Q12874	Splicing factor 3A subunit 3 OS=Homo sapiens GN=SF3A3 PE=1 SV=1 - [SF3A3_HUMAN]	1.80	1
Q8WVY3	U4/U6 small nuclear ribonucleoprotein Prp31 OS=Homo sapiens GN=PRPF31 PE=1 SV=2 - [PRP31_HUMAN]	1.60	1
Q9UQR0	Sex comb on midleg-like protein 2 OS=Homo sapiens GN=SCML2 PE=1 SV=1 - [SCML2_HUMAN]	1.57	1
Q8WYQ5	Microprocessor complex subunit DGCR8 OS=Homo sapiens GN=DGCR8 PE=1 SV=1 - [DGCR8_HUMAN]	1.55	1
Q92945	Far upstream element-binding protein 2 OS=Homo sapiens GN=KHSRP PE=1 SV=4 - [FUBP2_HUMAN]	1.55	1
Q9BZ25	Apoptosis inhibitor 5 OS=Homo sapiens GN=API5 PE=1 SV=3 - [API5_HUMAN]	1.53	1
P35579	Myosin-9 OS=Homo sapiens GN=MYH9 PE=1 SV=4 - [MYH9_HUMAN]	1.53	2
O43143	Putative pre-mRNA-splicing factor ATP-dependent RNA helicase DHX15 OS=Homo sapiens GN=DHX15 PE=1 SV=2 - [DHX15_HUMAN]	1.51	1
Q96T88	E3 ubiquitin-protein ligase UHRF1 OS=Homo sapiens GN=UHRF1 PE=1 SV=1 - [UHRF1_HUMAN]	1.51	1
O94776	Metastasis-associated protein MTA2 OS=Homo sapiens GN=MTA2 PE=1 SV=1 - [MTA2_HUMAN]	1.50	1
P12956	X-ray repair cross-complementing protein 6 OS=Homo sapiens GN=XRCC6 PE=1 SV=2 - [XRCC6_HUMAN]	1.48	2
Q9UM54	Pre-mRNA-processing factor 19 OS=Homo sapiens GN=PRPF19 PE=1 SV=1 - [PRP19_HUMAN]	1.39	1
P27816	Microtubule-associated protein 4 OS=Homo sapiens GN=MAP4 PE=1 SV=3 - [MAP4_HUMAN]	1.30	1
Q9H0M6	5'-3' exoribonuclease 2 OS=Homo sapiens GN=XRN2 PE=1 SV=1 - [XRN2_HUMAN]	1.26	2
P49711	Transcriptional repressor CTCF OS=Homo sapiens GN=CTCF PE=1 SV=1 - [CTCF_HUMAN]	1.24	2
Q14839	Chromodomain-helicase-DNA-binding protein 4 OS=Homo sapiens GN=CHD4 PE=1 SV=2 - [CHD4_HUMAN]	1.20	4
P34246	DNA mismatch repair protein Msh2 OS=Homo sapiens GN=MSH2 PE=1 SV=1 - [MSH2_HUMAN]	1.18	2
Q02768	Serum albumin OS=Homo sapiens GN=ALB PE=1 SV=2 - [ALBU_HUMAN]	1.15	2
E9PDR5	Inhibitor of Bruton tyrosine kinase OS=Homo sapiens GN=IBTK PE=1 SV=1 - [E9PDR5_HUMAN]	1.13	1
J3KNJ3	N-acetylated-alpha-linked acidic dipeptidase 2 OS=Homo sapiens GN=NAALAD2 PE=4 SV=1 - [J3KNJ3_HUMAN]	1.13	1
Q14980	Nuclear mitotic apparatus protein 1 OS=Homo sapiens GN=NUMA1 PE=1 SV=2 - [NUMA1_HUMAN]	1.13	2
Q8N1F7	Nuclear pore complex protein Nup93 OS=Homo sapiens GN=NUP93 PE=1 SV=2 - [NUP93_HUMAN]	1.10	1
P40939	Trifunctional enzyme subunit alpha, mitochondrial OS=Homo sapiens GN=HADHA PE=1 SV=2 - [ECHA_HUMAN]	1.05	1
P54886	Delta-1-pyrroline-5-carboxylate synthase OS=Homo sapiens GN=ALDH18A1 PE=1 SV=2 - [P5CS_HUMAN]	1.01	2
P05129	Protein kinase C gamma type OS=Homo sapiens GN=PRKCG PE=1 SV=3 - [KPCG_HUMAN]	1.00	1
Q15393	Splicing factor 3B subunit 3 OS=Homo sapiens GN=SF3B3 PE=1 SV=4 - [SF3B3_HUMAN]	0.99	2
Q14566	DNA replication licensing factor MCM6 OS=Homo sapiens GN=MCM6 PE=1 SV=1 - [MCM6_HUMAN]	0.97	1
Q1KMD3	Heterogeneous nuclear ribonucleoprotein U-like protein 2 OS=Homo sapiens GN=HNRNPUL2 PE=1 SV=1 - [HNRUL2_HUMAN]	0.94	1
P58215	Lysyl oxidase homolog 3 OS=Homo sapiens GN=LOXL3 PE=2 SV=1 - [LOXL3_HUMAN]	0.93	2
P55198	Protein AF-17 OS=Homo sapiens GN=MLL7 PE=1 SV=2 - [AF17_HUMAN]	0.91	1
Q8NEM7	Transcription factor SPT20 homolog OS=Homo sapiens GN=SUPT20H PE=1 SV=2 - [SP20H_HUMAN]	0.90	1
Q15477	Helicase SKI2W OS=Homo sapiens GN=SKI2L PE=1 SV=3 - [SKI2L_HUMAN]	0.80	1
Q00858	Epithelial cell-transforming sequence 2 oncogene-like OS=Homo sapiens GN=ECT2L PE=2 SV=2 - [ECT2L_HUMAN]	0.77	1
Q92878	DNA repair protein RAD50 OS=Homo sapiens GN=RAD50 PE=1 SV=1 - [RAD50_HUMAN]	0.76	2
O60264	SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A member 5 OS=Homo sapiens GN=SMARCA5 PE=1 SV=1 - [SMCA5_HUMAN]	0.76	1
Q92621	Nuclear pore complex protein Nup205 OS=Homo sapiens GN=NUP205 PE=1 SV=3 - [NU205_HUMAN]	0.75	2
P49792	E3 SUMO-protein ligase RanBP2 OS=Homo sapiens GN=RANBP2 PE=1 SV=2 - [RBP2_HUMAN]	0.71	4
O60241	Brain-specific angiogenesis inhibitor 2 OS=Homo sapiens GN=BAI2 PE=1 SV=2 - [BAI2_HUMAN]	0.69	2
P11388	DNA topoisomerase 2-alpha OS=Homo sapiens GN=TOP2A PE=1 SV=3 - [TOP2A_HUMAN]	0.65	2
O00411	DNA-directed RNA polymerase, mitochondrial OS=Homo sapiens GN=POLRMT PE=1 SV=2 - [RPOM_HUMAN]	0.65	1
Q8WXE0	Caskin-2 OS=Homo sapiens GN=CASKIN2 PE=1 SV=2 - [CSK12_HUMAN]	0.58	1
Q15149	Plectin OS=Homo sapiens GN=PLEC PE=1 SV=3 - [PLEC_HUMAN]	0.38	3
P12270	Nucleoprotein TPR OS=Homo sapiens GN=TPR PE=1 SV=3 - [TPR_HUMAN]	0.38	1
Q5VYK3	Proteasome-associated protein ECM29 homolog OS=Homo sapiens GN=ECM29 PE=1 SV=2 - [ECM29_HUMAN]	0.38	1
P02111	Collagen alpha-3(VI) chain OS=Homo sapiens GN=COL6A3 PE=1 SV=5 - [CO6A3_HUMAN]	0.22	1

**Table 20. Raw data of BioID experiment: BioID-DOT1L-mut infected HEK293T cells, 2<sup>nd</sup> reading**

Accession	Description	ΣCoverage	Σ# PSMs
P22626	Heterogeneous nuclear ribonucleoproteins A2/B1 OS=Homo sapiens GN=HNRNPA2B1 PE=1 SV=2 - [ROA2_HUMAN]	56.09	57
P62937	Peptidyl-prolyl cis-trans isomerase A OS=Homo sapiens GN=PPIA PE=1 SV=2 - [PPIA_HUMAN]	47.88	17
P62888	60S ribosomal protein L30 OS=Homo sapiens GN=RPL30 PE=1 SV=2 - [RL30_HUMAN]	46.96	9
P46781	40S ribosomal protein S9 OS=Homo sapiens GN=RPS9 PE=1 SV=3 - [RS9_HUMAN]	46.91	41
P46779	60S ribosomal protein L28 OS=Homo sapiens GN=RPL28 PE=1 SV=3 - [RL28_HUMAN]	43.07	17
P15880	40S ribosomal protein S2 OS=Homo sapiens GN=RPS2 PE=1 SV=2 - [RS2_HUMAN]	41.64	26
P23396	40S ribosomal protein S3 OS=Homo sapiens GN=RPS3 PE=1 SV=2 - [RS3_HUMAN]	41.15	23
P62280	40S ribosomal protein S11 OS=Homo sapiens GN=RPS11 PE=1 SV=3 - [RS11_HUMAN]	41.14	22
P49458	Signal recognition particle 9 kDa protein OS=Homo sapiens GN=SRP9 PE=1 SV=2 - [SRP9_HUMAN]	40.70	7
P62249	40S ribosomal protein S16 OS=Homo sapiens GN=RPS16 PE=1 SV=2 - [RS16_HUMAN]	40.41	18
P18621	60S ribosomal protein L17 OS=Homo sapiens GN=RPL17 PE=1 SV=3 - [RL17_HUMAN]	40.22	18
P04406	Glyceraldehyde-3-phosphate dehydrogenase OS=Homo sapiens GN=GAPDH PE=1 SV=3 - [G3P_HUMAN]	40.00	27
P62805	Histone H4 OS=Homo sapiens GN=HIST1H4A PE=1 SV=2 - [H4_HUMAN]	39.81	13
O60814	Histone H2B type 1-K OS=Homo sapiens GN=HIST1H2BK PE=1 SV=3 - [H2BK_HUMAN]	38.89	15
P62913	60S ribosomal protein L11 OS=Homo sapiens GN=RPL11 PE=1 SV=2 - [RL11_HUMAN]	38.20	26
P61978	Heterogeneous nuclear ribonucleoprotein K OS=Homo sapiens GN=HNRNPK PE=1 SV=1 - [HNRPK_HUMAN]	37.15	34
P22392	Nucleoside diphosphate kinase B OS=Homo sapiens GN=NME2 PE=1 SV=1 - [NDKB_HUMAN]	36.84	10
P09847	Poly [ADP-ribose] polymerase 1 OS=Homo sapiens GN=PARP1 PE=1 SV=4 - [PARP1_HUMAN]	36.79	70
P61247	40S ribosomal protein S3a OS=Homo sapiens GN=RPS3A PE=1 SV=2 - [RS3A_HUMAN]	36.36	21
P68431	Histone H3.1 OS=Homo sapiens GN=HIST1H3A PE=1 SV=2 - [H31_HUMAN]	36.03	10
P09651	Heterogeneous nuclear ribonucleoprotein A1 OS=Homo sapiens GN=HNRNPA1 PE=1 SV=5 - [ROA1_HUMAN]	36.02	34
P08670	Vimentin OS=Homo sapiens GN=VIM PE=1 SV=4 - [VIME_HUMAN]	35.84	40
P62701	40S ribosomal protein S4, X isoform OS=Homo sapiens GN=RPS4X PE=1 SV=2 - [RS4X_HUMAN]	34.60	18
P26373	60S ribosomal protein L13 OS=Homo sapiens GN=RPL13 PE=1 SV=4 - [RL13_HUMAN]	34.60	29
P60709	Actin, cytoplasmic 1 OS=Homo sapiens GN=ACTB PE=1 SV=1 - [ACTB_HUMAN]	33.33	24
Q8N257	Histone H2B type 3-B OS=Homo sapiens GN=HIST3H2BB PE=1 SV=3 - [H2B3B_HUMAN]	33.33	13
P62857	40S ribosomal protein S28 OS=Homo sapiens GN=RPS28 PE=1 SV=1 - [RS28_HUMAN]	33.33	4
Q96K65	Histone H2A type 1-H OS=Homo sapiens GN=HIST1H2AH PE=1 SV=3 - [H2A1H_HUMAN]	32.81	9
P05141	ADP/ATP translocase 2 OS=Homo sapiens GN=SLC25A5 PE=1 SV=7 - [ADT2_HUMAN]	32.55	26
P07910	Heterogeneous nuclear ribonucleoproteins C1/C2 OS=Homo sapiens GN=HNRNPC PE=1 SV=4 - [HNRNPC_HUMAN]	31.70	21
Q9BQE3	Tubulin alpha-1C chain OS=Homo sapiens GN=TUBA1C PE=1 SV=1 - [TBA1C_HUMAN]	31.40	30
P68363	Tubulin alpha-1B chain OS=Homo sapiens GN=TUBA1B PE=1 SV=1 - [TBA1B_HUMAN]	31.26	30
P84103	Serine/arginine-rich splicing factor 3 OS=Homo sapiens GN=SRSF3 PE=1 SV=1 - [SRSF3_HUMAN]	31.10	9
P62753	40S ribosomal protein S6 OS=Homo sapiens GN=RPS6 PE=1 SV=1 - [RS6_HUMAN]	30.52	19
Q9Y3C1	Nucleolar protein 16 OS=Homo sapiens GN=NOP16 PE=1 SV=2 - [NOP16_HUMAN]	30.34	12
P62241	40S ribosomal protein S8 OS=Homo sapiens GN=RPS8 PE=1 SV=2 - [RS8_HUMAN]	29.81	10
P60866	40S ribosomal protein S20 OS=Homo sapiens GN=RPS20 PE=1 SV=1 - [RS20_HUMAN]	29.41	13
P51991	Heterogeneous nuclear ribonucleoprotein A3 OS=Homo sapiens GN=HNRNPA3 PE=1 SV=2 - [ROA3_HUMAN]	29.37	21
P05204	Non-histone chromosomal protein HMG-17 OS=Homo sapiens GN=HMGN2 PE=1 SV=3 - [HMGN2_HUMAN]	28.89	4
P04264	Keratin, type II cytoskeletal 1 OS=Homo sapiens GN=KRT1 PE=1 SV=6 - [K2C1_HUMAN]	28.73	39
Q15366-7	Isoform 7 of Poly(rC)-binding protein 2 OS=Homo sapiens GN=PCBP2 - [PCBP2_HUMAN]	28.62	7
P07437	Tubulin beta chain OS=Homo sapiens GN=TUBB PE=1 SV=2 - [TBB5_HUMAN]	28.60	30
P62847	40S ribosomal protein S24 OS=Homo sapiens GN=RPS24 PE=1 SV=1 - [RS24_HUMAN]	28.57	7
P05165	Propionyl-CoA carboxylase alpha chain, mitochondrial OS=Homo sapiens GN=PCCA PE=1 SV=4 - [PCCA_HUMAN]	28.57	36
P46782	40S ribosomal protein S5 OS=Homo sapiens GN=RPS5 PE=1 SV=4 - [RS5_HUMAN]	27.94	11
Q13885	Tubulin beta-2A chain OS=Homo sapiens GN=TUBB2A PE=1 SV=1 - [TBB2A_HUMAN]	27.87	30
P29372	DNA-3-methyladenine glycosylase OS=Homo sapiens GN=MPG PE=1 SV=3 - [3MG_HUMAN]	27.85	12
P63241	Eukaryotic translation initiation factor 5A-1 OS=Homo sapiens GN=EIF5A PE=1 SV=2 - [IF5A1_HUMAN]	27.27	6
P26641	Elongation factor 1-gamma OS=Homo sapiens GN=EEF1G PE=1 SV=3 - [EF1G_HUMAN]	27.23	26
Q9P258	Protein RCC2 OS=Homo sapiens GN=RCC2 PE=1 SV=2 - [RCC2_HUMAN]	27.20	26
Q06830	Peroxisomal protein OS=Homo sapiens GN=PRDX1 PE=1 SV=1 - [PRDX1_HUMAN]	27.14	13
P63267	Actin, gamma-enteric smooth muscle OS=Homo sapiens GN=ACTG2 PE=1 SV=1 - [ACTH_HUMAN]	27.13	22
P39019	40S ribosomal protein S19 OS=Homo sapiens GN=RPS19 PE=1 SV=2 - [RS19_HUMAN]	26.90	8
P42766	60S ribosomal protein L35 OS=Homo sapiens GN=RPL35 PE=1 SV=2 - [RL35_HUMAN]	26.83	9
Q8TEK3	Histone-lysine N-methyltransferase, H3 lysine-79 specific OS=Homo sapiens GN=DOT1L PE=1 SV=2 - [DOT1L_HUMAN]	26.74	101
P84098	60S ribosomal protein L19 OS=Homo sapiens GN=RPL19 PE=1 SV=1 - [RL19_HUMAN]	26.53	12
P52272	Heterogeneous nuclear ribonucleoprotein M OS=Homo sapiens GN=HNRNPM PE=1 SV=3 - [HNRPM_HUMAN]	26.44	37
P68104	Elongation factor 1-alpha OS=Homo sapiens GN=EEF1A1 PE=1 SV=1 - [EF1A1_HUMAN]	26.19	35

## Appendix-II

Accession	Description	ΣCoverage	Σ# PSMs
Q15233	Non-POU domain-containing octamer-binding protein OS=Homo sapiens GN=NONO PE=1 SV=4 - [NONO_HUMAN]	26.11	30
Q98VP2	Guanine nucleotide-binding protein-like 3 OS=Homo sapiens GN=GNL3 PE=1 SV=2 - [GNL3_HUMAN]	25.87	25
P17844	Probable ATP-dependent RNA helicase DDX5 OS=Homo sapiens GN=DDX5 PE=1 SV=1 - [DDX5_HUMAN]	25.73	41
P08107	Heat shock 70 kDa protein 1A/1B OS=Homo sapiens GN=HSPA1A PE=1 SV=5 - [HSP71_HUMAN]	25.43	33
P62633	Cellular nucleic acid-binding protein OS=Homo sapiens GN=CNBP PE=1 SV=1 - [CNBP_HUMAN]	25.42	6
P27635	60S ribosomal protein L10 OS=Homo sapiens GN=RPL10 PE=1 SV=4 - [RL10_HUMAN]	25.23	14
P62910	60S ribosomal protein L32 OS=Homo sapiens GN=RPL32 PE=1 SV=2 - [RL32_HUMAN]	25.19	10
P83881	60S ribosomal protein L36a OS=Homo sapiens GN=RPL36A PE=1 SV=2 - [RL36A_HUMAN]	24.53	10
P63173	60S ribosomal protein L38 OS=Homo sapiens GN=RPL38 PE=1 SV=2 - [RL38_HUMAN]	24.29	3
P06733	Alpha-enolase OS=Homo sapiens GN=ENO1 PE=1 SV=2 - [ENOA_HUMAN]	24.19	23
P16403	Histone H1.2 OS=Homo sapiens GN=HIST1H1C PE=1 SV=2 - [H12_HUMAN]	23.94	14
P62244	40S ribosomal protein S15a OS=Homo sapiens GN=RPS15A PE=1 SV=2 - [RS15A_HUMAN]	23.85	6
P62318	Small nuclear ribonucleoprotein Sm D3 OS=Homo sapiens GN=SNRNP3 PE=1 SV=1 - [SMD3_HUMAN]	23.81	3
P04075	Fructose-bisphosphate aldolase A OS=Homo sapiens GN=ALDOA PE=1 SV=2 - [ALDOA_HUMAN]	23.63	18
P38159	RNA-binding motif protein, X chromosome OS=Homo sapiens GN=RBMX PE=1 SV=3 - [RBMX_HUMAN]	23.53	19
P11498	Pyruvate carboxylase, mitochondrial OS=Homo sapiens GN=PC PE=1 SV=2 - [PYC_HUMAN]	23.26	51
P05114	Non-histone chromosomal protein HMG-14 OS=Homo sapiens GN=HMGN1 PE=1 SV=3 - [HMGN1_HUMAN]	23.00	4
P42677	40S ribosomal protein S27 OS=Homo sapiens GN=RPS27 PE=1 SV=3 - [RS27_HUMAN]	22.62	2
P29692	Elongation factor 1-delta OS=Homo sapiens GN=EEF1D PE=1 SV=5 - [EF1D_HUMAN]	22.42	8
P62266	40S ribosomal protein S23 OS=Homo sapiens GN=RPS23 PE=1 SV=3 - [RS23_HUMAN]	22.38	11
Q13085	Acetyl-CoA carboxylase 1 OS=Homo sapiens GN=ACACA PE=1 SV=2 - [ACACA_HUMAN]	22.25	95
Q15365	Poly(rC)-binding protein 1 OS=Homo sapiens GN=PCBP1 PE=1 SV=2 - [PCBP1_HUMAN]	21.63	12
P62829	60S ribosomal protein L23 OS=Homo sapiens GN=RPL23 PE=1 SV=1 - [RL23_HUMAN]	21.43	5
P07195	L-lactate dehydrogenase B chain OS=Homo sapiens GN=LDHB PE=1 SV=2 - [LDHB_HUMAN]	21.26	11
P10809	60 kDa heat shock protein, mitochondrial OS=Homo sapiens GN=HSPD1 PE=1 SV=2 - [CHG0_HUMAN]	20.94	21
P62854	40S ribosomal protein S26 OS=Homo sapiens GN=RPS26 PE=1 SV=3 - [RS26_HUMAN]	20.87	3
O76021	Ribosomal L1 domain-containing protein 1 OS=Homo sapiens GN=RLD1 PE=1 SV=3 - [RLD1_HUMAN]	20.61	19
P25705	ATP synthase subunit alpha, mitochondrial OS=Homo sapiens GN=ATP5A1 PE=1 SV=1 - [ATPA_HUMAN]	20.61	21
Q16629	Serine/arginine-rich splicing factor 7 OS=Homo sapiens GN=SRSF7 PE=1 SV=1 - [SRSF7_HUMAN]	20.59	8
P23528	Cofilin-1 OS=Homo sapiens GN=CFL1 PE=1 SV=3 - [COF1_HUMAN]	20.48	5
P12236	ADP/ATP translocase 3 OS=Homo sapiens GN=SLC25A6 PE=1 SV=4 - [ADT3_HUMAN]	20.47	18
P0C055	Histone H2A.Z OS=Homo sapiens GN=H2AFZ PE=1 SV=2 - [H2AZ_HUMAN]	20.31	6
P62987	Ubiquitin-60S ribosomal protein L40 OS=Homo sapiens GN=UBA52 PE=1 SV=2 - [RL40_HUMAN]	20.31	6
Q96EY4	Translation machinery-associated protein 16 OS=Homo sapiens GN=TMA16 PE=1 SV=2 - [TMA16_HUMAN]	20.20	9
P35527	Keratin, type I cytoskeletal 9 OS=Homo sapiens GN=KRT9 PE=1 SV=3 - [K1C9_HUMAN]	20.06	22
P12277	Creatine kinase B-type OS=Homo sapiens GN=CKB PE=1 SV=1 - [KCRB_HUMAN]	19.95	11
P62891	60S ribosomal protein L39 OS=Homo sapiens GN=RPL39 PE=1 SV=2 - [RL39_HUMAN]	19.61	2
P46776	60S ribosomal protein L27a OS=Homo sapiens GN=RPL27A PE=1 SV=2 - [RL27A_HUMAN]	19.59	8
P27695	DNA-(apurinic or apyrimidinic site) lyase OS=Homo sapiens GN=APEX1 PE=1 SV=2 - [APEX1_HUMAN]	19.50	8
Q99623	Prohibitin-2 OS=Homo sapiens GN=PHB2 PE=1 SV=2 - [PHB2_HUMAN]	19.06	9
Q13151	Heterogeneous nuclear ribonucleoprotein A0 OS=Homo sapiens GN=HNRPA0 PE=1 SV=1 - [ROA0_HUMAN]	19.02	12
O00571	ATP-dependent RNA helicase DDX3X OS=Homo sapiens GN=DDX3X PE=1 SV=3 - [DDX3X_HUMAN]	18.58	26
P23246	Splicing factor, proline- and glutamine-rich OS=Homo sapiens GN=SFQ PE=1 SV=2 - [SFQ_HUMAN]	18.39	26
P11142	Heat shock cognate 71 kDa protein OS=Homo sapiens GN=HSPA8 PE=1 SV=1 - [HSP7C_HUMAN]	18.27	26
P22087	rRNA 2'-O-methyltransferase fibrillar OS=Homo sapiens GN=FBL PE=1 SV=2 - [FBRL_HUMAN]	18.07	11
Q02543	60S ribosomal protein L18a OS=Homo sapiens GN=RPL18A PE=1 SV=2 - [RL18A_HUMAN]	17.61	8
P06748	Nucleophosmin OS=Homo sapiens GN=NPM1 PE=1 SV=2 - [NPM_HUMAN]	17.35	11
P23969	60S ribosomal protein L9 OS=Homo sapiens GN=RPL9 PE=1 SV=1 - [RL9_HUMAN]	17.19	4
Q14103	Heterogeneous nuclear ribonucleoprotein D0 OS=Homo sapiens GN=HNRNP0 PE=1 SV=1 - [HNRPD_HUMAN]	17.18	13
P62269	40S ribosomal protein S18 OS=Homo sapiens GN=RPS18 PE=1 SV=3 - [RS18_HUMAN]	17.11	6
A8MW09	Putative small nuclear ribonucleoprotein G-like protein 15 OS=Homo sapiens GN=SNRPGP15 PE=5 SV=2 - [RUXGL_HUMAN]	17.11	2
P30050	60S ribosomal protein L12 OS=Homo sapiens GN=RPL12 PE=1 SV=1 - [RL12_HUMAN]	16.97	4
P30041	Peroxisome oxidin-6 OS=Homo sapiens GN=PRDX6 PE=1 SV=3 - [PRDX6_HUMAN]	16.96	2
P62861	40S ribosomal protein S30 OS=Homo sapiens GN=FAU PE=1 SV=1 - [RS30_HUMAN]	16.95	2
Q99729-3	Isoform 3 of Heterogeneous nuclear ribonucleoprotein A/B OS=Homo sapiens GN=HNRNPAB - [ROAA_HUMAN]	16.49	4
P14618	Pyruvate kinase PKM OS=Homo sapiens GN=PKM PE=1 SV=4 - [KPYM_HUMAN]	16.38	15
P84090	Enhancer of rudimentary homolog OS=Homo sapiens GN=ERH PE=1 SV=1 - [ERH_HUMAN]	16.35	2
Q92841	Probable ATP-dependent RNA helicase DDX17 OS=Homo sapiens GN=DDX17 PE=1 SV=2 - [DDX17_HUMAN]	16.32	23
P00338	L-lactate dehydrogenase A chain OS=Homo sapiens GN=LDHA PE=1 SV=2 - [LDHA_HUMAN]	16.27	10
P62424	60S ribosomal protein L7a OS=Homo sapiens GN=RPL7A PE=1 SV=2 - [RL7A_HUMAN]	16.17	7
P08238	Heat shock protein HSP 90-beta OS=Homo sapiens GN=HSP90AB1 PE=1 SV=4 - [HS90B_HUMAN]	16.16	21
P50990	T-complex protein 1 subunit theta OS=Homo sapiens GN=CCT8 PE=1 SV=4 - [TCPO_HUMAN]	16.06	15
Q9H5H4	Zinc finger protein 768 OS=Homo sapiens GN=ZNF768 PE=1 SV=2 - [ZN768_HUMAN]	15.74	13
Q00839	Heterogeneous nuclear ribonucleoprotein U OS=Homo sapiens GN=HNRNPU PE=1 SV=6 - [HNRPU_HUMAN]	15.64	28
Q15717	ELAV-like protein 1 OS=Homo sapiens GN=ELAVL1 PE=1 SV=2 - [ELAVL_HUMAN]	15.64	9
Q8IWS0	PHD finger protein 6 OS=Homo sapiens GN=PHF6 PE=1 SV=1 - [PHF6_HUMAN]	15.62	9
P31943	Heterogeneous nuclear ribonucleoprotein H OS=Homo sapiens GN=HNRNPH1 PE=1 SV=4 - [HNRH1_HUMAN]	15.59	11
P62851	40S ribosomal protein S25 OS=Homo sapiens GN=RPS25 PE=1 SV=1 - [RS25_HUMAN]	15.20	4
P32119	Peroxisome oxidin-2 OS=Homo sapiens GN=PRDX2 PE=1 SV=5 - [PRDX2_HUMAN]	15.15	6
P14866	Heterogeneous nuclear ribonucleoprotein L OS=Homo sapiens GN=HNRNPL PE=1 SV=2 - [HNRPL_HUMAN]	15.11	18
Q13148	TAR DNA-binding protein 43 OS=Homo sapiens GN=TARDBP PE=1 SV=1 - [TADBP_HUMAN]	14.98	10
P06576	ATP synthase subunit beta, mitochondrial OS=Homo sapiens GN=ATP5B PE=1 SV=3 - [ATPB_HUMAN]	14.74	10
O43175	D-3-phosphoglycerate dehydrogenase OS=Homo sapiens GN=PHGDH PE=1 SV=4 - [SERA_HUMAN]	14.45	12
P50914	60S ribosomal protein L14 OS=Homo sapiens GN=RPL14 PE=1 SV=4 - [RL14_HUMAN]	14.42	4
Q8N726	Cyclin-dependent kinase inhibitor 2A, isoform 4 OS=Homo sapiens GN=CDKN2A PE=1 SV=2 - [CD2A2_HUMAN]	14.39	2
P62826	GTP-binding nuclear protein Ran OS=Homo sapiens GN=RAN PE=1 SV=3 - [RAN_HUMAN]	14.35	8
Q8NC51	Plasminogen activator inhibitor 1 RNA-binding protein OS=Homo sapiens GN=SERBP1 PE=1 SV=2 - [PAIRB_HUMAN]	14.22	7
P62263	40S ribosomal protein S14 OS=Homo sapiens GN=RPS14 PE=1 SV=3 - [RS14_HUMAN]	13.91	4
B2RPK0	Putative high mobility group protein B1-like 1 OS=Homo sapiens GN=HMG1P1 PE=5 SV=1 - [HGB1A_HUMAN]	13.74	6

Accession	Description	ΣCoverage	Σ# PSMs
O60506	Heterogeneous nuclear ribonucleoprotein Q OS=Homo sapiens GN=SYNCRIP PE=1 SV=2 - [HNRPO_HUMAN]	13.64	18
O95232	Luc7-like protein 3 OS=Homo sapiens GN=LUC7L3 PE=1 SV=2 - [LC7L3_HUMAN]	13.43	16
Q96EL3	39S ribosomal protein L53, mitochondrial OS=Homo sapiens GN=MRPL53 PE=1 SV=1 - [RM53_HUMAN]	13.39	2
P83731	60S ribosomal protein L24 OS=Homo sapiens GN=RPL24 PE=1 SV=1 - [RL24_HUMAN]	13.38	4
Q07955	Serine/arginine-rich splicing factor 1 OS=Homo sapiens GN=SRSF1 PE=1 SV=2 - [SRSF1_HUMAN]	13.31	5
O43809	Cleavage and polyadenylation specificity factor subunit 5 OS=Homo sapiens GN=NUDT21 PE=1 SV=1 - [CPSF5_HUMAN]	13.22	4
Q02878	60S ribosomal protein L6 OS=Homo sapiens GN=RPL6 PE=1 SV=3 - [RL6_HUMAN]	13.19	10
P60842	Eukaryotic initiation factor 4A-1 OS=Homo sapiens GN=EIF4A1 PE=1 SV=1 - [IF4A1_HUMAN]	13.05	9
P43243	Matrin-3 OS=Homo sapiens GN=MATR3 PE=1 SV=2 - [MATR3_HUMAN]	12.99	24
P20719	Homeobox protein Hox-A5 OS=Homo sapiens GN=HOXA5 PE=1 SV=2 - [HXA5_HUMAN]	12.96	4
Q9Y265	RuvB-like 1 OS=Homo sapiens GN=RUVBL1 PE=1 SV=1 - [RUVBL1_HUMAN]	12.94	10
P42167	Lamina-associated polypeptide 2, isoforms beta/gamma OS=Homo sapiens GN=TMPO PE=1 SV=2 - [LAP2B_HUMAN]	12.56	8
P13639	Elongation factor 2 OS=Homo sapiens GN=EEF2 PE=1 SV=4 - [EF2_HUMAN]	12.47	20
Q9NY12	H/ACA ribonucleoprotein complex subunit 1 OS=Homo sapiens GN=GAR1 PE=1 SV=1 - [GAR1_HUMAN]	12.44	3
P61254	60S ribosomal protein L26 OS=Homo sapiens GN=RPL26 PE=1 SV=1 - [RL26_HUMAN]	12.41	6
O00483	Cytochrome c oxidase subunit NDUFA4 OS=Homo sapiens GN=NDUFA4 PE=1 SV=1 - [NDUA4_HUMAN]	12.35	2
Q9BZ64	Nucleolar GTP-binding protein 1 OS=Homo sapiens GN=GTPBP4 PE=1 SV=3 - [NOG1_HUMAN]	12.30	16
Q12906	Interleukin enhancer-binding factor 3 OS=Homo sapiens GN=ILF3 PE=1 SV=3 - [ILF3_HUMAN]	11.97	18
Q13247	Serine/arginine-rich splicing factor 6 OS=Homo sapiens GN=SRSF6 PE=1 SV=2 - [SRSF6_HUMAN]	11.92	8
P62258	14-3-3 protein epsilon OS=Homo sapiens GN=YWHAE PE=1 SV=1 - [1433E_HUMAN]	11.76	6
O43390	Heterogeneous nuclear ribonucleoprotein R OS=Homo sapiens GN=HNRNPR PE=1 SV=1 - [HNRPR_HUMAN]	11.69	16
P78549	Endonuclease III-like protein 1 OS=Homo sapiens GN=NTHL1 PE=1 SV=2 - [NTH_HUMAN]	11.54	6
Q12905	Interleukin enhancer-binding factor 2 OS=Homo sapiens GN=ILF2 PE=1 SV=2 - [ILF2_HUMAN]	11.54	10
Q8IYL3	UPF0688 protein Clorf174 OS=Homo sapiens GN=Clorf174 PE=1 SV=2 - [CA174_HUMAN]	11.52	4
P50454	Serpin H1 OS=Homo sapiens GN=SERP1H1 PE=1 SV=2 - [SERPH1_HUMAN]	11.48	8
Q13838	Spliceosome RNA helicase DDX39B OS=Homo sapiens GN=DDX39B PE=1 SV=1 - [DX39B_HUMAN]	11.45	10
P56270	Myc-associated zinc finger protein OS=Homo sapiens GN=MAZ PE=1 SV=1 - [MAZ_HUMAN]	11.32	13
Q9NR30	Nucleolar RNA helicase 2 OS=Homo sapiens GN=DDX21 PE=1 SV=5 - [DDX21_HUMAN]	11.11	13
Q5J7H9	RRP12-like protein OS=Homo sapiens GN=RRP12 PE=1 SV=2 - [RRP12_HUMAN]	11.10	26
Q9Y2X3	Nucleolar protein 58 OS=Homo sapiens GN=NOPS8 PE=1 SV=1 - [NOPS8_HUMAN]	10.96	8
P18077	60S ribosomal protein L35a OS=Homo sapiens GN=RPL35A PE=1 SV=2 - [RL35A_HUMAN]	10.91	2
P62277	40S ribosomal protein S13 OS=Homo sapiens GN=RPS13 PE=1 SV=2 - [RS13_HUMAN]	10.60	6
P26599	Polypyrimidine tract-binding protein 1 OS=Homo sapiens GN=PTBP1 PE=1 SV=1 - [PTBP1_HUMAN]	10.55	12
P37802	Transgelin-2 OS=Homo sapiens GN=TAGLN2 PE=1 SV=3 - [TAGL2_HUMAN]	10.55	4
P05387	60S acidic ribosomal protein P2 OS=Homo sapiens GN=RPL2 PE=1 SV=1 - [RLA2_HUMAN]	10.43	2
H0Y8P6	Peptidylprolyl cis-trans isomerase NIMA-interacting 4 (Fragment) OS=Homo sapiens GN=PIN4 PE=1 SV=1 - [H0Y8P6_HUMAN]	10.43	1
P60468	Protein transport protein Sec61 subunit beta OS=Homo sapiens GN=SEC61B PE=1 SV=2 - [SC61B_HUMAN]	10.42	1
P37108	Signal recognition particle 14 kDa protein OS=Homo sapiens GN=SRP14 PE=1 SV=2 - [SRP14_HUMAN]	10.29	1
P32568	60S ribosomal protein L22 OS=Homo sapiens GN=RPL22 PE=1 SV=2 - [RL22_HUMAN]	10.16	2
P00558	Phosphoglycerate kinase 1 OS=Homo sapiens GN=PGK1 PE=1 SV=3 - [PGK1_HUMAN]	10.07	6
P07355	Annexin A2 OS=Homo sapiens GN=ANXA2 PE=1 SV=2 - [ANXA2_HUMAN]	10.03	6
P20700	Lamin-B1 OS=Homo sapiens GN=LMBN1 PE=1 SV=2 - [LMBN1_HUMAN]	9.90	11
P48730	Casein kinase I isoform delta OS=Homo sapiens GN=CSNK1D PE=1 SV=2 - [KC1D_HUMAN]	9.88	8
P61513	60S ribosomal protein L37a OS=Homo sapiens GN=RPL37A PE=1 SV=2 - [RL37A_HUMAN]	9.78	3
P35908	Keratin, type II cytoskeletal 2 epidermal OS=Homo sapiens GN=KRT2 PE=1 SV=2 - [K2E_HUMAN]	9.55	12
Q9J3J0	CGG triplet repeat-binding protein 1 (Fragment) OS=Homo sapiens GN=CGGBP1 PE=1 SV=1 - [C9J3J0_HUMAN]	9.43	1
P47914	60S ribosomal protein L29 OS=Homo sapiens GN=RPL29 PE=1 SV=2 - [RL29_HUMAN]	9.43	2
Q05991	T-complex protein 1 subunit delta OS=Homo sapiens GN=CCT4 PE=1 SV=4 - [TCPD_HUMAN]	9.28	8
P31942	Heterogeneous nuclear ribonucleoprotein H3 OS=Homo sapiens GN=HNRNPH3 PE=1 SV=2 - [HNRH3_HUMAN]	9.25	7
P62314	Small nuclear ribonucleoprotein Sm D1 OS=Homo sapiens GN=SNRPD1 PE=1 SV=1 - [SMD1_HUMAN]	9.24	2
P15311	Ezrin OS=Homo sapiens GN=EZR PE=1 SV=4 - [EZRI_HUMAN]	9.22	11
P39748	Flap endonuclease 1 OS=Homo sapiens GN=FEN1 PE=1 SV=1 - [FEN1_HUMAN]	9.21	8
P14678	Small nuclear ribonucleoprotein-associated proteins B and B' OS=Homo sapiens GN=SNRNPB PE=1 SV=2 - [RSMB_HUMAN]	9.17	6
P17987	T-complex protein 1 subunit alpha OS=Homo sapiens GN=TCP1 PE=1 SV=1 - [TCPA_HUMAN]	8.99	10
P62917	60S ribosomal protein L8 OS=Homo sapiens GN=RPL8 PE=1 SV=2 - [RL8_HUMAN]	8.95	3
P49411	Elongation factor Tu, mitochondrial OS=Homo sapiens GN=TUFM PE=1 SV=2 - [EFTU_HUMAN]	8.85	5
P63244	Guanine nucleotide-binding protein subunit beta-2-like 1 OS=Homo sapiens GN=GNB2L1 PE=1 SV=3 - [GBLP_HUMAN]	8.83	5
P38919	Eukaryotic initiation factor 4A-III OS=Homo sapiens GN=EIF4A3 PE=1 SV=4 - [IF4A3_HUMAN]	8.76	6
Q01130	Serine/arginine-rich splicing factor 2 OS=Homo sapiens GN=SRSF2 PE=1 SV=4 - [SRSF2_HUMAN]	8.60	4
P13645	Keratin, type I cytoskeletal 10 OS=Homo sapiens GN=KRT10 PE=1 SV=6 - [K1C10_HUMAN]	8.56	12
P62316	Small nuclear ribonucleoprotein Sm D2 OS=Homo sapiens GN=SNRPD2 PE=1 SV=1 - [SMD2_HUMAN]	8.47	2
P07900	Heat shock protein HSP 90-alpha OS=Homo sapiens GN=HSP90AA1 PE=1 SV=5 - [H90A_HUMAN]	8.47	11
Q14684	Ribosomal RNA processing protein 1 homolog B OS=Homo sapiens GN=RRP1B PE=1 SV=3 - [RRP1B_HUMAN]	8.31	12
P62081	40S ribosomal protein S7 OS=Homo sapiens GN=RPS7 PE=1 SV=1 - [RS7_HUMAN]	8.25	4
Q92945	Far upstream element-binding protein 2 OS=Homo sapiens GN=KHFRP PE=1 SV=4 - [FUBP2_HUMAN]	8.16	9
Q96RQ3	Methylcrotonoyl-CoA carboxylase subunit alpha, mitochondrial OS=Homo sapiens GN=MCCCI PE=1 SV=3 - [MCCA_HUMAN]	8.14	10
Q13263	Transcription intermediary factor 1-beta OS=Homo sapiens GN=TRIM28 PE=1 SV=5 - [TIF1B_HUMAN]	8.14	10
P06493	Cyclin-dependent kinase 1 OS=Homo sapiens GN=CDK1 PE=1 SV=3 - [CDK1_HUMAN]	8.08	4
P36542	ATP synthase subunit gamma, mitochondrial OS=Homo sapiens GN=ATP5C1 PE=1 SV=1 - [ATPG_HUMAN]	8.05	3
C9J384	Protein CMSS1 (Fragment) OS=Homo sapiens GN=CMSS1 PE=1 SV=1 - [C9J384_HUMAN]	8.00	2
P62995	Transformer-2 protein homolog beta OS=Homo sapiens GN=TRA2B PE=1 SV=1 - [TRA2B_HUMAN]	7.99	4
Q147U1	Zinc finger protein 846 OS=Homo sapiens GN=ZNF846 PE=1 SV=2 - [ZN846_HUMAN]	7.88	2
P14174	Macrophage migration inhibitory factor OS=Homo sapiens GN=MIF PE=1 SV=4 - [MIF_HUMAN]	7.83	2
C9J8D0	KRAB domain-containing protein 1 OS=Homo sapiens GN=KRBOX1 PE=2 SV=1 - [KRBX1_HUMAN]	7.81	1
Q08211	ATP-dependent RNA helicase A OS=Homo sapiens GN=DHX9 PE=1 SV=4 - [DHX9_HUMAN]	7.72	18
P49207	60S ribosomal protein L34 OS=Homo sapiens GN=RPL34 PE=1 SV=3 - [RL34_HUMAN]	7.69	2
C9JPD0	Neurexophilin-1 (Fragment) OS=Homo sapiens GN=NXPH1 PE=4 SV=1 - [C9JPD0_HUMAN]	7.69	1
P18124	60S ribosomal protein L7 OS=Homo sapiens GN=RPL7 PE=1 SV=1 - [RL7_HUMAN]	7.66	2
P35637	RNA-binding protein FUS OS=Homo sapiens GN=FUS PE=1 SV=1 - [FUS_HUMAN]	7.60	4



Accession	Description	ΣCoverage	Σ# PSMs
P61077	Ubiquitin-conjugating enzyme E2 D3 OS=Homo sapiens GN=UBE2D3 PE=1 SV=1 - [UB2D3_HUMAN]	7.48	1
P35659	Protein DEK OS=Homo sapiens GN=DEK PE=1 SV=1 - [DEK_HUMAN]	7.47	6
Q15691	Microtubule-associated protein RP/EB family member 1 OS=Homo sapiens GN=MAPRE1 PE=1 SV=3 - [MARE1_HUMAN]	7.46	2
P23284	Peptidyl-prolyl cis-trans isomerase B OS=Homo sapiens GN=PP1B PE=1 SV=2 - [PP1B_HUMAN]	7.41	4
Q14974	Importin subunit beta-1 OS=Homo sapiens GN=KPMB1 PE=1 SV=2 - [IMB1_HUMAN]	7.31	10
P07477	Trypsin-1 OS=Homo sapiens GN=PRSS1 PE=1 SV=1 - [TRY1_HUMAN]	7.29	17
O15235	28S ribosomal protein S12, mitochondrial OS=Homo sapiens GN=MRPS12 PE=1 SV=1 - [RT12_HUMAN]	7.25	1
Q07666	KH domain-containing, RNA-binding, signal transduction-associated protein 1 OS=Homo sapiens GN=KHDRB51 PE=1 SV=1 - [KHDR1_HUMAN]	7.22	9
P61927	60S ribosomal protein L37 OS=Homo sapiens GN=RPL37 PE=1 SV=2 - [RL37_HUMAN]	7.22	2
P62750	60S ribosomal protein L23a OS=Homo sapiens GN=RPL23A PE=1 SV=1 - [RL23A_HUMAN]	7.05	2
P78527	DNA-dependent protein kinase catalytic subunit OS=Homo sapiens GN=PRKDC PE=1 SV=3 - [PRKDC_HUMAN]	7.05	52
Q96AE4	Far upstream element-binding protein 1 OS=Homo sapiens GN=FUBP1 PE=1 SV=3 - [FUBP1_HUMAN]	6.99	9
P40429	60S ribosomal protein L13a OS=Homo sapiens GN=RPL13A PE=1 SV=2 - [RL13A_HUMAN]	6.90	4
P18754	Regulator of chromosome condensation OS=Homo sapiens GN=RCC1 PE=1 SV=1 - [RCC1_HUMAN]	6.89	4
Q9NRW3	DNA dC->dU-editing enzyme APOBEC-3C OS=Homo sapiens GN=APOBEC3C PE=1 SV=2 - [ABC3C_HUMAN]	6.84	2
Q9NSQ0	Putative ribosomal RNA-processing protein 7 homolog B OS=Homo sapiens GN=RRP7B PE=5 SV=1 - [RRP7B_HUMAN]	6.80	1
P17039	Zinc finger protein 30 OS=Homo sapiens GN=ZNF30 PE=2 SV=5 - [ZNF30_HUMAN]	6.74	3
Q14119	Vascular endothelial zinc finger 1 OS=Homo sapiens GN=VEZF1 PE=1 SV=2 - [VEZF1_HUMAN]	6.72	6
Q57280	Putative methyltransferase C9orf114 OS=Homo sapiens GN=C9orf114 PE=1 SV=3 - [C114_HUMAN]	6.65	3
P08621	U1 small nuclear ribonucleoprotein 70 kDa OS=Homo sapiens GN=SNRNP70 PE=1 SV=2 - [RU17_HUMAN]	6.64	6
P61353	60S ribosomal protein L27 OS=Homo sapiens GN=RPL27 PE=1 SV=2 - [RL27_HUMAN]	6.62	2
P21333	Filamin-A OS=Homo sapiens GN=FLNA PE=1 SV=4 - [FLNA_HUMAN]	6.42	24
Q9NRX1	RNA-binding protein PNO1 OS=Homo sapiens GN=PNO1 PE=1 SV=1 - [PNO1_HUMAN]	6.35	2
P22234	Multifunctional protein ADE2 OS=Homo sapiens GN=PAICS PE=1 SV=3 - [PUR6_HUMAN]	6.35	3
J3KPG2	Transcriptionally-controlled tumor protein OS=Homo sapiens GN=TPT1 PE=1 SV=1 - [J3KPG2_HUMAN]	6.25	1
Q15424	Scaffold attachment factor B1 OS=Homo sapiens GN=SAFB PE=1 SV=4 - [SAFB1_HUMAN]	6.23	10
O00567	Nucleolar protein 56 OS=Homo sapiens GN=NOPS6 PE=1 SV=4 - [NOPS6_HUMAN]	6.23	6
Q86U42	Polyadenylate-binding protein 2 OS=Homo sapiens GN=PABPN1 PE=1 SV=3 - [PABP2_HUMAN]	6.21	4
Q96127	Zinc finger protein 625 OS=Homo sapiens GN=ZNF625 PE=2 SV=1 - [ZNF625_HUMAN]	6.21	3
Q92804	TATA-binding protein-associated factor 2N OS=Homo sapiens GN=TAF15 PE=1 SV=1 - [RBP56_HUMAN]	6.08	2
P26308	Moesin OS=Homo sapiens GN=MSN PE=1 SV=3 - [MOES_HUMAN]	6.07	6
P25398	40S ribosomal protein S12 OS=Homo sapiens GN=RPS12 PE=1 SV=3 - [RS12_HUMAN]	6.06	1
P20042	Eukaryotic translation initiation factor 2 subunit 2 OS=Homo sapiens GN=EIF2S2 PE=1 SV=2 - [IF2B_HUMAN]	6.01	1
Q9BRX2	Protein pelota homolog OS=Homo sapiens GN=PELO PE=1 SV=2 - [PELO_HUMAN]	5.97	2
F8W7Q3	WD repeat-containing protein 48 OS=Homo sapiens GN=WDR48 PE=1 SV=1 - [F8W7Q3_HUMAN]	5.97	1
P00492	Hypoxanthine-guanine phosphoribosyltransferase OS=Homo sapiens GN=HPRT1 PE=1 SV=2 - [HPRT_HUMAN]	5.96	2
P55265	Double-stranded RNA-specific adenosine deaminase OS=Homo sapiens GN=ADAR PE=1 SV=4 - [DSRAD_HUMAN]	5.95	12
Q6R954	Polymerase delta interacting protein 46 OS=Homo sapiens GN=PDIP46 PE=1 SV=1 - [Q6R954_HUMAN]	5.95	1
P61313	60S ribosomal protein L15 OS=Homo sapiens GN=RPL15 PE=1 SV=2 - [RL15_HUMAN]	5.88	1
Q9BQ60	Myb-binding protein 1A OS=Homo sapiens GN=MYBBP1A PE=1 SV=2 - [MBB1A_HUMAN]	5.87	15
O95196-3	Isoform 3 of Chondroitin sulfate proteoglycan 5 OS=Homo sapiens GN=CSPG5 - [CSPG5_HUMAN]	5.74	1
C9JGC1	Zinc finger protein neuro-d4 (Fragment) OS=Homo sapiens GN=DPF1 PE=4 SV=1 - [C9JGC1_HUMAN]	5.71	2
P33993	DNA replication licensing factor MCM7 OS=Homo sapiens GN=MCM7 PE=1 SV=4 - [MCM7_HUMAN]	5.56	8
Q9Y324	rRNA-processing protein FCF1 homolog OS=Homo sapiens GN=FCF1 PE=2 SV=1 - [FCF1_HUMAN]	5.56	2
P61326	Protein mago nashi homolog OS=Homo sapiens GN=MAGOHP PE=1 SV=1 - [MGN_HUMAN]	5.48	1
Q15084	Protein disulfide-isomerase A6 OS=Homo sapiens GN=PDIA6 PE=1 SV=1 - [PDIA6_HUMAN]	5.45	3
Q81ZP2	Putative protein FAM10A4 OS=Homo sapiens GN=ST13P4 PE=5 SV=1 - [ST134_HUMAN]	5.42	2
P78371	T-complex protein 1 subunit beta OS=Homo sapiens GN=CCT2 PE=1 SV=4 - [TCTP_HUMAN]	5.42	4
P49368	T-complex protein 1 subunit gamma OS=Homo sapiens GN=CCT3 PE=1 SV=4 - [TCTP_HUMAN]	5.32	6
P25292	Importin subunit alpha-1 OS=Homo sapiens GN=KPN2 PE=1 SV=1 - [IMA1_HUMAN]	5.29	4
F8VYV2	60S ribosomal protein L18 OS=Homo sapiens GN=RPL18 PE=1 SV=1 - [F8VYV2_HUMAN]	5.26	1
Q9Y3Y2	Chromatin target of PRMT1 protein OS=Homo sapiens GN=CHTOP PE=1 SV=2 - [CHTOP_HUMAN]	5.24	2
O60264	SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A member 5 OS=Homo sapiens GN=SMARCA5 PE=1 SV=1 - [SMCA5_HUMAN]	5.23	9
P39023	60S ribosomal protein L3 OS=Homo sapiens GN=RPL3 PE=1 SV=2 - [RL3_HUMAN]	5.21	4
A0A087WDX8	PH and SEC7 domain-containing protein 3 OS=Homo sapiens GN=PSD3 PE=4 SV=1 - [A0A087WDX8_HUMAN]	5.20	1
P78347	General transcription factor II-1 OS=Homo sapiens GN=GTF2I PE=1 SV=2 - [GTF2I_HUMAN]	5.11	7
Q6HA08	Astacin-like metalloendopeptidase OS=Homo sapiens GN=ASTL PE=1 SV=4 - [ASTL_HUMAN]	5.10	1
C9JYW2	Nucleolin (Fragment) OS=Homo sapiens GN=NCL PE=1 SV=1 - [C9JYW2_HUMAN]	4.96	1
P25205	DNA replication licensing factor MCM3 OS=Homo sapiens GN=MCM3 PE=1 SV=3 - [MCM3_HUMAN]	4.95	10
O00541	Pescadillo homolog OS=Homo sapiens GN=PES1 PE=1 SV=1 - [PESC_HUMAN]	4.93	6
Q9NVP1	ATP-dependent RNA helicase DDX18 OS=Homo sapiens GN=DDX18 PE=1 SV=2 - [DDX18_HUMAN]	4.93	7
P08579	U2 small nuclear ribonucleoprotein B' OS=Homo sapiens GN=SNRNP2 PE=1 SV=1 - [RU2B_HUMAN]	4.89	2
Q5C924	Nucleolar MIF4G domain-containing protein 1 OS=Homo sapiens GN=NOM1 PE=1 SV=1 - [NOM1_HUMAN]	4.88	5
O14602	Eukaryotic translation initiation factor 1A, Y-chromosomal OS=Homo sapiens GN=EIF1AY PE=1 SV=4 - [IF1AY_HUMAN]	4.86	2
P49959	Double-strand break repair protein MRE11A OS=Homo sapiens GN=MRE11A PE=1 SV=3 - [MRE11_HUMAN]	4.80	6
O14979	Heterogeneous nuclear ribonucleoprotein D-like OS=Homo sapiens GN=HNRNPDL PE=1 SV=3 - [HNRDL_HUMAN]	4.76	4
P82979	SAP domain-containing ribonucleoprotein OS=Homo sapiens GN=SARNP PE=1 SV=3 - [SARNP_HUMAN]	4.76	2
P12956	X-ray repair cross-complementing protein 6 OS=Homo sapiens GN=XRCC6 PE=1 SV=2 - [XRCC6_HUMAN]	4.76	6
P46087	Probable 28S rRNA (cytosine(4447)-C(5))-methyltransferase OS=Homo sapiens GN=NOP2 PE=1 SV=2 - [NOP2_HUMAN]	4.68	5
P07305	Histone H1.0 OS=Homo sapiens GN=H1FO PE=1 SV=3 - [H1O_HUMAN]	4.64	1
MQQZL6	Cytokine receptor-like factor 1 (Fragment) OS=Homo sapiens GN=CRLF1 PE=4 SV=1 - [MQQZL6_HUMAN]	4.62	1
O75533	Splicing factor 3B subunit 1 OS=Homo sapiens GN=SF3B1 PE=1 SV=3 - [SF3B1_HUMAN]	4.60	7
P11021	78 kDa glucose-regulated protein OS=Homo sapiens GN=HSPA5 PE=1 SV=2 - [GRP78_HUMAN]	4.59	6
Q9HOC8	Integrin-linked kinase-associated serine/threonine phosphatase 2C OS=Homo sapiens GN=ILKAP PE=1 SV=1 - [ILKAP_HUMAN]	4.59	2
O75367	Core histone macro-H2A.1 OS=Homo sapiens GN=H2AFY1 PE=1 SV=4 - [H2AY_HUMAN]	4.57	2
Q6P087	RNA pseudouridylylase synthase domain-containing protein 3 OS=Homo sapiens GN=RPUSD3 PE=1 SV=3 - [RUSD3_HUMAN]	4.56	4
P30101	Protein disulfide-isomerase A3 OS=Homo sapiens GN=PDIA3 PE=1 SV=4 - [PDIA3_HUMAN]	4.55	4
P60174	Triosephosphate isomerase OS=Homo sapiens GN=TP11 PE=1 SV=3 - [TPIS_HUMAN]	4.55	2
P11177	Pyruvate dehydrogenase E1 component subunit beta, mitochondrial OS=Homo sapiens GN=PDHB PE=1 SV=3 - [ODPB_HUMAN]	4.46	1

Accession	Description	ΣCoverage	Σ# PSMs
H7CON4	Splicing factor 1 (Fragment) OS=Homo sapiens GN=SF1 PE=1 SV=1 - [H7CON4_HUMAN]	4.44	1
Q9BYG3	MK167 FHA domain-interacting nucleolar phosphoprotein OS=Homo sapiens GN=NIFK PE=1 SV=1 - [MK671_HUMAN]	4.44	1
P46778	60S ribosomal protein L21 OS=Homo sapiens GN=RPL21 PE=1 SV=2 - [RL21_HUMAN]	4.38	1
Q9Y2R4	Probable ATP-dependent RNA helicase DDX52 OS=Homo sapiens GN=DDX52 PE=1 SV=3 - [DDX52_HUMAN]	4.34	4
Q9Y383	Putative RNA-binding protein Luc7-like 2 OS=Homo sapiens GN=LUC7L2 PE=1 SV=2 - [LC7L2_HUMAN]	4.34	3
Q86T24	Transcriptional regulator Kaiso OS=Homo sapiens GN=ZBTB33 PE=1 SV=2 - [KAISO_HUMAN]	4.32	6
Q99986	Serine/threonine-protein kinase VRK1 OS=Homo sapiens GN=VRK1 PE=1 SV=1 - [VRK1_HUMAN]	4.29	2
Q08170	Serine/arginine-rich splicing factor 4 OS=Homo sapiens GN=SRSF4 PE=1 SV=2 - [SRSF4_HUMAN]	4.25	4
Q99832	T-complex protein 1 subunit eta OS=Homo sapiens GN=CCT7 PE=1 SV=2 - [TCPH_HUMAN]	4.24	3
Q9UM54	Pre-mRNA-processing factor 19 OS=Homo sapiens GN=PRPF19 PE=1 SV=1 - [PRP19_HUMAN]	4.17	3
Q14980	Nuclear mitotic apparatus protein 1 OS=Homo sapiens GN=NUMA1 PE=1 SV=2 - [NUMA1_HUMAN]	4.16	14
Q00325	Phosphate carrier protein, mitochondrial OS=Homo sapiens GN=SLC25A3 PE=1 SV=2 - [MPCP_HUMAN]	4.14	6
Q15072	Zinc finger protein OZF OS=Homo sapiens GN=ZNF146 PE=1 SV=2 - [OZF_HUMAN]	4.11	1
000139	Kinesin-like protein KIF2A OS=Homo sapiens GN=KIF2A PE=1 SV=3 - [KIF2A_HUMAN]	4.11	6
Q5QNZZ	ATP synthase F(0) complex subunit B1, mitochondrial OS=Homo sapiens GN=ATP5F1 PE=1 SV=1 - [Q5QNZZ_HUMAN]	4.10	1
P43246	DNA mismatch repair protein Msh2 OS=Homo sapiens GN=MSH2 PE=1 SV=1 - [MSH2_HUMAN]	4.07	7
P35232	Prohibitin OS=Homo sapiens GN=PHB PE=1 SV=1 - [PHB_HUMAN]	4.04	2
P26368	Splicing factor U2AF 65 kDa subunit OS=Homo sapiens GN=U2AF2 PE=1 SV=4 - [U2AF2_HUMAN]	4.00	3
O15347	High mobility group protein B3 OS=Homo sapiens GN=HMGB3 PE=1 SV=4 - [HMGB3_HUMAN]	4.00	2
P46109	Crk-like protein OS=Homo sapiens GN=CRKL PE=1 SV=1 - [CRKL_HUMAN]	3.96	2
Q14498	RNA-binding protein 39 OS=Homo sapiens GN=RBM39 PE=1 SV=2 - [RBM39_HUMAN]	3.96	4
Q9NV77	ATPase family AAA domain-containing protein 3A OS=Homo sapiens GN=ATAD3A PE=1 SV=2 - [ATD3A_HUMAN]	3.94	5
O60832	H/JACA ribonucleoprotein complex subunit 4 OS=Homo sapiens GN=DKC1 PE=1 SV=3 - [DKC1_HUMAN]	3.89	3
P25789	Proteasome subunit alpha type-4 OS=Homo sapiens GN=PSMA4 PE=1 SV=1 - [PSA4_HUMAN]	3.83	2
P31689	DnaJ homolog subfamily A member 1 OS=Homo sapiens GN=DNAJ1 PE=1 SV=2 - [DNA1_HUMAN]	3.78	2
Q96T88	E3 ubiquitin-protein ligase UHRF1 OS=Homo sapiens GN=UHRF1 PE=1 SV=1 - [UHRF1_HUMAN]	3.78	5
O75475	PC4 and SFRS1-interacting protein OS=Homo sapiens GN=PSIP1 PE=1 SV=1 - [PSIP1_HUMAN]	3.77	6
C9JN71	Zinc finger protein 878 OS=Homo sapiens GN=ZNF878 PE=3 SV=2 - [ZNF878_HUMAN]	3.77	3
Q9BPX3	Condensin complex subunit 3 OS=Homo sapiens GN=NCAPG PE=1 SV=1 - [CND3_HUMAN]	3.74	3
P62906	60S ribosomal protein L10a OS=Homo sapiens GN=RPL10A PE=1 SV=2 - [RL10A_HUMAN]	3.69	2
Q9Y3A5	Ribosome maturation protein SBDS OS=Homo sapiens GN=SBDS PE=1 SV=4 - [SBDS_HUMAN]	3.60	1
P21796	Voltage-dependent anion-selective channel protein 1 OS=Homo sapiens GN=VDAC1 PE=1 SV=2 - [VDAC1_HUMAN]	3.53	4
P31948	Stress-induced-phosphoprotein 1 OS=Homo sapiens GN=STIP1 PE=1 SV=1 - [STIP1_HUMAN]	3.50	1
P43487	Ran-specific GTPase-activating protein OS=Homo sapiens GN=RANBP1 PE=1 SV=1 - [RANG_HUMAN]	3.48	1
P05388	60S acidic ribosomal protein P0 OS=Homo sapiens GN=RPLP0 PE=1 SV=1 - [RLA0_HUMAN]	3.47	4
Q00918	Carbonic anhydrase 2 OS=Homo sapiens GN=CA2 PE=1 SV=2 - [CAH2_HUMAN]	3.46	2
P82DN6	Ribosome biogenesis protein BRX1 homolog OS=Homo sapiens GN=BRX1 PE=1 SV=2 - [BRX1_HUMAN]	3.40	2
P18846	Cyclic AMP-dependent transcription factor ATF-1 OS=Homo sapiens GN=ATF1 PE=1 SV=2 - [ATF1_HUMAN]	3.32	2
A8M236	Envoplakin-like protein OS=Homo sapiens GN=EVPL PE=2 SV=1 - [EVPL_HUMAN]	3.32	1
P11388	DNA topoisomerase 2-alpha OS=Homo sapiens GN=TOP2A PE=1 SV=3 - [TOP2A_HUMAN]	3.27	10
Q9Y2L1	Exosome complex exonuclease RRP44 OS=Homo sapiens GN=DIS3 PE=1 SV=2 - [RRP44_HUMAN]	3.24	3
Q9Y5J1	U3 small nucleolar RNA-associated protein 18 homolog OS=Homo sapiens GN=UTP18 PE=1 SV=3 - [UTP18_HUMAN]	3.24	2
Q8WVY3	U4/U6 small nuclear ribonucleoprotein Prp31 OS=Homo sapiens GN=PRPF31 PE=1 SV=2 - [PRP31_HUMAN]	3.21	3
P49792	E3 SUMO-protein ligase RanBP2 OS=Homo sapiens GN=RANBP2 PE=1 SV=2 - [RBP2_HUMAN]	3.19	15
P24534	Elongation factor 1-beta OS=Homo sapiens GN=EEF1B2 PE=1 SV=3 - [EF1B_HUMAN]	3.11	1
Q9P016	Thymocyte nuclear protein 1 OS=Homo sapiens GN=THYN1 PE=1 SV=1 - [THYN1_HUMAN]	3.11	1
P36873	Serine/threonine-protein phosphatase PPI-1 gamma catalytic subunit OS=Homo sapiens GN=PPP1CC PE=1 SV=1 - [PPI1_HUMAN]	3.10	2
P38646	Stress-70 protein, mitochondrial OS=Homo sapiens GN=HSPA9 PE=1 SV=2 - [GRP75_HUMAN]	3.09	4
Q8IZJ6	Inactive L-threonine 3-dehydrogenase, mitochondrial OS=Homo sapiens GN=TDH PE=2 SV=1 - [TDH_HUMAN]	3.04	1
Q9UKM9	RNA-binding protein Raly OS=Homo sapiens GN=RALY PE=1 SV=1 - [RALY_HUMAN]	2.94	2
Q14204	Cytoplasmic dynein 1 heavy chain 1 OS=Homo sapiens GN=DYNC1H1 PE=1 SV=5 - [DYHC1_HUMAN]	2.93	26
P22314	Ubiquitin-like modifier-activating enzyme 1 OS=Homo sapiens GN=UBA1 PE=1 SV=3 - [UBA1_HUMAN]	2.93	5
P53396	Cleavage and polyadenylation specificity factor subunit 7 OS=Homo sapiens GN=CPSF7 PE=1 SV=1 - [CPSF7_HUMAN]	2.91	5
Q96EK4	THAP domain-containing protein 11 OS=Homo sapiens GN=THAP11 PE=1 SV=2 - [THA11_HUMAN]	2.87	2
Q5T6C4	Ataxin-7-like protein 2 OS=Homo sapiens GN=ATXN7L2 PE=4 SV=1 - [Q5T6C4_HUMAN]	2.87	1
Q53FP0-2	Isoform 2 of Fibronectin type III domain-containing protein 3B OS=Homo sapiens GN=FND3B - [FND3B_HUMAN]	2.86	1
P49915-2	Isoform 2 of GMP synthase [glutamine-hydrolyzing] OS=Homo sapiens GN=GMPS - [GUAA_HUMAN]	2.86	2
Q96PK6	RNA-binding protein 14 OS=Homo sapiens GN=RBM14 PE=1 SV=2 - [RBM14_HUMAN]	2.84	4
Q9Y230	RuvB-like 2 OS=Homo sapiens GN=RUVBL2 PE=1 SV=3 - [RUVB2_HUMAN]	2.81	2
Q9UQ80	Proliferation-associated protein 2G4 OS=Homo sapiens GN=PA2G4 PE=1 SV=3 - [PA2G4_HUMAN]	2.79	1
Q9NZI8	Insulin-like growth factor 2 mRNA-binding protein 1 OS=Homo sapiens GN=IGF2BP1 PE=1 SV=2 - [IF2B1_HUMAN]	2.77	4
Q8N684	Cleavage and polyadenylation specificity factor subunit 7 OS=Homo sapiens GN=CPSF7 PE=1 SV=1 - [CPSF7_HUMAN]	2.76	2
P12004	Proliferating cell nuclear antigen OS=Homo sapiens GN=PCNA PE=1 SV=1 - [PCNA_HUMAN]	2.68	1
Q9Y589	FACT complex subunit SPT16 OS=Homo sapiens GN=SUPT16H PE=1 SV=1 - [SPT16H_HUMAN]	2.67	5
B7ZM13	ITPR1 protein OS=Homo sapiens GN=ITPR1 PE=2 SV=1 - [B7ZM13_HUMAN]	2.64	1
Q9H7B2	Ribosome production factor 2 homolog OS=Homo sapiens GN=RPF2 PE=1 SV=2 - [RPF2_HUMAN]	2.61	1
P33992	DNA replication licensing factor MCM5 OS=Homo sapiens GN=MCM5 PE=1 SV=5 - [MCM5_HUMAN]	2.59	4
P36578	60S ribosomal protein L4 OS=Homo sapiens GN=RPL4 PE=1 SV=5 - [RL4_HUMAN]	2.58	2
Q16630	Cleavage and polyadenylation specificity factor subunit 6 OS=Homo sapiens GN=CPSF6 PE=1 SV=2 - [CPSF6_HUMAN]	2.54	2
P12270	Nucleoprotein TPR OS=Homo sapiens GN=TPR PE=1 SV=3 - [TPR_HUMAN]	2.54	10
Q75643	U5 small nuclear ribonucleoprotein 200 kDa helicase OS=Homo sapiens GN=SNRNP200 PE=1 SV=2 - [U520_HUMAN]	2.48	8
Q9UBX7	Kallikrein-11 OS=Homo sapiens GN=KLK11 PE=1 SV=2 - [KLK11_HUMAN]	2.48	1
O43684	Mitotic checkpoint protein BUB3 OS=Homo sapiens GN=BUB3 PE=1 SV=1 - [BUB3_HUMAN]	2.44	1
Q969G3	SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily E member 1 OS=Homo sapiens GN=SMARCE1 PE=1 SV=2 - [SMCE1_HUMAN]	2.43	2
O15391	Transcription factor YY2 OS=Homo sapiens GN=YY2 PE=2 SV=1 - [TY2_HUMAN]	2.42	2
Q9H6J7	UPF0705 protein C11orf49 OS=Homo sapiens GN=C11orf49 PE=2 SV=2 - [CK049_HUMAN]	2.42	1
P39656	Dolichyl-diphosphooligosaccharide-protein glycosyltransferase 48 kDa subunit OS=Homo sapiens GN=DDOST PE=1 SV=4 - [OST48_HUMAN]	2.41	2
P40939	Trifunctional enzyme subunit alpha, mitochondrial OS=Homo sapiens GN=HADHA PE=1 SV=2 - [ECHA_HUMAN]	2.36	3
Q86VP6	Cullin-associated NECD8-dissociated protein 1 OS=Homo sapiens GN=CAND1 PE=1 SV=2 - [CAND1_HUMAN]	2.36	6

Accession	Description	ΣCoverage	Σ# PSMs
P49711	Transcriptional repressor CTCF OS=Homo sapiens GN=CTCF PE=1 SV=1 - [CTCF_HUMAN]	2.34	3
P12268	Inosine-5'-monophosphate dehydrogenase 2 OS=Homo sapiens GN=IMPDH2 PE=1 SV=2 - [IMDH2_HUMAN]	2.33	1
P13010	X-ray repair cross-complementing protein 5 OS=Homo sapiens GN=XRCC5 PE=1 SV=3 - [XRCC5_HUMAN]	2.32	3
O43929	Origin recognition complex subunit 4 OS=Homo sapiens GN=ORC4 PE=1 SV=2 - [ORC4_HUMAN]	2.29	2
Q14683	Structural maintenance of chromosomes protein 1A OS=Homo sapiens GN=SMC1A PE=1 SV=2 - [SMC1A_HUMAN]	2.27	4
Q32CQ8	Mitochondrial import inner membrane translocase subunit TIM50 OS=Homo sapiens GN=TIM50 PE=1 SV=2 - [TIM50_HUMAN]	2.27	2
Q14807	Kinesin-like protein KIF22 OS=Homo sapiens GN=KIF22 PE=1 SV=5 - [KIF22_HUMAN]	2.26	1
Q99615	DnaJ homolog subfamily C member 7 OS=Homo sapiens GN=DNAJC7 PE=1 SV=2 - [DNJC7_HUMAN]	2.23	2
P13797	Plastin-3 OS=Homo sapiens GN=PLS3 PE=1 SV=4 - [PLST_HUMAN]	2.22	2
P42704	Leucine-rich PPR motif-containing protein, mitochondrial OS=Homo sapiens GN=LRPPRC PE=1 SV=3 - [LRPPRC_HUMAN]	2.22	6
P35579	Myosin-9 OS=Homo sapiens GN=MYH9 PE=1 SV=4 - [MYH9_HUMAN]	2.19	6
O00442	RNA 3'-terminal phosphate cyclase OS=Homo sapiens GN=RTCA PE=1 SV=1 - [RTCA_HUMAN]	2.19	1
P12532	Creatine kinase U-type, mitochondrial OS=Homo sapiens GN=CKMT1A PE=1 SV=1 - [KCRU_HUMAN]	2.16	2
Q09666	Neuroblast differentiation-associated protein AHNAK OS=Homo sapiens GN=AHNAK PE=1 SV=2 - [AHNK_HUMAN]	2.14	2
Q01844	RNA-binding protein EWS OS=Homo sapiens GN=EWSR1 PE=1 SV=1 - [EWS_HUMAN]	2.13	1
Q96JP5	E3 ubiquitin-protein ligase ZFP91 OS=Homo sapiens GN=ZFP91 PE=1 SV=1 - [ZFP91_HUMAN]	2.11	2
Q9NVA2	Septin-11 OS=Homo sapiens GN=SEPT11 PE=1 SV=3 - [SEP11_HUMAN]	2.10	2
P11387	DNA topoisomerase 1 OS=Homo sapiens GN=TOP1 PE=1 SV=2 - [TOP1_HUMAN]	2.09	4
P23526	Adenosylhomocysteinase OS=Homo sapiens GN=AHCY PE=1 SV=4 - [SAHH_HUMAN]	2.08	1
Q8N1F7	Nuclear pore complex protein Nup93 OS=Homo sapiens GN=NUP93 PE=1 SV=2 - [NUP93_HUMAN]	2.08	4
P40926	Malate dehydrogenase, mitochondrial OS=Homo sapiens GN=MDH2 PE=1 SV=3 - [MDHM_HUMAN]	2.07	1
Q14781	Chromobox protein homolog 2 OS=Homo sapiens GN=CBX2 PE=1 SV=2 - [CBX2_HUMAN]	2.07	1
K7ENW7	DNA (cytosine-5)-methyltransferase 1 (Fragment) OS=Homo sapiens GN=DNMT1 PE=1 SV=1 - [K7ENW7_HUMAN]	2.06	1
Q9NR56	Muscleblind-like protein 1 OS=Homo sapiens GN=MBNL1 PE=1 SV=2 - [MBNL1_HUMAN]	2.06	1
Q15637	Splicing factor 1 OS=Homo sapiens GN=SF1 PE=1 SV=4 - [SF01_HUMAN]	2.03	2
Q9H8H2	Probable ATP-dependent RNA helicase DDX31 OS=Homo sapiens GN=DDX31 PE=1 SV=2 - [DDX31_HUMAN]	2.00	2
Q8WXA9	Splicing regulatory glutamine/lysine-rich protein 1 OS=Homo sapiens GN=SREK1 PE=1 SV=1 - [SREK1_HUMAN]	1.97	1
P56182	Ribosomal RNA processing protein 1 homolog A OS=Homo sapiens GN=RRP1 PE=1 SV=1 - [RRP1_HUMAN]	1.95	1
O75390	Citrate synthase, mitochondrial OS=Homo sapiens GN=CS PE=1 SV=2 - [CISY_HUMAN]	1.93	2
P61619	Protein transport protein Sec61 subunit alpha isoform 1 OS=Homo sapiens GN=SEC61A1 PE=1 SV=2 - [S61A1_HUMAN]	1.89	1
Q92925	SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily D member 2 OS=Homo sapiens GN=SMARCD2 PE=1 SV=3 - [SMRD2_HUMAN]	1.88	1
Q96P11	Probable 28S rRNA (cytosine-C(5))-methyltransferase OS=Homo sapiens GN=NSUN5 PE=1 SV=2 - [NSUN5_HUMAN]	1.86	2
P57740	Nuclear pore complex protein Nup107 OS=Homo sapiens GN=NUP107 PE=1 SV=1 - [NU107_HUMAN]	1.84	1
Q14839	Chromodomain-helicase-DNA-binding protein 4 OS=Homo sapiens GN=CHD4 PE=1 SV=2 - [CHD4_HUMAN]	1.83	6
Q08945	FACT complex subunit SSRP1 OS=Homo sapiens GN=SSRP1 PE=1 SV=1 - [SSRP1_HUMAN]	1.83	2
Q95831	Apoptosis-inducing factor 1, mitochondrial OS=Homo sapiens GN=AIFM1 PE=1 SV=1 - [AIFM1_HUMAN]	1.79	2
P46013	Antigen KI-67 OS=Homo sapiens GN=MKI67 PE=1 SV=2 - [KI67_HUMAN]	1.78	8
P49756	RNA-binding protein 25 OS=Homo sapiens GN=RBM25 PE=1 SV=3 - [RBM25_HUMAN]	1.78	1
Q9Y3T9	Nucleolar complex protein 2 homolog OS=Homo sapiens GN=NOC2L PE=1 SV=4 - [NOC2L_HUMAN]	1.74	2
Q14008	Cytoskeleton-associated protein 5 OS=Homo sapiens GN=CKAP5 PE=1 SV=3 - [CKAP5_HUMAN]	1.72	4
H7C022	WD repeat-containing protein 60 (Fragment) OS=Homo sapiens GN=WDR60 PE=1 SV=1 - [H7C022_HUMAN]	1.72	1
Q9BX9-4	Isoform 2 of N-alpha-acetyltransferase 15, NatA auxiliary subunit OS=Homo sapiens GN=NAA15 - [NAA15_HUMAN]	1.71	1
P30153	Serine/threonine-protein phosphatase 2A 65 kDa regulatory subunit A alpha isoform OS=Homo sapiens GN=PPP2R1A PE=1 SV=4 - [2AAA_HUMAN]	1.70	2
Q9UBD5	Origin recognition complex subunit 3 OS=Homo sapiens GN=ORC3 PE=1 SV=1 - [ORC3_HUMAN]	1.69	2
F5H8D7	Forkhead box protein H1 OS=Homo sapiens GN=FOXH1 PE=4 SV=2 - [F5H8D7_HUMAN]	1.66	1
O15042	U2 snRNP-associated SURP motif-containing protein OS=Homo sapiens GN=U2SURP PE=1 SV=2 - [SR140_HUMAN]	1.65	2
D6RDY0	Tetrapeptide repeat protein 9A (Fragment) OS=Homo sapiens GN=TTTC9 PE=4 SV=3 - [D6RDY0_HUMAN]	1.65	1
Q15061	WD repeat-containing protein 43 OS=Homo sapiens GN=WDR43 PE=1 SV=3 - [WDR43_HUMAN]	1.62	2
Q8NFW8	N-acylneuraminate cytidylyltransferase OS=Homo sapiens GN=CMAS PE=1 SV=2 - [NEUA_HUMAN]	1.61	2
Q12788	Transducin beta-like protein 3 OS=Homo sapiens GN=TLB3 PE=1 SV=2 - [TLB3_HUMAN]	1.61	2
O43913	Origin recognition complex subunit 5 OS=Homo sapiens GN=ORC5 PE=1 SV=1 - [ORC5_HUMAN]	1.61	1
Q9UQR0	Sex comb on midleg-like protein 2 OS=Homo sapiens GN=SCML2 PE=1 SV=1 - [SCML2_HUMAN]	1.57	2
Q9HOA0	N-acetyltransferase 10 OS=Homo sapiens GN=NAT10 PE=1 SV=2 - [NAT10_HUMAN]	1.56	2
P49591	Serine-tRNA ligase, cytoplasmic OS=Homo sapiens GN=SARS PE=1 SV=3 - [SYSY_HUMAN]	1.56	1
Q8WYQ5	Microprocessor complex subunit DGCR8 OS=Homo sapiens GN=DGCR8 PE=1 SV=1 - [DGCR8_HUMAN]	1.55	3
O60716	Catenin delta-1 OS=Homo sapiens GN=CTNND1 PE=1 SV=1 - [CTND1_HUMAN]	1.55	2
O43143	Putative pre-mRNA-splicing factor ATP-dependent RNA helicase DHX15 OS=Homo sapiens GN=DHX15 PE=1 SV=2 - [DHX15_HUMAN]	1.51	2
O94776	Metastasis-associated protein MTA2 OS=Homo sapiens GN=MTA2 PE=1 SV=1 - [MTA2_HUMAN]	1.50	2
Q13573	SNW domain-containing protein 1 OS=Homo sapiens GN=SNW1 PE=1 SV=1 - [SNW1_HUMAN]	1.49	2

Accession	Description	ΣCoverage	Σ# PSMs
O43395	U4/U6 small nuclear ribonucleoprotein Prp3 OS=Homo sapiens GN=PRPF3 PE=1 SV=2 - [PRPF3_HUMAN]	1.46	1
P29401	Transketolase OS=Homo sapiens GN=TKT PE=1 SV=3 - [TKT_HUMAN]	1.44	1
Q96T37	Putative RNA-binding protein 15 OS=Homo sapiens GN=RBM15 PE=1 SV=2 - [RBM15_HUMAN]	1.43	2
Q9UHX1	Poly(U)-binding-splicing factor PUF60 OS=Homo sapiens GN=PUF60 PE=1 SV=1 - [PUF60_HUMAN]	1.43	1
Q9UHP1	Histone lysine demethylase PHF8 OS=Homo sapiens GN=PHF8 PE=1 SV=3 - [PHF8_HUMAN]	1.42	2
Q15459	Splicing factor 3A subunit 1 OS=Homo sapiens GN=SF3A1 PE=1 SV=1 - [SF3A1_HUMAN]	1.39	1
H7C1M2	Protein SON (Fragment) OS=Homo sapiens GN=SON PE=1 SV=1 - [H7C1M2_HUMAN]	1.38	2
P04843	Dolichyl-diphosphooligosaccharide-protein glycosyltransferase subunit 1 OS=Homo sapiens GN=RPN1 PE=1 SV=1 - [RPN1_HUMAN]	1.32	2
P27816	Microtubule-associated protein 4 OS=Homo sapiens GN=MAP4 PE=1 SV=3 - [MAP4_HUMAN]	1.30	2
Q8NE71	ATP-binding cassette sub-family F member 1 OS=Homo sapiens GN=ABCF1 PE=1 SV=2 - [ABCF1_HUMAN]	1.30	2
Q9H307	Pinin OS=Homo sapiens GN=PNN PE=1 SV=4 - [PININ_HUMAN]	1.26	2
Q9HD06	5'-3' exonuclease 2 OS=Homo sapiens GN=XRN2 PE=1 SV=1 - [XRN2_HUMAN]	1.26	1
O75400	Pre-mRNA-processing factor 40 homolog A OS=Homo sapiens GN=PRPF40A PE=1 SV=2 - [PR40A_HUMAN]	1.25	2
P49327	Fatty acid synthase OS=Homo sapiens GN=FASN PE=1 SV=3 - [FAS_HUMAN]	1.15	5
P02768	Serum albumin OS=Homo sapiens GN=ALB PE=1 SV=2 - [ALBU_HUMAN]	1.15	2
O00411	DNA-directed RNA polymerase, mitochondrial OS=Homo sapiens GN=POLRMT PE=1 SV=2 - [RPOM_HUMAN]	1.14	2
Q9UG63	ATP-binding cassette sub-family F member 2 OS=Homo sapiens GN=ABCF2 PE=1 SV=2 - [ABCF2_HUMAN]	1.12	2
O75152	Zinc finger CCH domain-containing protein 11A OS=Homo sapiens GN=ZC3H11A PE=1 SV=3 - [ZC11A_HUMAN]	1.11	2
P49916	DNA ligase 3 OS=Homo sapiens GN=LIG3 PE=1 SV=2 - [DNL3_HUMAN]	1.09	2
O14654	Insulin receptor substrate 4 OS=Homo sapiens GN=IRS4 PE=1 SV=1 - [IRS4_HUMAN]	1.03	1
Q63HK3	Zinc finger protein with KRAB and SCAN domains 2 OS=Homo sapiens GN=ZKSCAN2 PE=1 SV=2 - [ZKSC2_HUMAN]	1.03	1
Q8WVK9	Cytoskeleton-associated protein 2 OS=Homo sapiens GN=CKAP2 PE=1 SV=1 - [CKAP2_HUMAN]	1.02	1
P54886	Delta-1-pyrroline-5-carboxylate synthase OS=Homo sapiens GN=ALDH18A1 PE=1 SV=2 - [PSCS_HUMAN]	1.01	2
Q8WTT2	Nucleolar complex protein 3 homolog OS=Homo sapiens GN=NOC3L PE=1 SV=1 - [NOC3L_HUMAN]	1.00	1
Q15393	Splicing factor 3B subunit 3 OS=Homo sapiens GN=SF3B3 PE=1 SV=4 - [SF3B3_HUMAN]	0.99	2
P55197-1	Isoform 1 of Protein AF-10 OS=Homo sapiens GN=MLLT10 - [AF10_HUMAN]	0.97	1
Q14566	DNA replication licensing factor MCM6 OS=Homo sapiens GN=MCM6 PE=1 SV=1 - [MCM6_HUMAN]	0.97	2
Q1KMD3	Heterogeneous nuclear ribonucleoprotein U-like protein 2 OS=Homo sapiens GN=HNRNPUL2 PE=1 SV=1 - [HNRL2_HUMAN]	0.94	2
P55198	Protein AF-17 OS=Homo sapiens GN=MLLT6 PE=1 SV=2 - [AF17_HUMAN]	0.91	2
Q12769	Nuclear pore complex protein Nup160 OS=Homo sapiens GN=NUP160 PE=1 SV=3 - [NU160_HUMAN]	0.91	1
O60841	Eukaryotic translation initiation factor 5B OS=Homo sapiens GN=EIF5B PE=1 SV=4 - [IF2P_HUMAN]	0.90	2
A1X283	SH3 and FX domain-containing protein 2B OS=Homo sapiens GN=SH3PX2B PE=1 SV=3 - [SPD2B_HUMAN]	0.88	2
P49790	Nuclear pore complex protein Nup153 OS=Homo sapiens GN=NUP153 PE=1 SV=2 - [NU153_HUMAN]	0.88	2
Q13435	Splicing factor 3B subunit 2 OS=Homo sapiens GN=SF3B2 PE=1 SV=2 - [SF3B2_HUMAN]	0.78	2
Q92621	Nuclear pore complex protein Nup205 OS=Homo sapiens GN=NUP205 PE=1 SV=3 - [NU205_HUMAN]	0.75	1
Q8N2Y8	Iporin OS=Homo sapiens GN=RUSC2 PE=1 SV=3 - [RUSC2_HUMAN]	0.73	1
Q8N3U4	Cohesin subunit SA-2 OS=Homo sapiens GN=STAG2 PE=1 SV=3 - [STAG2_HUMAN]	0.73	1
O14776	Transcription elongation regulator 1 OS=Homo sapiens GN=TCERG1 PE=1 SV=2 - [TCRG1_HUMAN]	0.73	1
O60241	Brain-specific angiogenesis inhibitor 2 OS=Homo sapiens GN=BAI2 PE=2 SV=2 - [BAI2_HUMAN]	0.69	4
O15197	Ephrin type-B receptor 6 OS=Homo sapiens GN=EPHB6 PE=1 SV=4 - [EPHB6_HUMAN]	0.69	1
Q9NPP4	NLR family CARD domain-containing protein 4 OS=Homo sapiens GN=NLR4 PE=1 SV=2 - [NLR4_HUMAN]	0.68	1
Q9UQE7	Structural maintenance of chromosomes protein 3 OS=Homo sapiens GN=SMC3 PE=1 SV=2 - [SMC3_HUMAN]	0.66	2
Q9BZH6	WD repeat-containing protein 11 OS=Homo sapiens GN=WDR11 PE=1 SV=1 - [WDR11_HUMAN]	0.65	1
Q8WXE0	Caskin-2 OS=Homo sapiens GN=CASKIN2 PE=1 SV=2 - [CSK12_HUMAN]	0.58	2
Q5W79	Centrosomal protein of 170 kDa OS=Homo sapiens GN=CEP170 PE=1 SV=1 - [CE170_HUMAN]	0.57	2
Q8N127	THO complex subunit 2 OS=Homo sapiens GN=THOC2 PE=1 SV=2 - [THOC2_HUMAN]	0.56	2
Q9Y490	Talin-1 OS=Homo sapiens GN=TLN1 PE=1 SV=3 - [TLN1_HUMAN]	0.55	2
Q00610	Clastrin heavy chain 1 OS=Homo sapiens GN=CLTC PE=1 SV=5 - [CLH1_HUMAN]	0.54	1
Q14690	Protein RRP5 homolog OS=Homo sapiens GN=PDCC11 PE=1 SV=3 - [RRP5_HUMAN]	0.53	2
P51532	Transcription activator BRG1 OS=Homo sapiens GN=SMARCA4 PE=1 SV=2 - [SMCA4_HUMAN]	0.49	2
Q9HCJ0	Trinucleotide repeat-containing gene 6C protein OS=Homo sapiens GN=TNRC6C PE=1 SV=3 - [TNRC6C_HUMAN]	0.47	1
Q13813	Spectrin alpha chain, non-erythrocytic 1 OS=Homo sapiens GN=SPTAN1 PE=1 SV=3 - [SPTN1_HUMAN]	0.44	2
Q6P2Q9	Pre-mRNA-processing-splicing factor 8 OS=Homo sapiens GN=PRPF8 PE=1 SV=2 - [PRP8_HUMAN]	0.34	2
Q5UJP0	Telomere-associated protein RIF1 OS=Homo sapiens GN=RIF1 PE=1 SV=2 - [RIF1_HUMAN]	0.28	1
Q15149	Plectin OS=Homo sapiens GN=PLEC PE=1 SV=3 - [PLEC_HUMAN]	0.15	2

## APPENDIX-III



Deniz Ugurlu Cimen &lt;denizugurlu88@gmail.com&gt;

**Re: CSHL Press Reprint Permission Request Form - G&D, 2018 -> PhD Thesis**

3 iletii

**Mazzullo, Mala** <mazzullo@cshl.edu>  
Alici: "dugurlu13@ku.edu.tr" <dugurlu13@ku.edu.tr>

28 Ağustos 2018 21:46

Dear Deniz,

Thank you for submitting your permission request for the use of the material described below in your PhD thesis. As noted in the article:

© 2018 Zhang et al.; Published by Cold Spring Harbor Laboratory Press<<http://genesdev.cshlp.org/site/misc/terms.xhtml>>

This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see <http://genesdev.cshlp.org/site/misc/terms.xhtml>). After six months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

Although the material you are seeking permission for is copyrighted by the article authors, Zhang et al., we are within the first six months following full-issue publication. For the non-commercial purpose of your thesis, we are pleased to grant you permission to use Figure 1D and 2B. Please do, however, be sure to include the complete reference to the original Genes & Development article and indicate if any changes are made.

Best wishes for success with your thesis!

Kind regards,  
Mala  
Mala Shwe Mazzullo  
Executive Assistant to the Publisher  
Cold Spring Harbor Laboratory Press  
One Bungtown Road, Cold Spring Harbor, NY 11724  
Tel: 516.422.4005 / [www.cshlpress.org](http://www.cshlpress.org)

---

## REFERENCES

1. Gurdon, J. B., Laskey, R. a & Reeves, O. R. The developmental capacity of nuclei transplanted from keratinized skin cells of adult frogs. *J. Embryol. Exp. Morphol.* **34**, 93–112 (1975).
2. Stevens, L. & Little, S. Spontaneous testicular tumors in an inbred strain of mice. *PNAS* **40**, 1080–1087 (1954).
3. Kahan, B. W. & Ephrussi, B. Developmental potentialities of clonal in vitro cultures of mouse testicular teratoma. *J. Natl. Cancer Inst.* **44**, 1015–1036 (1970).
4. Davis, R. L., Weintraub, H. & Lassar, A. B. Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell* **51**, 987–1000 (1987).
5. Takahashi, K. & Yamanaka, S. Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell* **126**, 663–676 (2006).
6. Takahashi, K. *et al.* Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861–872 (2007).
7. Hanna, J. H., Saha, K. & Jaenisch, R. Pluripotency and Cellular Reprogramming: Facts, Hypotheses, Unresolved Issues. *Cell* **143**, 508–525 (2010).
8. De Los Angeles, A. *et al.* Hallmarks of pluripotency. *Nature* **525**, 469–478 (2015).
9. Nichols, J. & Smith, A. Pluripotency in the embryo and in culture. *Cold Spring Harb. Perspect. Biol.* **4**, a008128 (2012).
10. Guo, G. *et al.* Naive Pluripotent Stem Cells Derived Directly from Isolated Cells of the Human Inner Cell Mass. *Stem Cell Reports* **6**, 437–446 (2016).
11. Kilens, S. *et al.* Parallel derivation of isogenic human primed and naive induced pluripotent stem cells. *Nat. Commun.* **9**, 360 (2018).

12. Kisa, F. *et al.* Naive-like ESRRB+ iPSCs with the Capacity for Rapid Neural Differentiation. *Stem cell reports* **9**, 1825–1838 (2017).
13. Doi, A. *et al.* Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. *Nat. Genet.* **41**, 1350–1353 (2009).
14. Deng, J. *et al.* Targeted bisulfite sequencing reveals changes in DNA methylation associated with nuclear reprogramming. *Nat. Biotechnol.* **27**, 353–360 (2009).
15. Chin, M. H. *et al.* Induced Pluripotent Stem Cells and Embryonic Stem Cells Are Distinguished by Gene Expression Signatures. *Cell Stem Cell* **5**, 111–123 (2009).
16. Polo, J. M. *et al.* Cell type of origin influences the molecular and functional properties of mouse induced pluripotent stem cells. *Nat. Biotechnol.* **28**, 848–855 (2010).
17. Tomoda, K. *et al.* Derivation Conditions Impact X-Inactivation Status in Female Human Induced Pluripotent Stem Cells. *Cell Stem Cell* **11**, 91–99 (2012).
18. Guo, H. *et al.* The DNA methylation landscape of human early embryos. *Nature* **511**, 606–610 (2014).
19. Smith, Z. D. *et al.* DNA methylation dynamics of the human preimplantation embryo. *Nature* **511**, 611–615 (2014).
20. Takahashi, K. & Yamanaka, S. A developmental framework for induced pluripotency. doi:10.1242/dev.114249
21. Boland, M. J. *et al.* Adult mice generated from induced pluripotent stem cells. *Nature* **461**, 91–94 (2009).
22. Kang, L., Wang, J., Zhang, Y., Kou, Z. & Gao, S. iPS Cells Can Support Full-Term Development of Tetraploid Blastocyst-Complemented Embryos. *Cell Stem Cell* **5**, 135–138 (2009).
23. Zhao, X. *et al.* iPS cells produce viable mice through tetraploid complementation. *Nature* **461**, 86–90 (2009).

24. Liu, H. *et al.* Generation of Induced Pluripotent Stem Cells from Adult Rhesus Monkey Fibroblasts. *Cell Stem Cell* **3**, 587–590 (2008).
25. Liao, J. *et al.* Generation of Induced Pluripotent Stem Cell Lines from Adult Rat Cells. *Cell Stem Cell* **4**, 11–15 (2009).
26. Esteban, M. A. *et al.* Generation of Induced Pluripotent Stem Cell Lines from Tibetan Miniature Pig. *J. Biol. Chem.* **284**, 17634–17640 (2009).
27. Ogorevc, J., Orehek, S. & Dovč, P. Cellular reprogramming in farm animals: an overview of iPSC generation in the mammalian farm animal species. *J. Anim. Sci. Biotechnol.* **7**, 10 (2016).
28. Stadtfeld, M. & Hochedlinger, K. Induced pluripotency: history, mechanisms, and applications. *Genes Dev.* **24**, 2239–2263 (2010).
29. Nakagawa, M., Takizawa, N., Narita, M., Ichisaka, T. & Yamanaka, S. Promotion of direct reprogramming by transformation-deficient Myc. *Proc. Natl. Acad. Sci.* **107**, 14152–14157 (2010).
30. Han, J. *et al.* Tbx3 improves the germ-line competency of induced pluripotent stem cells. *Nature* **463**, 1096–1100 (2010).
31. Maekawa, M. *et al.* Direct reprogramming of somatic cells is promoted by maternal transcription factor Glis1. *Nature* **474**, 225–229 (2011).
32. Nakagawa, M. *et al.* Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. *Nat. Biotechnol.* **26**, 101–106 (2008).
33. Worringer, K. A. *et al.* The let-7/LIN-41 Pathway Regulates Reprogramming to Human Induced Pluripotent Stem Cells by Controlling Expression of Prodifferentiation Genes. *Cell Stem Cell* **14**, 40–52 (2014).
34. Yang, P. *et al.* RCOR2 Is a Subunit of the LSD1 Complex That Regulates ESC Property and Substitutes for SOX2 in Reprogramming Somatic Cells to Pluripotency. *Stem Cells* **29**, 791–801 (2011).
35. Feng, B. *et al.* Reprogramming of fibroblasts into induced pluripotent stem cells with orphan nuclear receptor Esrrb. *Nat. Cell Biol.* **11**, 197–203 (2009).



36. Onder, T. T. *et al.* Chromatin-modifying enzymes as modulators of reprogramming. *Nature* **483**, 598–602 (2012).
37. Yu, J. *et al.* Induced Pluripotent Stem Cell Lines Derived from Human Somatic Cells. *Science (80-. )*. **318**, 1917–1920 (2007).
38. Shinagawa, T. *et al.* Histone Variants Enriched in Oocytes Enhance Reprogramming to Induced Pluripotent Stem Cells. *Cell Stem Cell* **14**, 217–227 (2014).
39. Shu, J. *et al.* Induction of Pluripotency in Mouse Somatic Cells with Lineage Specifiers. *Cell* **153**, 963–975 (2013).
40. Montserrat, N. *et al.* Reprogramming of Human Fibroblasts to Pluripotency with Lineage Specifiers. *Cell Stem Cell* **13**, 341–350 (2013).
41. Heng, J.-C. D. *et al.* The nuclear receptor Nr5a2 can replace Oct4 in the reprogramming of murine somatic cells to pluripotent cells. *Cell Stem Cell* **6**, 167–74 (2010).
42. Buganim, Y. *et al.* Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase. *Cell* **150**, 1209–22 (2012).
43. Redmer, T. *et al.* E-cadherin is crucial for embryonic stem cell pluripotency and can replace OCT4 during somatic cell reprogramming. *EMBO Rep.* **12**, 720–726 (2011).
44. Gao, Y. *et al.* Replacement of Oct4 by Tet1 during iPSC Induction Reveals an Important Role of DNA Methylation and Hydroxymethylation in Reprogramming. *Cell Stem Cell* **12**, 453–469 (2013).
45. Mai, T. *et al.* NKX3-1 is required for induced pluripotent stem cell reprogramming and can replace OCT4 in mouse and human iPSC induction. *Nat. Cell Biol.* **20**, 900–908 (2018).
46. Miyoshi, N. *et al.* Reprogramming of Mouse and Human Cells to Pluripotency Using Mature MicroRNAs. *Cell Stem Cell* **8**, 633–638 (2011).
47. Anokye-Danso, F. *et al.* Highly Efficient miRNA-Mediated

- 
- Reprogramming of Mouse and Human Somatic Cells to Pluripotency. *Cell Stem Cell* **8**, 376–388 (2011).
48. Theunissen, T. W. & Jaenisch, R. Molecular Control of Induced Pluripotency. *Cell Stem Cell* **14**, 720–734 (2014).
  49. Esteban, M. A. *et al.* Vitamin C Enhances the Generation of Mouse and Human Induced Pluripotent Stem Cells. *Cell Stem Cell* **6**, 71–79 (2010).
  50. Kawamura, T. *et al.* Linking the p53 tumour suppressor pathway to somatic cell reprogramming. *Nature* **460**, 1140–1144 (2009).
  51. Menendez, S., Camus, S. & Belmonte, J. C. I. p53: Guardian of reprogramming. *Cell Cycle* **9**, 3887–3891 (2010).
  52. Hou, P. *et al.* Pluripotent Stem Cells Induced from Mouse Somatic Cells by Small-Molecule Compounds. *Science (80-. )*. **341**, 651–654 (2013).
  53. Li, Y. *et al.* Generation of iPSCs from mouse fibroblasts with a single gene, Oct4 and small molecules. *Cell Res.* **21**, 196–204 (2011).
  54. Zhao, Y. *et al.* A XEN-like State Bridges Somatic Cells to Pluripotency during Chemical Reprogramming. *Cell* **163**, 1678–1691 (2015).
  55. Onder, T. T. *et al.* Chromatin-modifying enzymes as modulators of reprogramming. *Nature* **483**, 598–602 (2012).
  56. Zhu, S. *et al.* Reprogramming of Human Primary Somatic Cells by OCT4 and Chemical Compounds. *Cell Stem Cell* **7**, 651–655 (2010).
  57. Zhou, W. & Freed, C. R. Adenoviral Gene Delivery Can Reprogram Human Fibroblasts to Induced Pluripotent Stem Cells. *Stem Cells* **27**, 2667–2674 (2009).
  58. Fusaki, N., Ban, H., Nishiyama, A., Saeki, K. & Hasegawa, M. Efficient induction of transgene-free human pluripotent stem cells using a vector based on Sendai virus, an RNA virus that does not integrate into the host genome. *Proc. Jpn. Acad. Ser. B. Phys. Biol. Sci.* **85**, 348–62 (2009).
  59. Yu, J., Chau, K. F., Vodyanik, M. A., Jiang, J. & Jiang, Y. Efficient

- 
- Feeder-Free Episomal Reprogramming with Small Molecules. *PLoS One* **6**, e17557 (2011).
60. Warren, L. *et al.* Highly Efficient Reprogramming to Pluripotency and Directed Differentiation of Human Cells with Synthetic Modified mRNA. *Cell Stem Cell* **7**, 618–630 (2010).
61. Bosnali, M. & Edenhofer, F. Generation of transducible versions of transcription factors Oct4 and Sox2. *Biol. Chem.* **389**, 851–61 (2008).
62. González, F., Boué, S. & Belmonte, J. C. I. Methods for making induced pluripotent stem cells: reprogramming à la carte. *Nat. Rev. Genet.* **12**, 231–242 (2011).
63. Samavarchi-Tehrani, P. *et al.* Functional Genomics Reveals a BMP-Driven Mesenchymal-to-Epithelial Transition in the Initiation of Somatic Cell Reprogramming. *Cell Stem Cell* **7**, 64–77 (2010).
64. David, L. & Polo, J. M. Phases of reprogramming. *Stem Cell Res.* **12**, 754–61 (2014).
65. Yoshida, Y., Takahashi, K., Okita, K., Ichisaka, T. & Yamanaka, S. Hypoxia Enhances the Generation of Induced Pluripotent Stem Cells. *Cell Stem Cell* **5**, 237–241 (2009).
66. Tanabe, K., Nakamura, M., Narita, M., Takahashi, K. & Yamanaka, S. Maturation, not initiation, is the major roadblock during reprogramming toward pluripotency from human fibroblasts. *Proc. Natl. Acad. Sci.* **110**, 12172–12179 (2013).
67. Stadtfeld, M., Maherali, N., Breault, D. T. & Hochedlinger, K. Defining Molecular Cornerstones during Fibroblast to iPS Cell Reprogramming in Mouse. *Cell Stem Cell* **2**, 230–240 (2008).
68. Maherali, N. *et al.* Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. *Cell Stem Cell* **1**, 55–70 (2007).
69. Wernig, M. *et al.* In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature* **448**, 318–324 (2007).
70. Mikkelsen, T. S. *et al.* Dissecting direct reprogramming through integrative genomic analysis. *Nature* **454**, 49–55 (2008).

71. Cahan, P. *et al.* CellNet: Network Biology Applied to Stem Cell Engineering. *Cell* **158**, 903–915 (2014).
72. Hill, P. W. S., Amouroux, R. & Hajkova, P. DNA demethylation, Tet proteins and 5-hydroxymethylcytosine in epigenetic reprogramming: an emerging complex story. *Genomics* **104**, 324–33 (2014).
73. Huangfu, D. *et al.* Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. *Nat. Biotechnol.* **26**, 795–7 (2008).
74. Bernstein, B. E. *et al.* A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells. *Cell* **125**, 315–326 (2006).
75. Guenther, M. G., Levine, S. S., Boyer, L. A., Jaenisch, R. & Young, R. A. A Chromatin Landmark and Transcription Initiation at Most Promoters in Human Cells. *Cell* **130**, 77–88 (2007).
76. Ang, Y.-S. *et al.* Wdr5 Mediates Self-Renewal and Reprogramming via the Embryonic Stem Cell Core Transcriptional Network. *Cell* **145**, 183–197 (2011).
77. Jiang, H. *et al.* Role for Dpy-30 in ES Cell-Fate Specification by Regulation of H3K4 Methylation within Bivalent Domains. *Cell* **144**, 513–525 (2011).
78. Adamo, A. *et al.* LSD1 regulates the balance between self-renewal and differentiation in human embryonic stem cells. *Nat. Cell Biol.* **13**, 652–659 (2011).
79. Xie, L. *et al.* KDM5B regulates embryonic stem cell self-renewal and represses cryptic intragenic transcription. *EMBO J.* **30**, 1473–1484 (2011).
80. Loh, Y.-H., Zhang, W., Chen, X., George, J. & Ng, H.-H. Jmjd1a and Jmjd2c histone H3 Lys 9 demethylases regulate self-renewal in embryonic stem cells. *Genes Dev.* **21**, 2545–2557 (2007).
81. Zhang, Z. *et al.* PRC2 Complexes with JARID2, MTF2, and esPRC2p48 in ES Cells to Modulate ES Cell Pluripotency and Somatic Cell Reprograming. *Stem Cells* **29**, 229–240 (2011).

- 
82. Zhao, W. *et al.* Jmjd3 Inhibits Reprogramming by Upregulating Expression of INK4a/Arf and Targeting PHF20 for Ubiquitination. *Cell* **152**, 1037–1050 (2013).
  83. Mansour, A. A. *et al.* The H3K27 demethylase Utx regulates somatic and germ cell epigenetic reprogramming. *Nature* **488**, 409–413 (2012).
  84. Wang, T. *et al.* The Histone Demethylases Jhdm1a/1b Enhance Somatic Cell Reprogramming in a Vitamin-C-Dependent Manner. *Cell Stem Cell* **9**, 575–587 (2011).
  85. Singhal, N. *et al.* Chromatin-Remodeling Components of the BAF Complex Facilitate Reprogramming. *Cell* **141**, 943–955 (2010).
  86. Hu, X. *et al.* Tet and TDG Mediate DNA Demethylation Essential for Mesenchymal-to-Epithelial Transition in Somatic Cell Reprogramming. *Cell Stem Cell* **14**, 512–522 (2014).
  87. Chen, J. J. *et al.* H3K9 methylation is a barrier during somatic cell reprogramming into iPSCs. *Nat. Genet.* **45**, 34–42 (2013).
  88. Gaspar-Maia, A. *et al.* MacroH2A histone variants act as a barrier upon reprogramming towards pluripotency. *Nat. Commun.* **4**, 1565 (2013).
  89. Xu, Y. *et al.* Transcriptional Control of Somatic Cell Reprogramming. *Trends Cell Biol.* **26**, 272–288 (2016).
  90. Barry, E. R. *et al.* ES Cell Cycle Progression and Differentiation Require the Action of the Histone Methyltransferase Dot1L. *Stem Cells* **27**, 1538–1547 (2009).
  91. Min, J., Feng, Q., Li, Z., Zhang, Y. & Xu, R. M. Structure of the catalytic domain of human Dot1L, a non-SET domain nucleosomal histone methyltransferase. *Cell* **112**, 711–723 (2003).
  92. Singer, M. S. *et al.* Identification of high-copy disruptors of telomeric silencing in *Saccharomyces cerevisiae*. *Genetics* **150**, 613–632 (1998).
  93. Nguyen, A. T. & Zhang, Y. The diverse functions of Dot1 and H3K79 methylation. *Genes and Development* **25**, 1345–1358 (2011).
  94. Feng, Q. *et al.* Methylation of H3-lysine 79 is mediated by a new family

- of HMTases without a SET domain. *Curr. Biol.* **12**, 1052–1058 (2002).
95. Dlakić, M. Chromatin silencing protein and pachytene checkpoint regulator Dot1p has a methyltransferase fold. *Trends Biochem. Sci.* **26**, 405–7 (2001).
  96. Sun, Z.-W. & Allis, C. D. Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. *Nature* **418**, 104–108 (2002).
  97. McGinty, R. K., Kim, J., Chatterjee, C., Roeder, R. G. & Muir, T. W. Chemically ubiquitylated histone H2B stimulates hDot1L-mediated intranucleosomal methylation. *Nature* **453**, 812–816 (2008).
  98. Lacoste, N., Utley, R. T., Hunter, J. M., Poirier, G. G. & Côté, J. Disruptor of Telomeric Silencing-1 Is a Chromatin-specific Histone H3 Methyltransferase. *J. Biol. Chem.* **277**, 30421–30424 (2002).
  99. Ng, H. H. *et al.* Lysine methylation within the globular domain of histone H3 by Dot1 is important for telomeric silencing and Sir protein association. *Genes Dev.* **16**, 1518–1527 (2002).
  100. van Leeuwen, F., Gafken, P. R. & Gottschling, D. E. Dot1p modulates silencing in yeast by methylation of the nucleosome core. *Cell* **109**, 745–56 (2002).
  101. Katan-Khaykovich, Y. & Struhl, K. Heterochromatin formation involves changes in histone modifications over multiple cell generations. *EMBO J.* **24**, 2138–2149 (2005).
  102. Chen, C.-W. *et al.* DOT1L inhibits SIRT1-mediated epigenetic silencing to maintain leukemic gene expression in MLL-rearranged leukemia. *Nat. Med.* **21**, 335–343 (2015).
  103. Bitoun, E., Oliver, P. L. & Davies, K. E. The mixed-lineage leukemia fusion partner AF4 stimulates RNA polymerase II transcriptional elongation and mediates coordinated chromatin remodeling. *Hum. Mol. Genet.* **16**, 92–106 (2007).
  104. Mueller, D. *et al.* Misguided Transcriptional Elongation Causes Mixed Lineage Leukemia. *PLoS Biol.* **7**, e1000249 (2009).
  105. Mohan, M. *et al.* Linking H3K79 trimethylation to Wnt signaling through a novel Dot1-containing complex (DotCom). *Genes Dev.* **24**,

- 574–589 (2010).
106. Lin, C. *et al.* AFF4, a Component of the ELL/P-TEFb Elongation Complex and a Shared Subunit of MLL Chimeras, Can Link Transcription Elongation to Leukemia. *Mol. Cell* **37**, 429–437 (2010).
  107. Yokoyama, A., Lin, M., Naresh, A., Kitabayashi, I. & Cleary, M. L. A Higher-Order Complex Containing AF4 and ENL Family Proteins with P-TEFb Facilitates Oncogenic and Physiologic MLL-Dependent Transcription. *Cancer Cell* **17**, 198–212 (2010).
  108. Huyen, Y. *et al.* Methylated lysine 79 of histone H3 targets 53BP1 to DNA double-strand breaks. *Nature* **432**, 406–411 (2004).
  109. Ward, I. M., Minn, K., van Deursen, J. & Chen, J. p53 Binding protein 53BP1 is required for DNA damage responses and tumor suppression in mice. *Mol. Cell. Biol.* **23**, 2556–63 (2003).
  110. Wysocki, R. *et al.* Role of Dot1-Dependent Histone H3 Methylation in G1 and S Phase DNA Damage Checkpoint Functions of Rad9. *Mol. Cell. Biol.* **25**, 8430–8443 (2005).
  111. Lazzaro, F. *et al.* Histone methyltransferase Dot1 and Rad9 inhibit single-stranded DNA accumulation at DSBs and uncapped telomeres. *EMBO J.* **27**, 1502–12 (2008).
  112. Bostelman, L. J., Keller, A. M., Albrecht, A. M., Arat, A. & Thompson, J. S. Methylation of histone H3 lysine-79 by Dot1p plays multiple roles in the response to UV damage in *Saccharomyces cerevisiae*. *DNA Repair (Amst)*. **6**, 383–395 (2007).
  113. Wakeman, T. P., Wang, Q., Feng, J. & Wang, X.-F. Bat3 facilitates H3K79 dimethylation by DOT1L and promotes DNA damage-induced 53BP1 foci at G1/G2 cell-cycle phases. *EMBO J.* **31**, 2169–2181 (2012).
  114. Fu, H. *et al.* Methylation of Histone H3 on Lysine 79 Associates with a Group of Replication Origins and Helps Limit DNA Replication Once per Cell Cycle. *PLoS Genet.* **9**, e1003542 (2013).
  115. Schulze, J. M. *et al.* Linking Cell Cycle to Histone Modifications: SBF and H2B Monoubiquitination Machinery and Cell-Cycle Regulation of H3K79 Dimethylation. *Mol. Cell* **35**, 626–641 (2009).

116. Feng, Y. *et al.* Early mammalian erythropoiesis requires the Dot1L methyltransferase. *Blood* **116**, 4483–4491 (2010).
117. Shanower, G. A. *et al.* Characterization of the grappa Gene, the Drosophila Histone H3 Lysine 79 Methyltransferase. *Genetics* **169**, 173–184 (2004).
118. Ooga, M. *et al.* Changes in H3K79 Methylation During Preimplantation Development in Mice. *Biol. Reprod.* **78**, 413–424 (2008).
119. Jones, B. *et al.* The histone H3K79 methyltransferase Dot1L is essential for mammalian development and heterochromatin structure. *PLoS Genet.* **4**, (2008).
120. Jbara, M., Guttmann-Raviv, N., Maity, S. K., Ayoub, N. & Brik, A. Total chemical synthesis of methylated analogues of histone 3 revealed KDM4D as a potential regulator of H3K79me3. *Bioorg. Med. Chem.* **25**, 4966–4970 (2017).
121. Okada, Y. *et al.* hDOT1L links histone methylation to leukemogenesis. *Cell* **121**, 167–178 (2005).
122. Okada, Y. *et al.* Leukaemic transformation by CALM–AF10 involves upregulation of Hoxa5 by hDOT1L. *Nat. Cell Biol.* **8**, 1017–1024 (2006).
123. Chen, C.-W. & Armstrong, S. A. Targeting DOT1L and HOX gene expression in MLL-rearranged leukemia and beyond. *Exp. Hematol.* **43**, 673–84 (2015).
124. McLean, C. M., Karemaker, I. D. & van Leeuwen, F. The emerging roles of DOT1L in leukemia and normal development. *Leukemia* **28**, 2131–2138 (2014).
125. Park, G., Gong, Z., Chen, J. & Kim, J. E. Characterization of the DOT1L network: Implications of diverse roles for DOT1L. *Protein J.* **29**, 213–223 (2010).
126. Shah, S. & Henriksen, M. A. A Novel Disrupter of Telomere Silencing 1-like (DOT1L) Interaction Is Required for Signal Transducer and Activator of Transcription 1 (STAT1)-activated Gene Expression. *J. Biol. Chem.* **286**, 41195–41204 (2011).



- 
127. Benedikt, A. *et al.* The leukemogenic AF4–MLL fusion protein causes P-TEFb kinase activation and altered epigenetic signatures. *Leukemia* **25**, 135–144 (2011).
128. Mueller, D. *et al.* A role for the MLL fusion partner ENL in transcriptional elongation and chromatin modification. *Blood* **110**, 4445–54 (2007).
129. Castelli, G., Pelosi, E. & Testa, U. Targeting histone methyltransferase and demethylase in acute myeloid leukemia therapy. *Onco. Targets. Ther.* **11**, 131–155 (2018).
130. Chen, S. *et al.* The PZP Domain of AF10 Senses Unmodified H3K27 to Regulate DOT1L-Mediated Methylation of H3K79. *Mol. Cell* **60**, 319–327 (2015).
131. Sanjana, N. E., Shalem, O. & Zhang, F. Improved vectors and genome-wide libraries for CRISPR screening. *Nat. Methods* **11**, 783–784 (2014).
132. Park, I.-H. *et al.* Reprogramming of human somatic cells to pluripotency with defined factors. *Nature* **451**, 141–146 (2008).
133. Firat-Karalar, E. N., Rauniyar, N., Yates, J. R. & Stearns, T. Proximity interactions among centrosome components identify regulators of centriole duplication. *Curr. Biol.* **24**, 664–670 (2014).
134. Özkan Küçük, N. E., Şanal, E., Tan, E., Mitchison, T. & Özlü, N. Labeling Carboxyl Groups of Surface-Exposed Proteins Provides an Orthogonal Approach for Cell Surface Isolation. *J. Proteome Res.* **17**, 1784–1793 (2018).
135. Subramanian, A. *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci.* **102**, 15545–15550 (2005).
136. Wang, P. *et al.* Global analysis of H3K4 methylation defines MLL family member targets and points to a role for MLL1-mediated H3K4 methylation in the regulation of transcriptional initiation by RNA polymerase II. *Mol. Cell. Biol.* **29**, 6074–85 (2009).
137. Wang, J. *et al.* A protein interaction network for pluripotency of embryonic stem cells. *Nature* **444**, 364–368 (2006).

- 
138. Dundas, C. M., Demonte, D. & Park, S. Streptavidin–biotin technology: improvements and innovations in chemical and biological applications. *Appl. Microbiol. Biotechnol.* **97**, 9343–9353 (2013).
139. Roux, K. J., Kim, D. I., Raida, M. & Burke, B. A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *J. Cell Biol.* **196**, 801–810 (2012).
140. Kim, D. I. *et al.* An improved smaller biotin ligase for BioID proximity labeling. *Mol. Biol. Cell* **27**, 1188–96 (2016).
141. Li, P., Li, J., Wang, L. & Di, L.-J. Proximity Labeling of Interacting Proteins: Application of BioID as a Discovery Tool. *Proteomics* **17**, 1700002 (2017).
142. Han, S. *et al.* Proximity Biotinylation as a Method for Mapping Proteins Associated with mtDNA in Living Cells. *Cell Chem. Biol.* **24**, 404–414 (2017).
143. Branon, T. C. *et al.* Directed evolution of TurboID for efficient proximity labeling in living cells and organisms. *bioRxiv* 196980 (2017). doi:10.1101/196980
144. Morriswood, B. *et al.* Novel Bilobe Components in *Trypanosoma brucei* Identified Using Proximity-Dependent Biotinylation. *Eukaryot. Cell* **12**, 356–367 (2013).
145. Chen, S. *et al.* The PZP Domain of AF10 Senses Unmodified H3K27 to Regulate DOT1L-Mediated Methylation of H3K79. *Mol. Cell* **60**, 319–327 (2015).
146. Chen, L. *et al.* Abrogation of MLL–AF10 and CALM–AF10-mediated transformation through genetic inactivation or pharmacological inhibition of the H3K79 methyltransferase Dot1l. *Leukemia* **27**, 813–822 (2013).
147. Zhang, H. *et al.* Structural and functional analysis of the DOT1L–AF10 complex reveals mechanistic insights into MLL–AF10-associated leukemogenesis. *Genes Dev.* **32**, 341–346 (2018).
148. Ogoh, H. *et al.* Mllt10 knockout mouse model reveals critical role of Af10-dependent H3K79 methylation in midfacial development. *Sci. Rep.* **7**, 11922 (2017).

- 
149. Deshpande, A. J. *et al.* AF10 Regulates Progressive H3K79 Methylation and HOX Gene Expression in Diverse AML Subtypes. *Cancer Cell* **26**, 896–908 (2014).
  150. DiMartino, J. F. *et al.* The AF10 leucine zipper is required for leukemic transformation of myeloid progenitors by MLL-AF10. *Blood* **99**, 3780–5 (2002).
  151. Suzuki, M. *et al.* Nuclear export signal within CALM is necessary for CALM-AF10-induced leukemia. *Cancer Sci.* **105**, 315–323 (2014).
  152. Conway, A. E., Scotland, P. B., Lavau, C. P. & Wechsler, D. S. A CALM-derived nuclear export signal is essential for CALM-AF10-mediated leukemogenesis. *Blood* **121**, 4758–4768 (2013).
  153. Cecere, G., Hoersch, S., Jensen, M. B., Dixit, S. & Grishok, A. The ZFP-1(AF10)/DOT-1 Complex Opposes H2B Ubiquitination to Reduce Pol II Transcription. *Mol. Cell* **50**, 894–907 (2013).
  154. Okada, Y. *et al.* hDOT1L Links Histone Methylation to Leukemogenesis. *Cell* **121**, 167–178 (2005).
  155. Ziemer-van der Poel, S. *et al.* Identification of a gene, MLL, that spans the breakpoint in 11q23 translocations associated with human leukemias. *Proc. Natl. Acad. Sci. U. S. A.* **88**, 10735–9 (1991).
  156. Yagi, H. *et al.* Growth disturbance in fetal liver hematopoiesis of Mll-mutant mice. *Blood* **92**, 108–17 (1998).
  157. Milne, T. A. *et al.* MLL targets SET domain methyltransferase activity to Hox gene promoters. *Mol. Cell* **10**, 1107–17 (2002).
  158. Takeda, S. *et al.* Proteolysis of MLL family proteins is essential for taspase1-orchestrated cell cycle progression. *Genes Dev.* **20**, 2397–409 (2006).
  159. Yamashita, M. *et al.* Crucial Role of MLL for the Maintenance of Memory T Helper Type 2 Cell Responses. *Immunity* **24**, 611–622 (2006).
  160. Krivtsov, A. V. & Armstrong, S. A. MLL translocations, histone modifications and leukaemia stem-cell development. *Nat. Rev. Cancer* **7**, 823–833 (2007).

- 
161. Meyer, C. *et al.* New insights to the MLL recombinome of acute leukemias. *Leukemia* **23**, 1490–1499 (2009).
162. Milne, T. A., Martin, M. E., Brock, H. W., Slany, R. K. & Hess, J. L. Leukemogenic MLL fusion proteins bind across a broad region of the Hox a9 locus, promoting transcription and multiple histone modifications. *Cancer Res.* **65**, 11367–74 (2005).
163. Bernt, K. M. *et al.* MLL-Rearranged Leukemia Is Dependent on Aberrant H3K79 Methylation by DOT1L. *Cancer Cell* **20**, 66–78 (2011).
164. Gu, B. & Lee, M. G. Histone H3 lysine 4 methyltransferases and demethylases in self-renewal and differentiation of stem cells. *Cell Biosci.* **3**, 39 (2013).
165. Hsieh, J. J.-D., Cheng, E. H.-Y. & Korsmeyer, S. J. Taspase1: a threonine aspartase required for cleavage of MLL and proper HOX gene expression. *Cell* **115**, 293–303 (2003).
166. Hsieh, J. J.-D., Ernst, P., Erdjument-Bromage, H., Tempst, P. & Korsmeyer, S. J. Proteolytic cleavage of MLL generates a complex of N- and C-terminal fragments that confers protein stability and subnuclear localization. *Mol. Cell. Biol.* **23**, 186–94 (2003).
167. Dharmarajan, V. & S., M. Biochemistry of the Mixed Lineage Leukemia 1 (MLL1) Protein and Targeted Therapies for Associated Leukemia. in *Acute Leukemia - The Scientist's Perspective and Challenge* (InTech, 2011). doi:10.5772/19754
168. Armstrong, S. A. *et al.* MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia. *Nat. Genet.* **30**, 41–47 (2001).
169. Dou, Y. *et al.* Regulation of MLL1 H3K4 methyltransferase activity by its core components. *Nat. Struct. Mol. Biol.* **13**, 713–719 (2006).
170. Li, Y. *et al.* Structural basis for activity regulation of MLL family methyltransferases. *Nature* **530**, 447–452 (2016).
171. Shilatifard, A. The COMPASS Family of Histone H3K4 Methylases: Mechanisms of Regulation in Development and Disease Pathogenesis. *Annu. Rev. Biochem.* **81**, 65–95 (2012).

- 
172. Yokoyama, A. *et al.* The Menin Tumor Suppressor Protein Is an Essential Oncogenic Cofactor for MLL-Associated Leukemogenesis. *Cell* **123**, 207–218 (2005).
173. Koche, R. P. *et al.* Reprogramming Factor Expression Initiates Widespread Targeted Chromatin Remodeling. *Cell Stem Cell* **8**, 96–105 (2011).
174. Sindhu, C., Samavarchi-Tehrani, P. & Meissner, A. Transcription Factor-mediated Epigenetic Reprogramming. *J. Biol. Chem.* **287**, 30922–30931 (2012).
175. Yang, Z., Augustin, J., Hu, J. & Jiang, H. Physical Interactions and Functional Coordination between the Core Subunits of Set1/MLL Complexes and the Reprogramming Factors. *PLoS One* **10**, e0145336 (2015).
176. Zhang, H. *et al.* MLL1 Inhibition Reprograms Epiblast Stem Cells to Naive Pluripotency. *Cell Stem Cell* **18**, 481–94 (2016).
177. Nguyen, A. T. & Zhang, Y. The diverse functions of Dot1 and H3K79 methylation. *Genes Dev.* **25**, 1345–58 (2011).
178. Wysocka, J., Myers, M. P., Laherty, C. D., Eisenman, R. N. & 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.* **17**, 896–911 (2003).
179. Mueller, D. *et al.* A role for the MLL fusion partner ENL in transcriptional elongation and chromatin modification. *Blood* **110**, 4445–4454 (2007).
180. Ma, C. *et al.* Nono, a Bivalent Domain Factor, Regulates Erk Signaling and Mouse Embryonic Stem Cell Pluripotency. *Cell Rep.* **17**, 997–1007 (2016).
181. Bond, C. S. & Fox, A. H. Paraspeckles: nuclear bodies built on long noncoding RNA. *J. Cell Biol.* **186**, 637–644 (2009).
182. Ghosal, S., Das, S. & Chakrabarti, J. Long Noncoding RNAs: New Players in the Molecular Mechanism for Maintenance and Differentiation of Pluripotent Stem Cells. *Stem Cells Dev.* **22**, 2240–2253 (2013).

183. Wickramasinghe, V. O. & Laskey, R. A. Control of mammalian gene expression by selective mRNA export. *Nat. Rev. Mol. Cell Biol.* **16**, 431–442 (2015).
184. Xiao, Y. & Weaver, D. T. Conditional gene targeted deletion by Cre recombinase demonstrates the requirement for the double-strand break repair Mre11 protein in murine embryonic stem cells. *Nucleic Acids Res.* **25**, 2985–2991 (1997).
185. Momcilovic, O. *et al.* DNA damage responses in human induced pluripotent stem cells and embryonic stem cells. *PLoS One* **5**, e13410 (2010).
186. Hörmanseder, E. *et al.* H3K4 Methylation-Dependent Memory of Somatic Cell Identity Inhibits Reprogramming and Development of Nuclear Transfer Embryos. *Cell Stem Cell* **21**, 135–143.e6 (2017).
187. Zhang, Z. *et al.* Down-Regulation of H3K4me3 by MM-102 Facilitates Epigenetic Reprogramming of Porcine Somatic Cell Nuclear Transfer. *Cell Physiol Biochem* **45**, 1529–1540 (2018).