Examining the Interaction Partners of Essential Chromatin Regulators in Triple- Negative Breast Cancer

by

Beril Esin

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

Master of Science

in

Cellular and Molecular Medicine

KOÇ
ÜNİVERSİTESİ

August 9, 2023

Examining the Interaction Partners of Essential Chromatin Regulators in Triple- Negative Breast Cancer

Koç University

Graduate School of Health Sciences

This is to certify that I have examined this copy of a master's thesis by

Beril Esin

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:

Prof. Tuğba Bağcı- Önder (Advisor)

Prof. Tamer Önder

Assoc. Prof. Adam Cribbs

Date: 09.08.2023

I dedicate my thesis,

to my parents,

Prof. Melek Nihal Esin,

and Dr. Murat Esin

ABSTRACT

Examining the Interaction Partners of Essential Chromatin Regulators in Triple- Negative Breast Cancer

Beril Esin

Master of Science in Cellular and Molecular Medicine

August 9, 2023

Triple Negative Breast Cancer (TNBC) is a highly invasive breast cancer subtype, which occurs with the absence of HER2, Estrogen and Progesterone receptors. The lack of specific hormone receptors causes the conventional targeted therapies to fail. The absence of such therapeutic options and the complexity of the genetic pathways led us to study the epigenetics of TNBC. From our previous studies, a novel epigenetic regulator SS18L2 was identified to be affecting TNBC cell viability and cell cycle. SS18L2 (Synovial Sarcoma Translocation Gene on Chromosome 18- Like Protein 2) is a homolog of SS18 and SS18L1, chromatin regulators found in BAF (SWI/SNF) complex. There is limited knowledge on SS18L2 and the roles of SS18L2 in cancer are elusive. To decipher the roles of SS18L2 in TNBC cell survival, we aimed to reveal the interaction partners of SS18L2 in a TNBC cell line model.

To examine the interactome of SS18L2, BioID (Proximity Dependent Biotin Identification) method was utilized. The BioID fusion proteins were engineered for SS18L2 and its homologs, SS18 and SS18L1; which were all further examined with cell-based assays. First, western blot experiments revealed efficient and specific production of engineered BioID fusion proteins. The SS18L2, SS18L1, and SS18 fusion proteins were mainly localized to nucleus, as gauged by immunofluorescence staining. Biotinylation efficiencies of the cell lines overexpressing the fusion proteins and the pull-down efficiency were all tested; and N-BioID fusion constructs were chosen for further experiments. To identify the unique interaction partners of SS18L2, different than its homologs, SS18 and SS18L1, mass spectrometry was performed for all three homologs.

The results showed 169 interacting partners of SS18L2, with LogFC higher than 2 and 15 hit proteins with LogFC higher than 5 and p value smaller than 0.001. These interactions were mainly from the BAF (SWI/SNF) complex, such as SMARCC1, SMARCD1, SMARCA2/4, ARID1B, GLTSCR1, DPF2, ARID1A, SMARCB1, SMARCC2 and BCL7. The interactomes of SS18 and SS18L1 were similar to that of SS18L2, although a higher number of proteins were identified (374 and 267 respectively). There were 17 unique interacting proteins for SS18L2, however the significance values for these interactions were low, prompting further studies. Overall, the study shows that SS18L2 forms complexes within BAF complex in TNBC cells. The mechanisms behind the selective cell survival regulatory role of SS18L2 remains unanswered, prompting more studies with different approaches. This thesis provides a strong groundwork for deciphering the role of SS18L2 in cancer.

ÖZETÇE

Üçlü Negatif Meme Kanserinde Temel Kromatin Düzenleyicilerinin Etkileşim Partnerlerinin İncelenmesi

Beril Esin

Hücresel ve Moleküler Tıp, Yüksek Lisans

9 Ağustos 2023

Üçlü negatif meme kanseri (TNBC) oldukça invazif olan ve HER2, östrojen ve progesteron reseptörlerinin yokluğu ile ortaya çıkan bir meme kanseri türüdür. Spesifik hormon reseptörlerinin eksikliği, geleneksel hedefe yönelik tedavilerin başarısız olmasına sebep olmaktadır. Tedavi seçeneklerinin yetersizliği ve genetik yolakların kompleksliği bizi üçlü negatif meme kanserinde epigenetik çalışmalara yönlendirmiştir. Önceki çalışmalarımızda, özgün bir epigenetik düzenleyici olan SS18L2'nun TNBC hücre canlılığında etkileri olduğu gözlemlenmiştir. SS18L2 (Synovial Sarcoma Translocation Gene on Chromosome 18- Like Protein 2), BAF (SWI/SNF) kompleksinde yer alan SS18 ve SS18L1 kromatin düzenleyicilerinin homologudur. Literatürde SS18L2 hakkında sınırlı bilgi bulunmaktadır ve SS18L2'nin kanser türleri üzerindeki etkileri henüz bilinmemektedir. SS18L2'nin TNBC hücrelerinin hayatta kalması üzerindeki etkilerini açığa çıkartmak için SS18L2'nin etkileşim partnerleri TNBC hücre modelinde incelenmiştir.

SS18L2 etkileşim ağını incelemek için BioID (Proximity Dependent Biotin Identification) metodu uygulanmıştır. BioID füzyon proteinleri, SS18L2 ve homologları SS18 ve SS18L1 için tasarlanmıştır ve hücre temelli deneyler ile teyit edilmiştir. Öncelikle, western blot deneyleri tasarlanmış olan BioID füzyon proteinlerinin etkili ve spesifik olarak üretildiğini göstermiştir. SS18L2, SS18 ve SS18L1 füzyon proteinleri immünofloresan boyamalar ile tespit edilmiş ve temel olarak çekirdekte lokalize olduğu gösterilmiştir. Füzyon proteinlerinin stabil ekspresyonu ile oluşan hücre hatlarının biyotinlenme verimliliği ve pull-down deneyleri test edilmiş, N-BioID füzyonları gelecek deneyler için seçilmiştir. SS18L2'nin özgün ve homologları SS18 ve SS18L1'den farklı olan etkileşim partnerlerinin seçilmesi için kütle spektrometresi deneyleri üç homolog için de gerçekleştirilmiştir.

Sonuçlar SS18L2 için LogFC değeri 2'den yüksek 169 adet etkileşim partneri ile birlikte 15 adet LogFC değeri 5' den yüksek ve p değeri 0.001'den düşük olan etkileşim partneri belirlemiştir. Bu etkileşimler çoğunlukla SMARCC1, SMARCD1, SMARCA2/4, ARID1B, GLTSCR1, DPF2, ARID1A, SMARCB1, SMARCC2 ve BCL7 gibi BAF (SWI/SNF) kompleksi üyelerinden oluşmaktadır. SS18 ve SS18L1' in etkileşim partnerleri SS18L2 ile benzer çıkmış, ancak daha fazla protein tanımlanmıştır (sırası ile 374 ve 267). SS28L2 için 17 adet özgün etkileşen protein bulunmuştur, ancak etkileşimlerin önem değerleri düşük olduğundan çalışmalar ilerletilmemiştir. Tümü ile ele alındığında, bu çalışma SS18L2'nin TNBC hücrelerinde BAF kompleks ile etkileştiğini göstermiştir. SS18L2'nin özgün hücrelerde hayatta kalım düzenleyici rolünün arkasındaki mekanizmalar henüz yanıtlanmamıştır ve farklı yaklaşımlar ile daha fazla çalışmaya ışık tutmaktadır. Bu tez, SS18L2'nin kanser üzerindeki rolünü çözmek için güçlü bir temel sağlamaktadır.

ACKNOWLEDGEMENTS

I remember crying from happiness when I first received the acceptance letter from Prof. Tuğba- Bağcı Önder. I wish to express my appreciation towards her, my PI over the last two years, who welcomed me into the lab with open arms. I am sincerely grateful for her, for giving me the opportunity to do what I love. Her constant support and guidance, both scientifically and personally, have evolved me into the scientist that I am today. I would also like to extend my heartfelt gratitude towards Prof. Tamer Önder, who has been the person that I go to whenever I have questions on my mind. His exceptional guidance broadened my perspective on science, and I cannot thank him enough for that. I am also thankful for Assoc. Prof. Adam Cribbs, for being a member of my thesis defence committee.

I thank Koç University Research Center for Translational Medicine (KUTTAM) and The Scientific and Technological Research Council of Turkey (TÜBİTAK) for their financial support in my master's thesis. I would also like to thank Dr. Nazlı Ezgi Özkan-Küçük and Prof. Nurhan Özlü for their contributions to my thesis.

I would like to thank Dr. Özlem Yedier- Bayram deeply, for everything she has done for me. She is the person that guided me through my very first steps into academia. Thanks to her, I learned to never give up and that anything is possible. She has always been my biggest support and the best teammate anybody could ever ask for. I am genuinely fortunate to have her not only as a mentor but also as a lifelong friend.

I would also like to thank all the past and present TBO lab members, who have all brought joy into my life. Dr. Ahmet Cingöz for bringing laughter to my mornings, Dr. Alişan Kayabölen for being my cloning mentor and Dr. Ali Cenk Aksu for all his support. I would also like to thank Nareg Pınarbaşı- Değirmenci for all the wonderful times we have shared during IF experiments, for being there for me when I needed it the most and for putting a smile on my face with her presence.

I would like to thank Göktuğ Karabıyık, for creating the funniest memories with me and for making the thesis writing journey as fun as it could be. I am grateful for Ezgi Yağmur Kala for her heartwarming friendship and Ebru Yılmaz for being an amazing travel buddy. I would like to thank Canan Bayraktar for our countless fun memories and our perfect dinner gatherings. I would like to thank Fulya Mina Küçüktaş for our long car rides, deepest talks, vacations, and hours of study sessions that I cannot fit into one sentence. I thank Beyza Nur Köseoğlu for her continuous love that has always lift me up,

and for our beautiful memories covered with our favourite songs. In addition, I thank every TBO lab member, Aksel Çelikkol, İpek Kök, Sedef Saraç, Ayşegül Durgun and Gizem Doğan Yılmaz for making our times in the lab incredibly enjoyable.

I would like to acknowledge my friends from the next-door labs, Ayça Açar and Enes Çiçek for being my safe places to go to in this journey. Their friendship has given me comfort and strength.

I would also like to express special thanks to my beloved undergrads, Baturay Altay and Elif Ayça Güvener, who have stayed with me along the late-night experiments and have given me the motivation to carry on every day.

I would also like to thank my closest friends; Elif, Begüm, Kıvanç, Zeynep, Öykü, Mirey and Umut for being the light in my life whenever I feel stuck. Finally, I would like to thank my mom, Prof. Melek Nihal Esin for being the reason that I love science, and my dad Dr. Murat Esin, for being my biggest fan. Knowing that they are always there to support me has been a constant source of happiness and strength throughout my life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

BAF	BRG/BRM-associated factor
BioID	Proximity dependent biotin identification
BirA*	Promiscuous BirA ligase
BRD	Bromodomain
BSA	Bovine Serum Albumin
CRISPR	Clustered Regularly Interspaced Palindromic Repeats
DMEM	Dulbecco's Modified Eagle Medium
DNMTs	DNA Methyltransferases
ECM	Extracellular Matrix
EPIKOL	Epigenetic Knockout Library
ER	Estrogen Receptor
HAT	Histone Acetyltransferase
HDAC	Histone Deacetylase
HEK293T	Human Embryonic Kidney 293T Cells
HER2	Human Epidermal Growth Factor
HKDM	Histone Lysine Demethylases
HRP	Horseradish Peroxidase
kDa	kilo Dalton
KMT	Lysine methyltransferase
KO	Knockout
LC-MS/MS	Liquid Chromatography-Tandem Mass Spectrometry
LogFC	Log Fold Change
LR	attL-attR Gateway Cloning
MBD	Methyl CpG Binding Domain
MDA-MB-231	MD Anderson-Metastatic Breast-231 Cells
OE	Overexpression
PBS	Phosphate-Buffered Saline
PFA	Paraformaldehyde
PMSF	Phenylmethylsulfonyl Fluoride
PPI	Protein- Protein Interactions
PR	Progesterone Receptor

PRMT	Protein Arginine Methyltransferase
QPGY	Glutamine, Proline, Glycine, and Tyrosine Rich domain
SNH	SYT N-Terminal homology domain
SS18	Synovial Sarcoma Translocation Gene on Chromosome 18
SS18L1	SS18- Like Protein 1
SS18L2	SS18- Like Protein 2
SWI/SNF	Switch/Sucrose Non-Fermentable
TET	Ten-Eleven Translocation
TNBC	Triple-Negative Breast Cancer

Chapter 1:
INTRODUCTION

1.1 Triple-Negative Breast Cancer

Triple- Negative Breast Cancer (TNBC) is the most aggressive breast cancer subtype, which comprises 10-15% of the diagnosed breast cancer patients (Won & Spruck, 2020). TNBC occurs in the absence of three specific hormone receptors: ER (estrogen receptor), PR (progesterone receptor) and HER2/ERBB2 (human epidermal growth factor receptor type 2) (Foulkes et al., 2010). Breast cancer patients that do not express the three hormone receptors have a high metastatic state, poor prognosis, and a higher tendency to relapse when compared to other breast cancer subtypes (Yin et al., 2020).

TNBC is a heterogeneous cancer type with distinct molecular subtypes, which supports the complexity of the disease. In a study conducted by Lehmann et al., tumor samples of 587 TNBC patients were collected, and gene expression profiling of the tumor samples were performed (Yin et al., 2020). The gene expression profiling resulted classified TNBC into six specific subtypes (TNBC type) (**Figure 1.1**). The subtypes identified in the study comprise of basal- like 1 (BL1), basal- like 2 (BL2), mesenchymal (M), mesenchymal stem cell like (MSL), immunomodulatory (IM), and luminal androgen receptor (LAR) subtypes (Lehmann et al., 2011). BL1 subtype are mostly enriched in cell division or cell cycle related components. In addition, DNA damage response pathways are enriched with the proliferation pathways in this subtype. BL2 subtype is mostly associated to growth factor signaling pathways (Wnt/β-catenin, EGF pathway and NGF pathway) glycolysis and gluconeogenesis. Moreover, M and MSL subtypes have specific gene ontologies that are enriched in cell motility pathways (Rho pathway), cell differentiation pathways and ECM receptor interaction. Different from the M subtype, MSL subtype has a lower expression of the proliferation related genes and claudins. The IM subtype has a higher abundancy in gene ontologies related to immune cell processes such as immune cell signaling or cytokine signaling. Finally, the LAR subtype is indicated to be the most distinctive subtype. Even though the subtype is ER negative, gene ontologies were found to be expressing hormone regulated pathways. Moreover, the study indicates androgen receptor (AR) was highly expressed in mRNA levels compared to

other subtypes, which might explain the hormone regulated pathways (Lehmann et al., 2011). The intrinsic molecular subtypes of TNBC and the six specific subtypes were also studied by Lehmann and Pietenpol in 2014. The study resulted in separating TNBC subtypes into classifications related to PAM50 intrinsic subtypes (Luminal A, Luminal B, HER2-enriched, Basal-like, and Normal-like) and categorized into groups such as basal- like and non-basal-like, low claudin and high claudin (Lehmann & Pietenpol, 2014). Further research of Lehmann in 2016 the TNBC molecular subtypes were simplified into four tumor- specific subtypes entitled TNBCtype-4 (BL1, BL2, M and LAR) using publicly available breast cancer gene expression data sets (Lehmann et al., 2016).

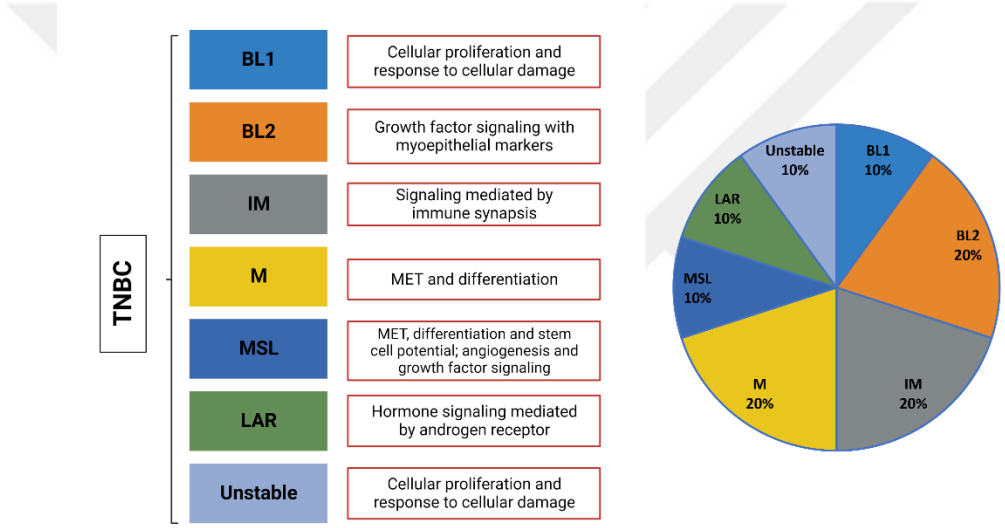

Figure 1.1. Lehmann's Triple Negative Breast Cancer subtypes (TNBCtype). Pie chart indicates the distribution of the subtypes among TNBC. MET stands for mesenchymal epidermal transition. Adapted from (Uscanga-Perales et al., 2016). Figure created with BioRender.com.

In addition to the various subtypes, there are multiple molecular signaling pathways essential for TNBC cell invasion, proliferation, or survival. The multiple signaling cascades are initiated with the GPCR's (G-Protein Coupled Receptors), receptor tyrosine kinases and integrins with their downstream effectors (Iorio et al., 2016). The signaling pathways include Ras- mediated signaling through receptor tyrosine kinases. Ras interacts with Raf and activates it, while the phosphorylation of Raf activates MEK, phosphorylating MAPKs and resulting in TNBC cell proliferation. Integrins initiate signaling pathways associated with FAK, cdc42 and Src. Moreover, activating receptor

tyrosine kinases stimulate PI3K pathway, activating AKT and mTOR pathway. In addition, the G-proteins activate phospholipase C and NF-κB (nuclear factor kappa B) (Iorio et al., 2016). All the molecular signaling pathways result in the progression and survival of TNBC.

TNBC is chemotherapy sensitive; therefore, the current treatment for the disease includes chemotherapy drugs like anthracycline, anti-microtubule agent taxane alkylating agents, and an anti-metabolite fluorouracil, followed by surgery (Won & Spruck, 2020). However, the aggressive cancer type does not respond to the conventional chemotherapies. Moreover, due to the lack of the specific hormone receptors, endocrine, HER2 treatments and other molecular targeted therapies fail to work (Yin et al., 2020). In addition, TNBC is diagnosed in the advanced stages of the tumor that reduces the effects of the treatments drastically (Wiggs et al., 2022). The low efficiency of the existing treatment strategies creates a need for novel therapies for a higher survival rate of the patients.

Due to the uniqueness of the triple negative breast cancer subtypes, complexity of the molecular signaling pathways and treatment options with low efficiency, we studied the epigenetic regulations of triple- negative breast cancer in this study.

1.2 Epigenetics

Epigenetics is the study of the heritable changes in the gene expression which do not alter the genomic sequence of the DNA. On the contrary, epigenetic modifications modify chromatin structure or the accessibility of the DNA (Handy et al., 2011). These alterations regulate gene expression, which can also be linked to cell cycle or tumor progression. Modifications to DNA and chromatin form the basis of epigenetics and epigenetic modifications. Epigenetic mechanisms can be listed as DNA methylation, non-coding RNA, loss of imprinting and chromatin modifications (Hamilton, 2011) (**Figure 1.2**).

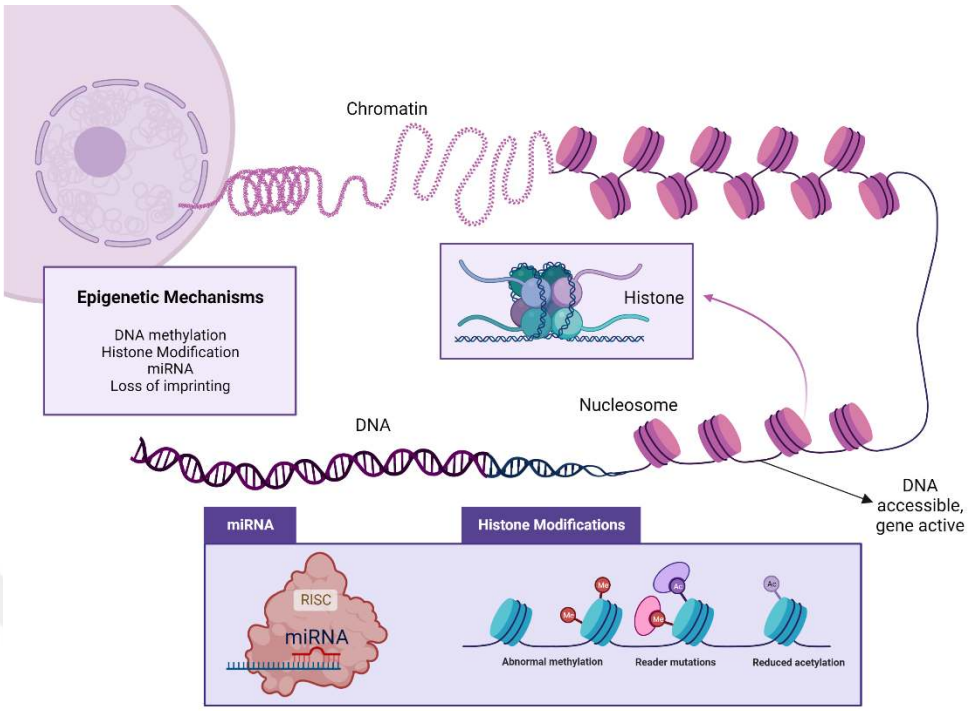

Figure 1.2. Schematic representation of epigenetic modifications. Adapted from (Audia & Campbell, 2016). Figure created with BioRender.com.

1.2.1 DNA Methylation

DNA methylation is one of the most essential epigenetic modifications in the genome. The addition or removal of a methyl group on the DNA occurs from S-adenyl methionine (SAM) to the cytosine ring at the C-5 position (Jin et al., 2011). Methylation has a role in repressing or activating the gene expression, which is regulated by DNA methyltransferases (DNMTs) (Dhar et al., 2021). The DNMT family comprises of DNMT1, DNMT2, DNMT3A, DNMT3B, and DNMT3L (Jin et al., 2011). DNMT2 is a homolog of methyltransferases that function as tRNA transferases. DNMT3A and DNMT3B form a new pattern of methylation to the unmodified and non-methylated DNA, which are called “de novo” enzymes. Conversely, DNMT1 takes place in DNA replication, where the enzyme copies the DNA methylation pattern from the parental strand to the daughter strand (Moore et al., 2013). The modifications indicate that methylation is mostly associated with gene silencing, while demethylation of the DNA is associated with gene activation (Wei et al., 2017).

Methyl group addition occurs only in CpG islands, which are regions in the genome where guanine follows a cytosine. CpG islands are highly abundant in nearly half of the promoters of all of the genes (Hamilton, 2011). Under normal circumstances, CpG islands that are outside the promoter region of the gene are methylated, which results in various levels or versions of mRNA. On the other hand, DNA hypomethylation (demethylation, loss of a methyl group) of CpG island outside of the promoter region can be leading to translocations, chromosomal rearrangements or deletions (Hamilton, 2011). Uncommon changes in methylation on the DNA can cause cancer formation or tumor progression (Ehrlich, 2019). Moreover, DNMT deficiency or the loss of DNA methylation can be associated with the progression of tumor growth (Ehrlich, 2009). Hypermethylation, the addition of methyl groups, can cause transcriptional silencing followed by the loss of protein expression. Consequently, hypermethylation of the genes could be resulting in gene-silencing or allelic loss, creating the gene to act as a tumor suppressor (Hamilton, 2011).

1.2.2 Non-Coding RNA

Non- coding RNAs are RNA clusters, which do not code for proteins. They are indicated to being functional in the post-translational modifications; therefore, being an essential epigenetic modifier (Wei et al., 2017). Studies have shown that only 2-3% of the human genome is translated into proteins and the non-coding genes are transcribed as non-coding RNAs. Non-coding RNAs differ in size as short non-coding RNAs (miRNA, siRNA, piRNA) and long non-coding RNAs (Yang et al., 2020). Even though the non-coding RNAs do not translate into gene expression in terms of protein formation, it was emphasized that they play an essential part in epigenetic modifications (Wei et al., 2017).

The most common epigenetic modification by the non-coding RNAs is from microRNAs (miRNAs). miRNAs are short non-coding RNAs with the length of 19 to 22 bp average (Yang et al., 2020). miRNAs can bind to the direct sequence of an mRNA and degrades the mRNA with the aid of RISC complex. In addition, if the miRNA and mRNA do not have a perfect sequence match, miRNA binds to the 3’ end of the mRNA partly (Hamilton, 2011). This causes the inhibition of the function of the transfer RNA. Both mechanisms of miRNA affect the structure or the functions of the RNAs causing modifications in RNA in eukaryotic cells (Yang et al., 2020).

1.2.3 Loss of Imprinting

Genomic imprinting occurs when the offspring inherits two alleles, one from the mother and one from the father, and one of the alleles is turned off (Bajrami & Spiroski, 2016). This could occur with DNA methylation or histone acetylation existing on the imprinted gene (Hamilton, 2011). Therefore, the gene that is turned off is silenced or "imprinted" in the offspring. However, only one allele from one of the parents is expressed (Bajrami & Spiroski, 2016). Therefore, loss of imprinting implies the gene is either expressed in a bi-allelic manner, or both copies are not expressed. The silencing of the gene expression results in epigenetic alterations, which might lead to diseases such as Angelman syndrome and Prader-Willi syndrome (Butler, 2009).

1.2.4 Chromatin Modifications

One of the fundamental components of epigenetics is chromatin, where the epigenetic alterations occur. The foundation element of chromatin is nucleosomes. A nucleosome consists of a 147 bp of DNA, which can be found packaged around a histone octamer: H2A, H2B, H3 and H4, two of each histone protein (Dawson & Kouzarides, 2012). Spacing between the nucleosomes determine the chromatin structure followed by the formation of chromosomes. Nucleosome spacing when the nucleosomes are in an open structure means that the chromatin is active (Hamilton, 2011). In the active state of chromatin, there are several post-translational modifications affecting the histone tails. There are two mechanisms of chromatin structure that lead to DNA and histone modifications. First, chromatin packing can be modified by the change of charge or with the interactions between nucleosomes. This results in the opening or the closing of DNA polymer, which manages the access of DNA binding proteins like transcription factors. Second, chemical moieties can alter the nucleosome surface and initiate interactions with chromatin binding proteins (Berger, 2007). These modifications comprise of methylation, acetylation, phosphorylation, ubiquitination and SUMOylation. All histone post translational modifications are removable (Kouzarides, 2007).

1.2.5 Epigenetic Modifiers: Writers, Readers, and Erasers

The epigenic modifiers causing chromatin and histone modifications are categorized as writers, readers and erasers (Audia & Campbell, 2016) (**Figure 1.3**). Writers introduce modifications on DNA and histone tails, mostly by methylation and acetylation. The modifications alter transcriptional activation or repression. The epigenetic writers are listed as: DNA methyl transferases (DNMTs), histone lysine methyltransferases (HKMTs), protein arginine methyltransferases (PRMTs) and histone acetyltransferases (HATs) (Biswas & Rao, 2018). SET1 family and their correlation to histone 3 lysine 4 (H3K4) associated with transcriptional activation can be given as an example of HKMTs (Sugeedha et al., 2021).

Readers are epigenetic modifiers that recognize the epigenetic writers. They bind to different covalent modifications on the DNA by utilizing specific docking domains (Damiani et al., 2020). The epigenetic readers are listed as: Methyl CpG binding proteins (MBPs), histone methylation readers and histone acetylation readers (Biswas & Rao, 2018). Specialized domains for readers are bromodomain (BRDs), chromodomain, Tudor domain, plant homeodomain (PHD) and Methyl CpG binding domain (MBD) (Damiani et al., 2020). Histone acetylated lysines bind to bromodomains that are often demonstrated in chromatin remodeling complexes (Bannister & Kouzarides, 2011). Essential proteins from the reader proteins related to our study are most significantly bromodomain proteins BRD7, BRD9 and BRD4 and their association with chromatin remodeling. Moreover, SMARCA2/4, which are essential members of BAF (SWI/SNF) complex, are significant for cell proliferation and differentiation (Audia & Campbell, 2016). In addition, ZnF domain can also recognize methylated DNA regions with proteins such as ZBTB33/Kaiso, ZBTB4, ZBTB38 and ZFP57 (Biswas & Rao, 2018).

Readers can be categorized into subgroups, considering their differences in recognizing distinct modified amino acid residues. Different modification states of the same amino acid can be emphasized by histone H3 lysine 27 unmethylation (H3K27), monomethylation (H3K27me), dimethylation (H3K27me2), trimethylation (H3K27me3), and acetylation (H3K27ac) (Damiani et al., 2020).

Erasers work on histones and remove the small covalent modifications; therefore, eliminating the modifications conducted by the writer proteins. Epigenetic erasers can be listed as Ten-eleven Translocation (TET), histone deacetylases (HDACs) and histone

lysine demethylases (HKDMs) (Biswas & Rao, 2018). One of the essential members of the HKDM family is LSD protein family. They include histone demethylase LSD1 (KDM1A) and LSD2 (KDM1B), which remove a methyl group from histone lysines. In addition, HDACs deacetylate histones from lysine residues and are involved in metabolic regulation (Yang et al., 2016).

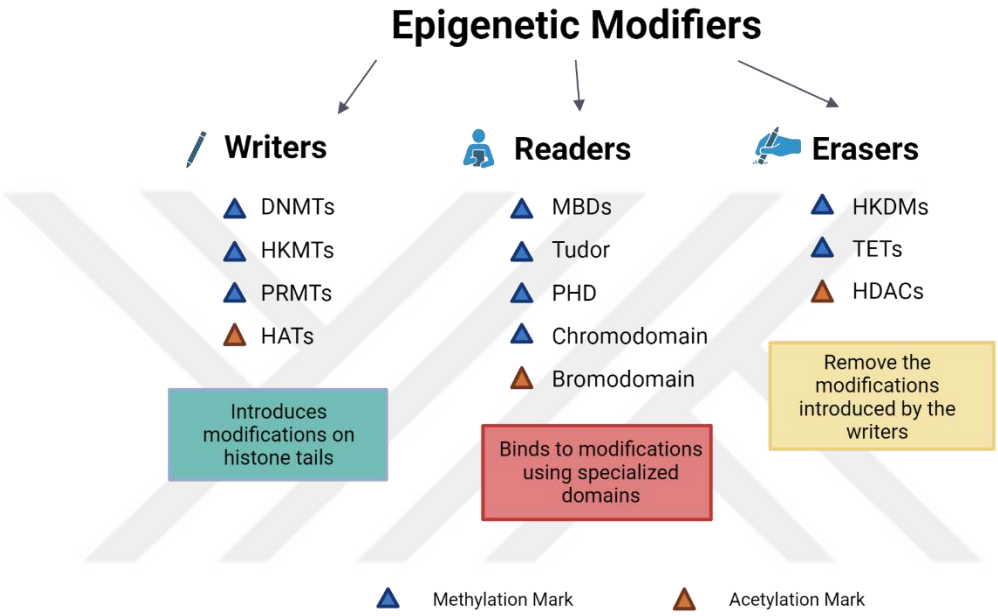

Figure 1.3. Summary of Writer, Reader, and Eraser epigenetic modifiers. Modifications related to methylation are marked with blue triangle, modifications related to acetylation are marked with orange triangle. Adapted from (Biswas & Rao, 2018). Figure created with Biorender.com.

1.3 BAF (SWI/SNF) Complex

One of the most crucial chromatin regulators in epigenetics are the ATP-dependent chromatin remodeling factors, such as BAF chromatin remodeling complex (Sokpor et al., 2017) . BAF (BRG/BRM-associated factor) complex, also known as mammalian SWI/SNF (SWItch/Sucrose Non-Fermentable) complex, is an ATP- dependent complex that alters chromatin dynamics and regulates gene expression (Alfert et al., 2019). Mammalian BAF complexes are conserved evolutionarily. The complex was first discovered in yeast (SWI/SNF), then in drosophila (BRM- associated protein BAP), and later in mammals as BAF complex (Pagliaroli & Trizzino, 2021).

Mammalian BAF complex utilizes the energy in ATP hydrolysis to modify the chromatin structure. First, BAF complex binds to nucleosomes in a central cavity, where DNA is accessible. With ATP, the complex breaks the bonds between histones and the DNA (Pagliaroli & Trizzino, 2021). This forms a transient DNA loop around the nucleosomes, which could result in sliding or ejecting the nucleosomes, affecting chromatin accessibility to the binding of transcription factors (Mittal & Roberts, 2020).

1.3.1 BAF complex Subtypes

BAF complexes are macromolecular assemblies that is composed of numerous diverse subunits (Mittal & Roberts, 2020). There are three distinctive subunits of the BAF complex: cBAF (canonical BAF), pBAF (polybromo-associated BAF) and ncBAF (non-canonical BAF) (Wanior et al., 2021). All of the subunits contain core proteins; SMARCC1, SMARCC2 and SMARCD1 and few more as can be seen in **Figure 1.4**. Moreover, SMARCA2 (BRM) and SMARCA4 (BRG1) subunits, catalytic subunits catalyzing ATP hydrolysis, are both available in all subunits (Mashtalir et al., 2018). All core subunits share different configuration. Differentially cBAF contains ARID1A and ARID1B, ncBAF contains GLTSCR1, GLTSCR1L and BRD9 and pBAF contains ARID2, PHF10, BRD7 and PRBM1 (Wanior et al., 2021). Conversely, pBAF does not have SS18 or SS18L1 as a subunit while two subunits are both part of cBAF and ncBAF. The differences in the BAF complex members result in various transcriptional expressions. Whereas pBAF and ncBAF have a higher number of promoter regions, cBAF is associated with binding to enhancers (Wanior et al., 2021).

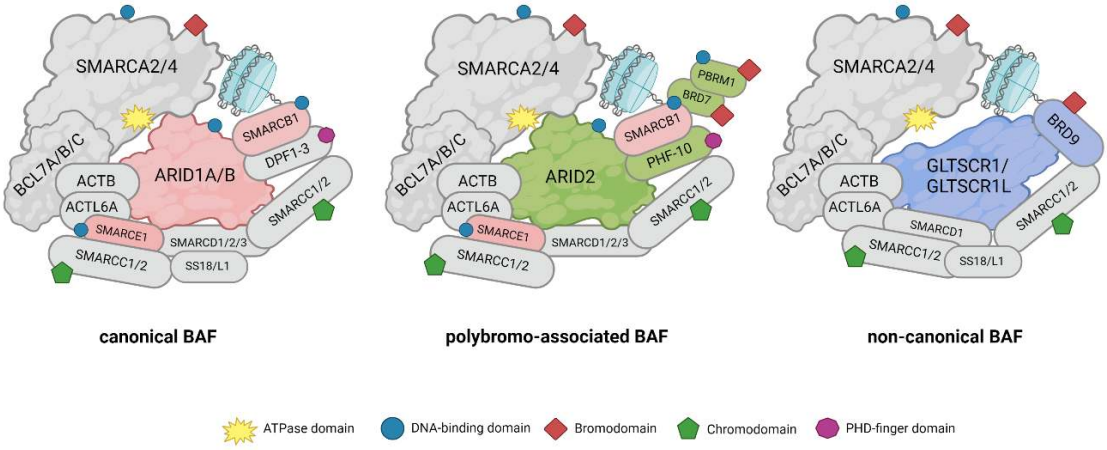

Figure 1.4. Subunits of BAF(SWI/SNF) complex. Specific domains and the shapes associated to them are indicated in the figure. All shared subunits are in grey, subunits that are not found in ncBAF are red, subunits that are unique to pBAF are green and subunits that are unique to ncBAF are blue. Adapted from (Wanior et al., 2021). Figure created with BioRender.com.

1.3.2 BAF Complex in Cancer

Among the roles of BAF complex in different transcriptional activities and its effects on gene expression, it was expected that a chromatin remodeler like BAF (SWI/SNF) complex might play a role in malignancy formation. One of the most essential members of the BAF complex, SMARCB1 (Baf47), which is only found in cBAF and pBAF and has a role in the maintenance of enhancer targeting of the BAF complex (Wang et al., 2017). Studies have shown that SMARCB1 loss due to genetic mutations can promote cancer (Mittal & Roberts, 2020). Consequently, SMARCB1 has been identified as having a tumor-suppressive role in humans, being mutated in 25% of the cancers (Cooper & Hong, 2022). Moreover, it has been suggested that higher expressions of SMARCB1, regulating super-enhancers, may inhibit the formation of ncBAF (Wanior et al., 2021). The complete biallelic loos of SMARCB1 was defined to be causing various tumor types, which malignant rhabdoid tumor of the kidney and atypical teratoid rhabdoid tumor can be given as examples (Cooper & Hong, 2022; Mittal & Roberts, 2020). The role of SMARCB1 in regulating super-enhancer binding is essential for rhabdoid tumor survival, including some super-enhancers like *SOX2* in brain- derived rhabdoid tumors that are repressed by polycomb activation (Tamaki et al., 2015; Wang et al., 2017). Loss of SMARCB1 results in the reversion of the repression of *SOX2,* leading to a decrease in

cell differentiation and resulting in rhabdoid tumor progression (Kadoch & Crabtree, 2013).

Another essential role of SMARCB1 loss, is in the formation of synovial sarcoma (Alfert et al., 2019). Synovial sarcoma is a subtype of soft tissue sarcomas. The sarcomas have a differential characteristic due to the chromosomal translocation t(X;18). In addition to the chromosomal translocation, synovial sarcoma occurs with the formation of SS18-SSX fusion protein and its relation between *PRC2* and BAF for chromatin silencing and remodeling (Gazendam et al., 2021). Eight C-Terminal amino acids of SS18 in SS18-SSX complex is replaced by 78 amino acids of the C-terminus of SSX1, SSX2 or SSX4 (Gure et al., 2002). SSX proteins were demonstrated to control repression while SS18 takes place in transcriptional activation (Hale et al., 2019). The presence of SS18-SSX replaces the wild type SS18 predominantly in BAF complex, which leads to the removal of SMARCB1 (McBride et al., 2018). In the proposed model for SS18-SSX, McBride et al. demonstrates that the binding of the fusion protein and the removal of SMARCB1 directs BAF complex to the broad polycomb domains away from enhancers. Polycomb proteins are transcriptional repressors created by multi-protein complexes; moreover, they mediate the chromatin compaction resulting the chromatin silencing of the *HOX* genes (Hale et al., 2019; Morey & Helin, 2010). The relocation of BAF complex to the broad polycomb domains suppresses the *PRC2* (polycomb repressive complex 2) gene and activates the bivalent genes depending on the levels of H3K4me3 levels. *PRC2* carries out transcriptional silencing with the functions of a catalytic subunit EZH2. EZH2 catalyzes the tri-methylation of H3K27. Binding of BAF complex to the broad polycomb domain results in the activation of gene expression, which is dependent on the increasing levels of H3K4me3. (Hale et al., 2019) (**Figure 1.5**).

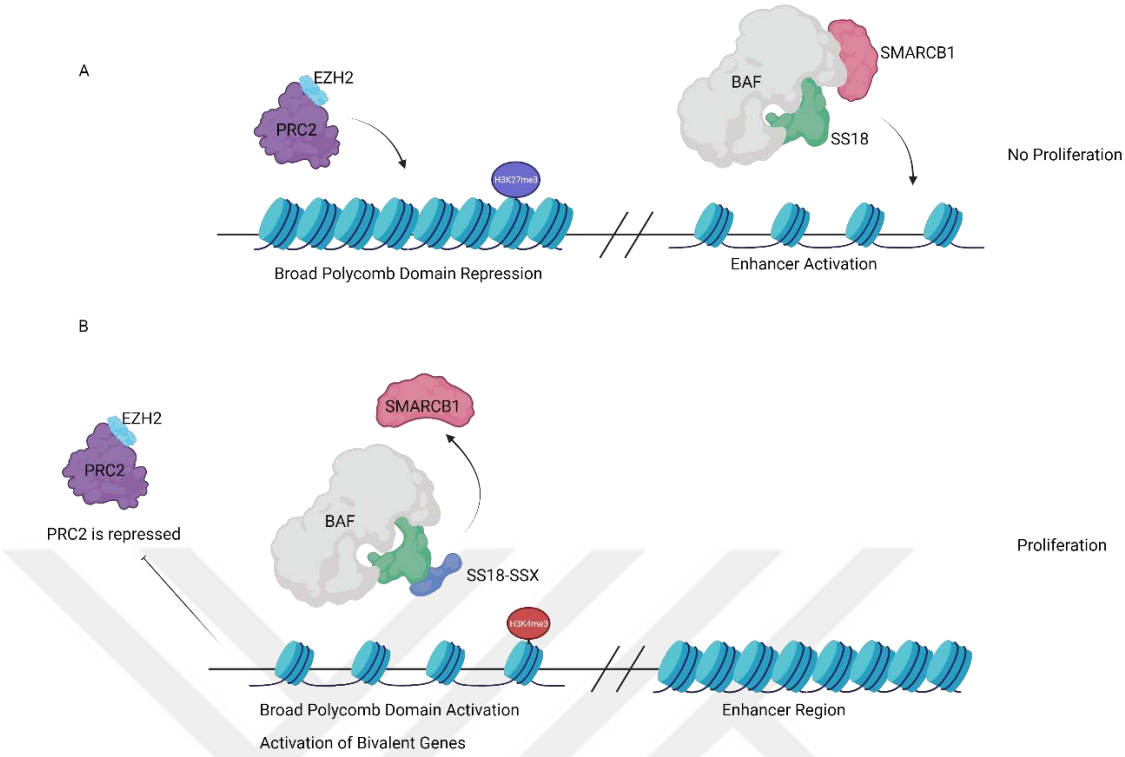

Figure 1.5. The mechanism of SS18-SSX. A. Normal conditions where wild type SS18 and SMARB1 are present. Broad polycomb domain is repressed. PRC2 mediates H3K27me3 leading to gene silencing. **B.** The binding of SS18-SSX to the BAF complex with the loss of SMARCB1 results in broad polycomb domain activation. PRC2 is repressed. Bivalent gene activation occurs, H3K4me3 expression increases. Figure created with Biorender.com.

1.4 EPIKOL (Epigenetic Knock-Out Library)

To investigate the roles of epigenetic modifiers in cancer cells, a custom created epigenome- wide CRISPR-Cas9 knockout library was generated by the joint efforts of TBO, TO and NL labs at Koç University School of Medicine. The knock-out library aims to target a high range of epigenetic modifiers that regulate triple-negative breast (TNBC) and prostate cancer (Yedier-Bayram et al., 2022). The library comprises of 7820 sgRNAs that target 719 epigenetic modifiers, which have different roles in various chromatin associated pathways. In addition, EPIKOL content includes 25 genes from distinct families for internal control, 35 essential genes like ribosomal protein-encoding genes and finally 80 non-targeting sgRNAs.

1.4.1 Screen Validations and SS18L2

CRISPR screens are efficient methods utilized for diverse applications in biology. In a typical pooled CRISPR-Cas9 KO screen, a library that contains various sgRNAs is introduced into cells (Bock et al., 2022). To decipher the specific epigenetic modifiers essential for TNBC cell fitness, three TNBC cell lines were screened from the library: SUM159PT, SUM149PT and MDA-MB-231. Non-malignant human mammary epithelium cells (HMLE) were also screened as a non-cancer cell type (Yedier-Bayram et al., 2022). The sgRNAs will only result in depletion of the gene if the target gene is cleaved in most of cells within the population (Wang et al., 2014). From the validations of the EPIKOL screens, 15 specific genes were identified as essential modifiers and were determined to be depleted in all of the three TNBC cell lines. Four highest hits regarding the depletion scores from the screen were *KAT8, KANSL3*, *KANSL2* and *SS18L2* (Yedier-Bayram et al., 2022). Apart from *SS18L2*, the selected hits were members of the NSL (Non-Specific Lethal) complex, which are indicated to be localizing in the nucleus to gene promoters and enhancers (Sheikh et al., 2019). Competition assays and long-term colony formation assays showed TNBC cell fitness defects significantly when *SS18L2* and NSL members (*KAT8*, *KANSL3, KANSL2*) are knocked-out. Knock- outs of the epigenetic modifiers were shown to induce apoptosis by the significant cleavage of PARP (poly ADP ribose polymerase). Additionally, with the knock-out of *SS18L2*, cell number was observed to be decreasing in the G0/G1 and S phase, which leads to cell accumulation at the G2/M phase (Yedier-Bayram et al., 2022). The results indicated a cell cycle arrest upon SS18L2 knock out at the G2/M phase.

1.4.2 SS18L2 and Cell Cycle

To further analyze the notion that SS18L2 could be related to cell cycle, RNA sequencing was performed to determine the changes in the transcriptome in MDA-MB-231 cells (Yedier-Bayram et al., 2022). The knock-out of *SS18L2* led to a greater number of upregulated genes than the downregulated genes (**Figure 1.6A**). However, in the downregulated genes, gene set enrichment analysis led to the identification of biological process like chromosome organization, and segregation of the related genes (**Figure 1.6B**). Moreover, the differential expression of the cell cycle related genes was shown in **Figure 1.6C** with late G2/M phase, early G1/S phase and DREAM complex related genes

(Yedier-Bayram et al., 2022). DREAM complex members are genes that plays a part in the cell cycle, where the gene expression increases in G1/S phase and G2/M phase (Sadasivam & DeCaprio, 2013). Moreover, to verify the relation of the downregulated genes to cell cycle, PIP-FUCCI experiments with fluorescence imaging demonstrated that *SS18L2* knockout lead to a lower number of cells compared to control (**Figure 1.6D, E**). Cells that indicated G2/M phase (with the mCherry fluorescence) were significantly higher with the knockout of *SS18L2* (Yedier-Bayram et al., 2022). Therefore, it was concluded that upon the knockout of *SS18L2*, the cells can enter mitosis; however, get arrested at the G2/M phase causing a decrease in TNBC cell viability.

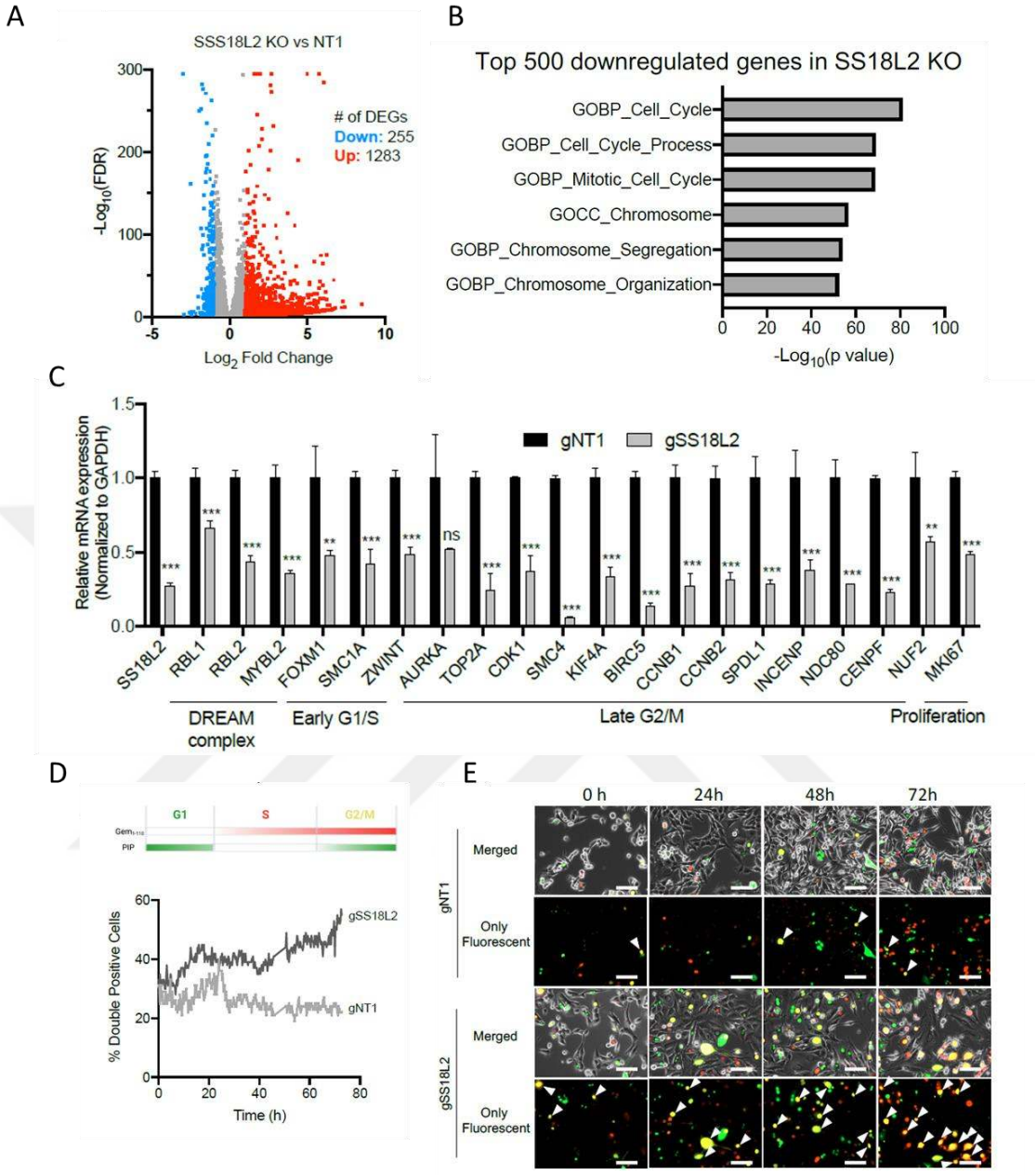

Figure 1.6. The effects of SS18L2 knockout on cell cycle. A. Volcano plot after RNA sequencing, indicating differentially expressed genes with the knock-out of SS18L2. **B.** Gene set enrichment analysis (GSEA), and biological processes related to top 500 downregulated genes with the knockout of SS18L2. **C.** QPCR (quantitative real-time PCR) analysis of the cell cycle related genes with the knock-out of SS18L2. **D.** PIP-FUCCI analysis schematic representation. The expression rated of cells that have mCherry (red) and mVenus (green). Both fluorescence signals indicate G2/M phase. **E.** PIP-FUCCI experiment for 72 hours, representative images. Arrows indicate yellow fluorescent. P values statistically analyzed by two-tailed Student's *t*-test, $*p < 0.05$, $**p < 0.01$, $***p < 0.001$. Figure from (Yedier-Bayram et al., 2022).

Following the results regarding the role of SS18L2 in cell cycle, knockout of the homologs of SS18L2, which are BAF complex members SS18 and SS18L1, were investigated. Interestingly, it was shown that the KO of the homologs did not affect the cell viability while the KO of SS18L2 significantly decreased TNBC cell fitness (**Figure 1.7**).

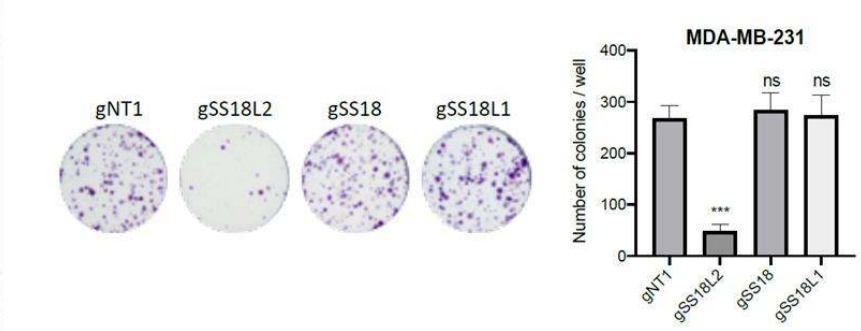

Figure 1.7. Long term colony formation assay indicating differences between the effects of SS18L2 KO and the KO of its homologs. Figure from (Yedier-Bayram et al., 2022).

1.5 SS18L2 in Literature

The SS18L2 gene, Synovial Sarcoma Translocation Gene on Chromosome 18- Like 2), is a homolog of SS18 and SS18L1 (de Bruijn & Geurts van Kessel, 2006). SS18L2 was shown to be a pseudogene of SS18 (de Bruijn et al., 2001). As illustrated in Figure 1.8, SS18L2 has differences from its homologs. In the literature, studies have indicated varieties in terms of the domains of SS18L2, SS18 (SYT) and SS18L1 (CREST). While SS18 and SS18L1 has a similar domain structure with 11 exons and similar intron-exon lengths, SS18L2 comprises of only 3 exons (**Figure 1.8**) (de Bruijn & Geurts van Kessel, 2006). Moreover, SS18 and SS18L1 consists of 2 domains, SNH (SYT N-Terminal homology domain) and QPGY (Glutamine, Proline, Glycine, and Tyrosine rich) domain. SNH domain is a conserved domain found in many species, at the N terminal of SS18 and SS18L1 (Perani et al., 2003). Moreover, SNH domain is related to protein- protein interactions, most of the proteins that are engaging in epigenetic regulations (de Bruijn & Geurts van Kessel, 2006). The domain was shown to bind to SNF11 (SNF11 BD) binding domain regions, most importantly interacting with BRM (SMARCA2) and BRG1 (SMARCA4) proteins (Perani et al., 2005). Moreover, the QPGY domain, present at the C-Terminus of SS18 and SS18L1, plays a role in transcriptional activation. It was also shown that SS18 interacts with p300 histone acetyltransferase, and does not interact with

CBP, which are homologous to each other (Martire et al., 2020). Since SS18L2 lacks the QPGY domain, it was suggested that SS8L2 might be interfering the functions of SS18 and SS18L1 competitively (de Bruijn & Geurts van Kessel, 2006).

Through the small number of research conducted on SS18L2, it is feasible to conclude that we have very little knowledge on SS18L2 in the literature, which comprises of older research papers. Besides the findings from our previous study, the role of SS18L2 in cancer is still unknown and needs further research. In a study of identifying transcriptional factors, SS18L2 was utilized as an unknown cofactor and shown to be interacting with some of the BAF complex members in HEK293T cells (Alerasool et al., 2022). However, no functional validation assays were performed focusing on the interactions of SS18L2. Furthermore, the role of SS18L2 in cancers, specifically in TNBC is unknown.

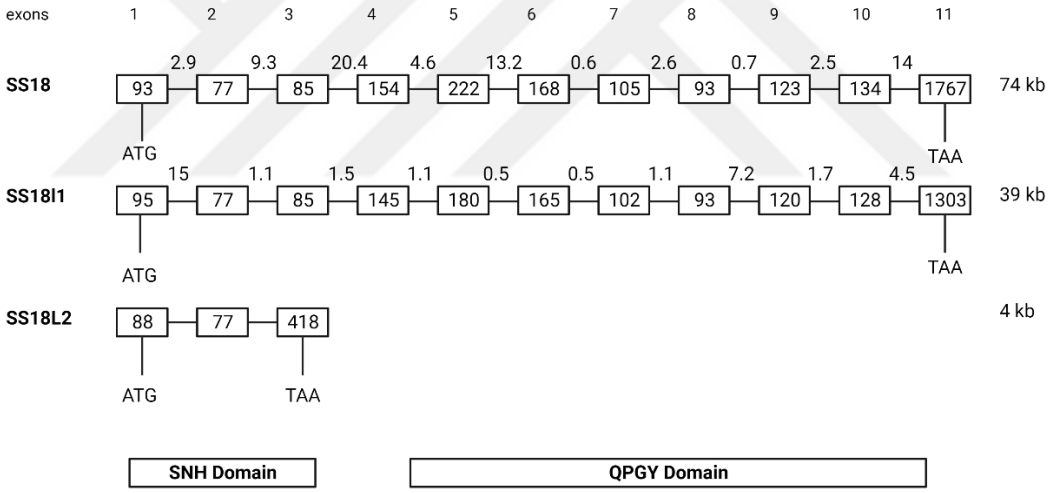

Figure 1.8. Domains of SS18, SS8L1 and SS18L2 genes. Small boxes indicate exons in base pairs. Bars indicate introns in kilobases. Adapted from (de Bruijn & Geurts van Kessel, 2006). Figure created with BioRender.com.

1.6 Proteomic Approaches & BioID

Proteomics is the investigation of the interaction of proteins, finding out their function, structure, or their activities. It is a highly complex approach due to the variety of changes in protein expression dependent on different causes like external conditions and time (Al-Amrani et al., 2021). However, examining the interactions between proteins can result in new findings that could not be achieved otherwise. The differences in the

proteome than the exact genomic sequence can lead to understanding the function an organism in a deeper level, even to the understanding of post translational modifications (Murtaza et al., 2020). Therefore, to understand the functions and interactions of an epigenetic modifier such as SS18L2, proteomic approaches are effective ways for further investigation and functional characterization.

Most commonly used techniques in proteomics utilizes Mass Spectrometry as a tool, to decipher the protein profiles of the living cells (Al-Amrani et al., 2021). One of the techniques utilized to identify the protein- protein interactions (PPIs) is the BioID (proximity dependent biotin identification) method. The method is used to identify the candidate proteins, taking the interactome of a specific protein into account, also called the "bait" protein (Sears et al., 2019). The procedure of BioID is composed of the generation of the BioID fusion protein (insert-BirA*), the overexpression of the protein in the target cells, biotinylation, lysis, pull down and identification with the Mass Spectrometer (**Figure 1.9**) (Roux et al., 2018).

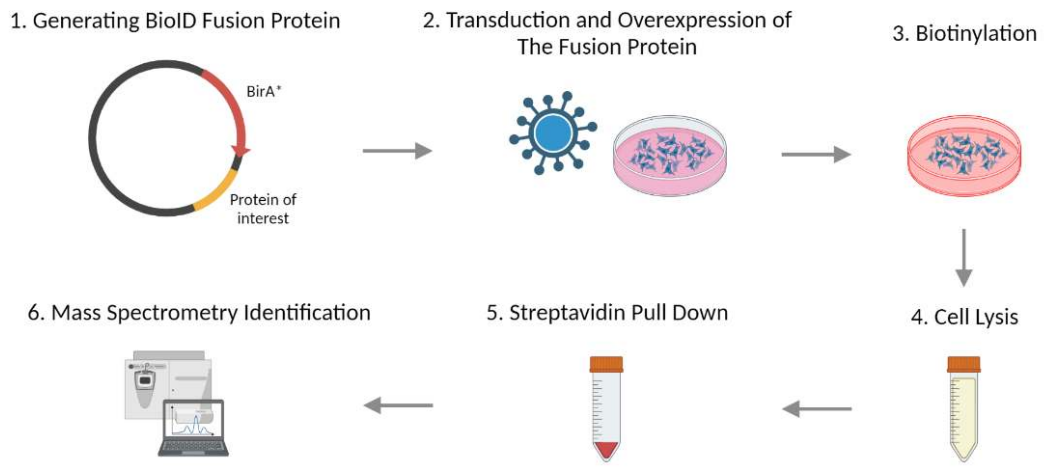

Figure 1.9. Schematic representation of the BioID Procedure. Figure created with BioRender.com.

The generated fusion protein includes BirA*, a mutant version of biotin protein ligase BirA. Wild type BirA binds to a protein subunit of acetyl-CoA carboxylase, a biotin carboxyl carrier. Moreover, BirA forms an amide between a specific lysine side chain and biotin on the surface of a specific protein (Sears et al., 2019). The biotinylated lysine residue is biotin labelled with a 15 amino-acid peptide, which is also called the Avi-tag

(Li & Sousa, 2012). While the BirA ligase is specific to a substrate, the mutated version BirA (R118G) promiscuously biotinylates the proteins that the bait protein interacts with (Sears et al., 2019). Further, in the presence of biotin, BirA* can promiscuously biotinylate the proteins close to 10 nm radius of the bait protein. The high affinity between biotin and streptavidin creates an effective tool of biotin in labelling proteins, aiding the pull down of the biotinylated proteins for further analysis with mass spectrometry (Oostdyk et al., 2019).

There are several other proteomic approaches besides BioID such as TurboID, APEX or EMARS. TurboID, which has a shorter time period than BioID, with 10-minute incubation rather than 24 hours of robust biotinylation (May et al., 2020). APEX (engineered ascorbate peroxidase) on the other hand, has a labelling method in the presence of H2O2, oxidizing biotin-phenols to turn into biotin-phenoxyl radicals. These radicals are short lived and can biotinylate amino acids that are rich in electrons from a few nanometers (Nguyen et al., 2020). EMARS is an antibody-based approach, which is mostly applied in identifying the molecules on the cellular membrane (Honke, 2018).

1.7 Aim of the study

This study aims to find the interacting partners of a novel epigenetic modifier, SS18L2 in triple- negative breast cancer by BioID proteomic approach. The aim was subdivided into three objectives.

First objective is the generation of the fusion constructs, that include insert-BirA* fusion. The objective includes the validation of the fusion constructs SS18L2-BirA*, SS18-BirA* and SS18L1-BirA*, with functional assays and the overexpression of MDA-MB-231 cell line. Second objective is the biotinylation and the pull-down steps of the BioID procedure. The objective comprises of the optimization and confirmation of the MDA-MB-231 SS18L2 overexpressed cells, MDA-MB-231 SS18 overexpressed cells and MDA-MB-231 SS18L1 overexpressed cells with the related control groups. Finally, the identification of the possible interactome of SS18L2 with mass spectrometry was aimed. All three proteins were analyzed by mass spectrometry. The final objective of the study includes the comparison of the interactomes of the homologous proteins.

The study of the interaction partners of an uncharacterized protein SS18L2 in TNBC is crucial for deciphering the roles and characteristics of the protein. Due to the limited

research of this epigenetic modifier and its relation to its homologs, our findings could potentially provide a deeper understanding of SS18L2 in TNBC and lead a way for further investigations.

Chapter 2:
MATERIALS AND METHOD

2.1 Cloning of Insert-BirA Fusion Plasmids*

Inserts *SS18* and *SS18L1* were first amplified from U87 cDNA with PCR (polymerase chain reaction), Phusion® High-Fidelity PCR Kit (BioLabs). PCR conditions were (98°C 30 sec, 98°C 10 sec, 61-66°C 15 sec, 72°C 45sec, 72°C 5min, 4°C hold, 35X cycle). *SS18L2* was amplified from SS18L2 plasmid (DNAsu # HsCD00044038), which was a kind gift from Tamer Önder Lab. PCR conditions were (98°C 30 sec, 98°C 10 sec, 67°C 15 sec, 72°C 30 sec, 72°C 5 min, 4°C hold, 35X cycle). PCR amplifications were conducted with hot start. PCR products and the plasmids were cut with NheI (Neb) and EcoRI (Neb) for pcDNA3.1 MCS-BirA(R118G)-HA (CBioID, Addgene #36047) plasmid and inserts amplified accordingly, and XhoI (Neb) and EcoRI (Neb) for pcDNA3.1 mycBioID (NBioID, Addgene #35700) and inserts amplified accordingly. The enzyme cut reactions took place in Thermo Fisher Thermocycler, conditions were (37°C 2h, 65°C 20 min, 4°C hold). The products were purified with NucleoSpin Gel and PCR Clean-up Kit (Macherey-Nagel) according to the manufacturer's instructions. Moreover, the first reaction explained in **Figure 2.1** was conducted with the ligation of the cut N-C BioID plasmids with cut inserts. Ligation was conducted by T4 DNA Ligase Enzyme (Neb), with Thermo Fisher Thermocycler, conditions conditions (25°C 2h, 65°C 10 min, 4°C hold). The ligation products were than transformed to Stbl3 competent bacteria and the resulting plasmid was purified by NucleoSpin Plasmid MiniPrep Kit (Macherey-Nagel) according to the manufacturer's instructions. The first reaction will result in fusion proteins, *SS18-BirA*, SS18L1-BirA** and *SS18-BirA** in NBioID and CBioID plasmids.

For the second reaction, the fusion sequence was amplified, by adding new enzyme cut sites and tags with PCR. PCR conditions were (98°C 30 sec, 98°C 10 sec, 62-63°C 15 sec, 72°C 1 min 10 sec, 72°C 5 min, 4°C hold, 35X cycle). Annealing and extension times differ with different length of inserts. PCR amplifications were conducted with hot start. After amplification, inserts from C-BioID all have double HA tags and inserts from N-BioID all have double HA tags with a myc tag coming from the original plasmid. Amplified fusion sequences were cut with SalI (Neb) and XBaI (Neb) for the sequences

coming from both C-BioID and N-BioID. Moreover, the inserts are cloned into an entry vector to than be cloned into a lentiviral destination vector. The entry vector, pENTR1A no ccDB (w48-1) (pENTR1A, Addgene #17398) was cut with SalI (Neb) and XBaI (Neb) for the desired inserts. The enzyme cut reactions took place in Thermo Fisher Thermocycler, conditions were (37°C 2h, 65°C 20 min, 4°C hold). The products were purified with NucleoSpin Gel and PCR Clean-up Kit (Macherey-Nagel) according to the manufacturer's instructions. Furthermore, ligation of the amplified insert-*BirA** fusions was conducted by T4 DNA Ligase Enzyme (Neb), with Thermo Fisher Thermocycler, conditions (25°C 2h, 65°C 10 min, 4°C hold). The ligation products were transformed to Stbl3 competent bacteria, and the resulting plasmid was purified by NucleoSpin Plasmid MiniPrep Kit (Macherey-Nagel) according to the manufacturer's instructions. The second reaction will result in double HA tagged insert-*BirA** sequences in pENTRIA plasmid, different for both CBioID and NBioID constructs.

For the final reaction LR reaction or Gateway Cloning into a lentiviral vector pLEX_307 (Adddgene #41392) was conducted. LR cloning was performed with Gateway™ LR Clonase™ II Enzyme mix (Invitrogen) according to the manufacturer's instructions, with Thermo Fisher Thermocycler, conditions (25°C 16h, 4°C hold). Proteinase K was added to the reaction after 16h for 10 minutes at 37°C.

For a BirA* only control group, *BirA** only sequence was amplified from NBioID plasmid with Thermo Fisher Thermocycler, conditions (98°C 30 sec, 98°C 10 sec, 72°C 15 sec, 72°C 45 sec, 72°C 5 min, 4°C hold, 35X cycle). The insert was ligated to pENTRIA plasmid. The insert and the plasmid were cut with SalI (Neb) and XBaI (Neb). The insert with a myc tag was cloned into the lentiviral destination vector pLEX_307 as described in the method above. Detailed cloning of BirA* final plasmid is illustrated in **Figure 2.2** below. Primers that are utilized during the cloning procedures are given in the **Table 2.1** and primers to specific reactions were summarized in **Table 2.2.** Primers were designed with Dr. Alişan Kayabölen. All of the generated plasmids were confirmed by sanger sequencing and maps are reported in the (**Appendix A: Maps Generated for Fusion Plasmids**).

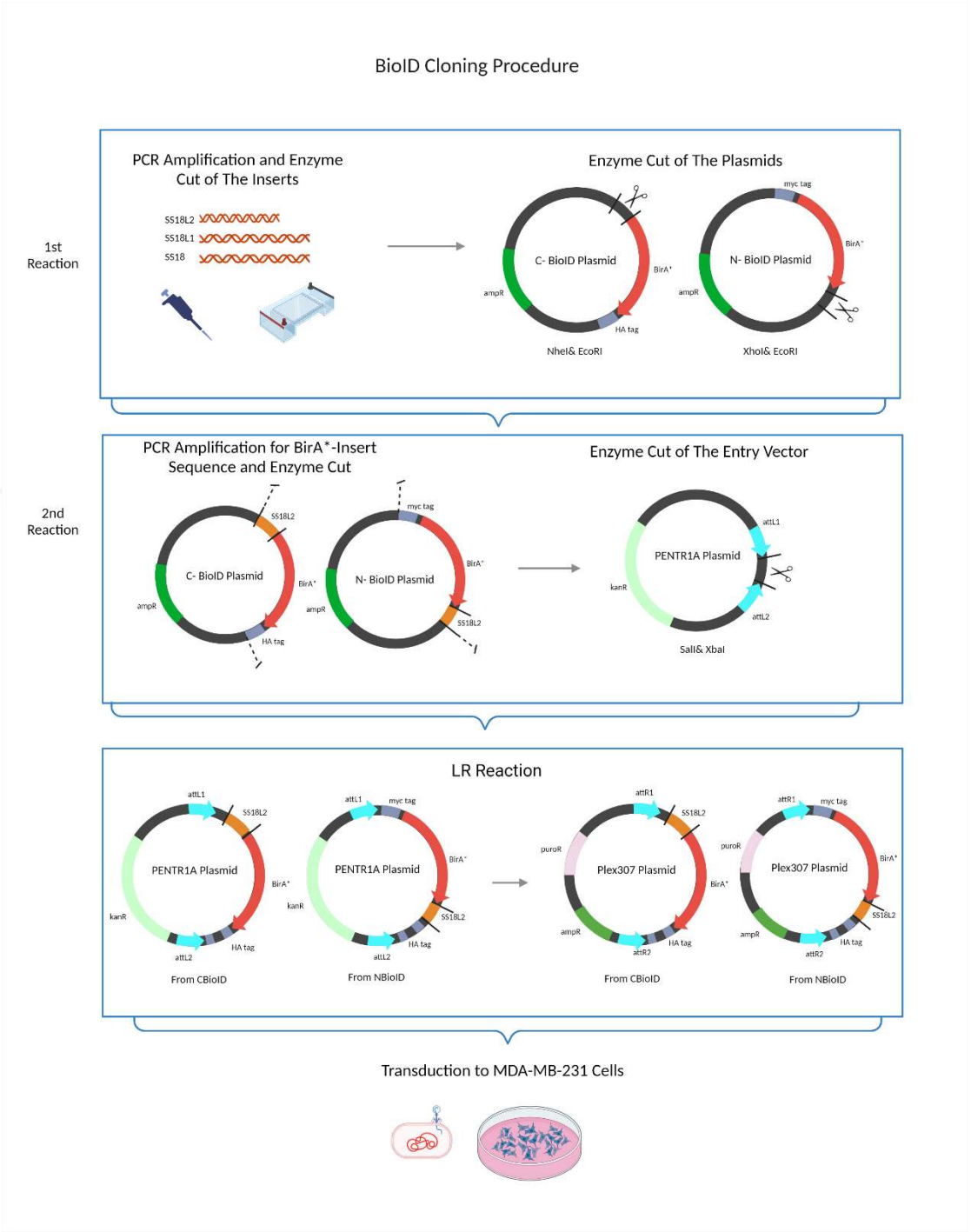

Figure 2.1. Schematic representation of the construction of the fusion plasmids. BirA* indicates mutated *BirA* gene, ampR indicates ampicillin resistance, kanR indicates kanamycin resistance and puroR indicates puromycin resistance genes. Dotted lines in second reaction indicate the sequence that is amplified by PCR. Figure created with BioRender.com.

Figure 2.2. Schematic representation of the construction of the BirA* only control fusion plasmid. BirA* indicates mutated *BirA* gene, ampR indicates ampicillin resistance, kanR indicates kanamycin resistance and puroR indicates puromycin resistance genes. Dotted lines in the first reaction indicate the sequence that is amplified by PCR. Figure created with BioRender.com.

Table 2.1. Primer sequences used in insert-BirA* fusion plasmids.

	PRIMER NAME	PRIMER SEQUENCE 5' TO 3'
1	SS18-F-NheI-Kozak	GTCGATGCTAGCGCCACC**ATGTCTGTGGCTTTCGCGGC**
2	SS18L1-F-NheI-Kozak	GTCGATGCTAGCGCCACC**ATGTCCGTGGCCTTCGCGTC**
3	SS18L2-F-NheI-Kozak	GTCGATGCTAGCGCCGCC**ATGTCGGTGGCCTTCGTACC**
4	SS18-R-EcoRI-linker (GS), SS18L1-R-EcoRI-linker (GS)	GTGCATGAATTCGCTGCC**CTGCTGGTAATTTCCATACTG**
5	SS18L2-R-EcoRI-linker (GS)	GTGCATGAATTCGCTGCC**TTCCATTGCTTTTGAAGTGC**
6	SS18-F2-SalI	GTGCATGTCGAC**GCCACCATGTCTGTGGCTTTCG**

7	SS18L1-F2-SalI	GTGCATGTCGAC**GCCACCATGTCCGTGGCCTTCG**
8	SS18L2-F2-SalI	GTGCATGTCGAC**GCCGCCATGTCGGTGGCCTTC**
9	SS18&L1&L2-R2-XbaI-HA-Stop	GTCGATTCTAGATTAgcctgcatagtccgggacatcatacggatagcc**tgcgtaatccggtacatcg**
10	SS18-F-XhoI-linker-BirA	GTGCATCTCGAGggctcc**TCTGTGGCTTTCGCGGCCC**
11	SS18L1-F-XhoI-linker-BirA	GTGCATCTCGAGggctcc**TCCGTGGCCTTCGCGTCTG**
12	SS18L2-F-XhoI-linker-BirA	GTGCATCTCGAGggctcc**TCGGTGGCCTTCGTACCG**
13	SS18-R-EcoRI-linker (GS)-HA-Stop, SS18L1-R-EcoRI-linker (GS)-HA-Stop	GTCGAT**GAATTC**TCAtgcgtaatccggtacatcgtaagggtaTCCGCT**CTGCTGGTAATTTCCATACTG**
14	SS18L2-R-EcoRI-linker (GS)-HA-Stop	GTCGATGAATTCTCAtgcgtaatccggtacatcgtaagggtaTCCGCT**TTCCATTGCTTTTGAAGTGC**
15	SS18&L1&L2-F2-SalI-BirA, SS18&L1&L2-R2-XbaI-HA-Stop	GTGGATGTCGAC**GCCACCATGgaacaaaaactcatc**
16	BirA-R-XbaI-Stop	TACGTCTAGAtcacttctctgcgcttctcaggg

Table 2.2. Primers and enzymes used in insert-BirA* fusion plasmids. Numbers correspond to the primer sequences indicated in **Table 2.1.**

CBioID		**Forward**	**Reverse**	**Enzyme**
1st Reaction	SS18	1	4	NheI and EcoRI
	SS18L1	2	4	
	SS18L2	3	5	
2nd Reaction	SS18	6	9	SalI and XbaI
	SS18L1	7	9	
	SS18L2	8	9	
NBioID		**Forward**	**Reverse**	**Enzyme**
1st Reaction	SS18	10	13	XhoI and EcoRI
	SS18L1	11	13	
	SS18L2	12	14	
2nd Reaction	SS18	15	9	SalI and XbaI
	SS18L1	15	9	
	SS18L2	15	9	

2.2 Cell Culture

MDA-MB-231 cell line was utilized throughout the experiments, which was a kind gift to our lab from Robert Weinberg (MIT, Boston, USA). The cells are cultured in DMEM (Gibco, USA) with %1 PS (Penicillin/Streptomycin, Gibco, USA) and 10% fetal bovine serum (Gibco, USA). The culture conditions are 37°C with 5°C CO_2 in a humidified incubator.

2.3 Immunofluorescence Staining (IF)

For immunofluorescence staining, cells were seeded on cover slips that were placed into 24-well plates. Cells were seeded as 100.000 cells per well. The next day, medium is removed, and cells are washed with PBS. Cells were moreover fixed with 4% PFA for 10 minutes in room temperature. PFA is then removed, and cells are washed with PBS 2 times. After the washings, cells are permeabilized with 1% Triton X-100 for 10 minutes in room temperature. The cells are washed with PBS again for 2 times. After removing PBS, cells are blocked with Superblock, and incubated in room temperature for 15 minutes. Coverslips are again washed with PBS 2 times. Cells were incubated with their related primary antibodies overnight at dark. After 16h, cells are incubated with the related secondary antibodies after 2 times PBS wash. The incubation lasts 1 hour in the dark. The coverslips are then mounted with DAPI (abcam). The images were taken at 40x with Zeiss Axio Imager M1. Antibodies utilized in the staining are listed in **Table 2.3** below.

Table 2.3. Primary and secondary antibody list utilized in immunofluorescence staining.

Antibody	Brand/Catalog Number	Dilution Rate
anti-HA Tag Antibody	Biolegend, 901513	1:2000
anti-MYCTtag Antibody	Protein-tech/16286-1-AP	1:400
53BP1	Novus/NB100-304	1:1000
Alexa Fluor® 594 Anti- IgG Secondary Antibody	Invitrogen	1:1000

Alexa Fluor® 488 Anti-IgG Secondary Antibody	Invitrogen	1:1000

2.4 *Lentiviral Packaging*

For viral packaging, 4 million Hek293T cells were seeded to 10 cm plates. Next day, 2250 ng psPAX2 (Addgene #12260), 225 ng VSV-G (Addgene #8454) and 2500 ng of the viral plasmid were prepared and 200 µl of serum free DMEM was added. After maximum 15 minutes of waiting, the DNA mixture was added to the PEI mixture with 15 µl of PEI and 185 µl serum free DMEM. After 30 minutes, all 400 µl of the mixtures were equally given to the 10 cm plates and left overnight. The medium was changed after 16 hours. The medium was collected 48 hours and 72 hours post transfection. The collected medium was filtered with 45-µm filters. The filtered medium containing viruses were added onto 1x PEG8000 and left overnight at +4°C. The next day, the viruses were centrifuged at 2500 rpm for 20 minutes, resuspended with PBS and concentrated for 100x. The concentrated viruses were stored at -80°C.

Viruses were transfected to cells in 6 well plates, 100 000 cell per well seeded the previous day. Transfection was conducted by adding 1 ml medium with PS (protamine sulfate) and infecting 20 µl of 100x concentrated viruses. 16 hours later, the medium is changed. The following day cells are introduced to 2 µg/ml puromycin, the infected cells survive the selection. The generated cells were amplified and frozen for further experiments.

2.5 *Western Blotting*

2.5.1 Protein Fractionation

The western blot experiments were conducted with the fractionation of the proteins into total, cytosolic and nuclear fractions. To obtain whole cell lysate (total fraction), pellets from MDA-MB-231 cells, MDA-MB-231 SS18L2 OE cells, MDA-MB-231 SS18 OE cells, MDA-MB-231 SS18L1 OE cells and MDA-MB-231 BirA* OE cells were lysed with RIPA Buffer containing 1mM PMSF, 1x Protease Inhibitor and 1x Phosphatase Inhibitor for 30 minutes. The samples were centrifuged at 13000xg for 10 minutes. The

supernatant was collected and stored at -80. For cytosolic lysis, pellets were resuspended in cytosolic lysis buffer (20 mM 7.9 pH HEPES, 10 mM KCl, 0.4% NP40, 0.1 mM EDTA and 1x Protease Inhibitor and 1x Phosphatase Inhibitor) and the samples wait on ice for 15 minutes. Samples were then centrifuged at 3000xg for 3 minutes. The supernatant was stored as the cytosolic fraction. Moreover, the remaining pellet was washed with the cytosolic buffer one more time and centrifuged at 3000xg for 3 minutes, supernatant was removed. The pellet was then resuspended with nuclear buffer (10 mM 7.9 pH HEPES, 0.4 M NaCl, 10% Glycerol, 1 mM EDTA and 1x Protease Inhibitor and 1x Phosphatase Inhibitor). The samples were sonicated on ice (2x10 sec, amplitude 40). Sonicated samples were centrifuged at 15000xg for 5 minutes. The supernatant was stored in -80 as the nuclear fraction. In order to determine the protein concentrations, Pierce's BCA Protein Assay Kit (Thermo Scientific, #23225, USA) was utilized according to according to the manufacturer's instructions.

2.5.2 Western Blot Protocol

40 µg of protein was mixed with 4X Laemmli Buffer (Bio-Rad, #1610747) containing 10% BME (Beta Mercaptoethanol). Samples were incubated at 95°C for 10 minutes. 4-12% Mini Protean TGX Precast Gels (Biorad, #456-1044) were utilized for the loading and the running of the samples. For the transfer of the proteins, PVDF membrane was utilized. Tranfer was conducted with Trans-Blot® Turbo™ Transfer System (Bio-Rad, USA), at 25 volt for 30 minutes. After transfer, the membrane was blocked with 5% Blotting-Grade Blocker (Bio-Rad, #1706404) nonfat dry milk for 1 hour in room temperature. After blocking, the membrane was washed with TBS-T for 3x15minutes. Membrane is then incubated with the respective antibodies overnight (16h) at +4°C. The membrane was further washed with TBS-T for 3x15minutes and incubated in secondary antibodies for 1 hour in room temperature. Membrane was again washed with TBS-T 3x15minutes and developed with Pierce™ ECL Western Blotting Substrate (Thermo-Fisher, # 32106) and imaged with Odyssey® Fc Imaging System by LI-COR Biosciences. Antibody list for all western blot experiments can be found table **Table 2.4** below.

Table 2.4. Primary antibody list utilized in western blotting.

Antibody	Brand/Catalog Number	Dilution Rate	Dilution Solution
anti-HA Tag Antibody	Biolegend/ 901513	1:1000	TBS-T
anti-MYCTtag Antibody	Protein-tech/16286-1-AP	1:1000	TBS-T
anti-Vinculin Antibody	Abcam/ ab73412	1:1000	TBS-T
anti- Histone H3 (D1H2) Antibody	Cell Signaling/ 4499	1:1000	TBS-T
anti- β Tubulin Antibody	Abcam/ ab6046	1:1000	TBS-T
anti- GLTSCR1 (H-10) Antibody	Santa Cruz/ sc-515086	1:1000	TBS-T
Streptavidin HRP Antibody	BioLegend/ 405210	1:1000	BSA

2.5.3 Western Blotting After Biotinylation or Pull Down

For the confirmation westerns after the biotinylation of the proteins, same procedure was applied. In the biotinylation western, washing steps are conducted with PBS to prevent excess binding of biotin. Moreover, for blocking 5%BSA in PBS was utilized. The antibody incubation is conducted with Streptavidin HRP Antibody without any primary antibody incubation, for 1 hour in room temperature. 60 µg protein is loaded.

For western blotting after pull-down experiments, same procedure was applied. Differently, washing steps are conducted with PBS and blocking was conducted with 5% BSA in PBS. The antibody incubation is conducted with Streptavidin HRP Antibody (**Table 2.4**) for 1 hour in room temperature. The loading conditions will further be explained in the **2.7** section in **Materials and Method**.

2.6 Biotinylation and Fractionation Before Mass Spectrometry

For biotinylation of the samples sent to mass spectrometry, 6500000 cells per 15 cm plates were seeded for yielding higher amounts of protein. 10 15 cm plates were seeded for every overexpressed and control cell lines. The next day, medium with 50μM D-Biotin (Supelco, #47868) will be introduced to the cells. After 24 hours of incubation with biotin, cells were collected with cell scrapers, to eliminate the degradation effect of trypsin. Every procedure must be on ice.

2.6.1 Protein Fractionation for Mass Spectrometry

The protein concentration was aimed to be minimum 4.5 mg of biotinylated proteins, to be pulled down and sent to mass spectrometry analysis. Pellets were fractionated with the procedure in section **2.5.1.** Due to the greater mass of proteins of the samples, pellets were collected in 15 cm falcons. After resuspension with the cytosolic buffer indicated in **2.5.1**, samples wait on ice for 15 minutes. Samples were then centrifuged at 3000rpm for 10 minutes. The supernatant is collected as the cytosolic fraction. Samples were washed with the cytosolic buffer and centrifuged at 3000rpm for 10 minutes. Supernatant was removed and the pellets were resuspended with nuclear buffer indicated in **2.5.1.** After resuspension, sonication was conducted with The Diagenode Biorupter® Sonicator (30 sec on 30 sec off, 30 cycle), a different sonicator due to higher mass. Moreover, the samples were centrifuged at 14000xg for 30 minutes at +4°C. Supernatant was taken as the nuclear fraction.

Every protein before sending them to mass spectrometry must be preserved on ice. Samples must not be stored at -80°C, to prevent the freeze thawing of the proteins. Protein concentrations were quantified by Pierce's BCA Protein Assay Kit (Thermo Scientific, #23225, USA) according to the manufacturer's instructions.

2.7 Pull Down Experiments

At the first day of the procedure, 4.5 mg of the biotinylated and fractionated proteins were calculated from each sample, and the respective buffers (cytosolic or nuclear) was completed to 1 ml for each sample, preparing a protein and a lysis buffer mix. On the other hand, Pierce™ Streptavidin UltraLink™ Resin (Thermo-Fisher, #53114) Volume

of the beads were calculated by 10 µl of beads for every 250 µg of protein. Beads were washed with cytosolic or nuclear buffers depending on the samples for 3 times, centrifuged at 1000xg for 2 minutes each time and supernatant is removed. The protein and buffer solutions were added onto the washed beads. The samples stay overnight in the rotator at the cold room, +4°C.

The next day, 16 hours later, the samples were centrifuged at 1000xg for 5 minutes, the supernatant was removed. Beads should not be disturbed during the process. Since beads are bound to proteins, the bead and protein complexes were washed with their respective buffers (cytosolic or nuclear) and were rotated for 8 minutes in the rotator. Then the samples were centrifuged at 1000xg for 2 minutes. The washing steps were repeated 4 times.

If the samples are being sent to mass spectrometry, the procedure at **2.8** will be followed. If the samples are not being sent to mass spectrometry, and confirmation westerns were to be conducted, the supernatant would be removed from the beads. 4X Laemmli buffer would be added, and the samples would be boiled at 95°C for 10 minutes. The resulting samples were centrifuged for 1000xg for 2 minutes and the supernatant can be restored at -80°C or directly loaded for western blot experiments.

2.8 Mass Spectrometry Analysis

After four repeated washes in section **2.7**, 100 µl of 100 mM ammonium bicarbonate, which contains 55mM iodoacetamide was added onto the beads. The samples were incubated for 1 hour in a dark room. Moreover, the samples were incubated with trypsin and incubated at 37°C overnight. After 16 hours, formic acid was added onto the samples ans the samples were centrifuged at 14000xg for 30 minutes. The supernatant was collected to be analyzed in the mass spectrometry. Mass spectrometry analysis was conducted with Thermo Scientific Q Exactive HF-X Hybrid Quadrupole-Orbitrap Mass Spectormeter. LC-MS/MS analysis was conducted in collaboration with Prof. Nurhan Ozlu and her team.

2.9 Bioinformatic Analysis of the Mass Spectrometry Results

The raw data obtained from the mass spectrometry analysis was utilized to analyze the peptid counts with MaxQuant (Max Planck Institute of Biochemistry) software. The

parameters “match between runs”, “LFQ”, “unique+razor” and “iBAQ” were enabled and “fast LFQ” was disabled. The digestion parameter was chosen to be trypsin. Other parameters were set as default.

After the analysis, the output which contains the protein groups detected was analyzed with Cassiopeia LFQ software. Cassiopeia is an open-source software, which is an R script that analyses the workflow in Max Perutz Labs mass spectrometry facility in Vienna. With the software, statistical label free quantification was analyzed. All groups were normalized relative to the control group, eliminating any known contaminant proteins. The protein quantities are identified with at least two distinct peptides, which were considered statistically significant, if correlates with the calculated p values. Proteins with LogFC values >1.5 and p values<0.05 were considered as the top interacting partners. The output interactomes were also analyzed in STRING and GO platforms to correlate the data with the known pathways. Proteins with LogFC values>1.5 were analyzed in the STRING platform.

Graphical figures were created in Python and R software.

Chapter 3:
RESULTS

3.1 Generation of SS18L2_BirA Fusion Protein and its Functional Validation*

3.1.1 Cloning of the SS18L2-BirA Fusion Protein*

In order to construct the fusion proteins needed for BioID a three stepped cloning method was applied as described in the methods section, **2.1.** *SS18L2* was first cloned into C-BioID and N-BioID plasmids. After amplifying the SS18L2-BirA* fusion sequence and adding HA Tags during the amplification, the fused insert was cloned into an entry vector, pENTR1A plasmid. The SS18L2-BirA* insert that is located between the attL1 and attL2 sites is cloned into the destination vector PLEX307, between aatR1 and attR2 sites with LR (gateway) cloning. With the same methodology, *SS18* and *SS18L1* inserts were cloned into N-BioID plasmid and into the PLEX307 lentiviral vector as a final step.

As a positive control, a destination vector which only includes *BirA** without any inserts is needed. The cloning procedure of BirA* only overexpression plasmid is described in the methods section, **2.1.** The *BirA** gene was first amplified from N-BioID backbone, amplifying the myc-tag sequence as well. The amplified sequence was then cloned into an entry vector, pENTR1A plasmid. The sequence was later then cloned into PLEX307 plasmid with LR cloning as the final lentiviral vector.

The plasmids that were engineered during the cloning process are demonstrated in the **Table 3.1** below.

Table 3.1. The Insert-BirA* plasmids generated after the cloning procedure. Steps of the cloning are assigned as the cloning steps mentioned in **Figure 2.1** and **Figure 2.2**.

Number	Generated Plasmid	Cloning Reaction Steps	Procedure Figure
1	NBioID-SS18L2 plasmid	First Reaction	Figure 2.1

2	SS18L2-NBioID-pENRTR1A plasmid	Second Reaction	Figure 2.1
3	SS18L2-NBioID-PLEX307 plasmid	Third Reaction (LR Cloning)	Figure 2.1
4	CBioID-SS18L2 plasmid	First Reaction	Figure 2.1
5	SS18L2-CBioID-pENRTR1A plasmid	Second Reaction	Figure 2.1
6	SS18L2-CBioID-PLEX307 plasmid	Third Reaction (LR Cloning)	Figure 2.1
7	NBioID-SS18 plasmid	First Reaction	Figure 2.1
8	SS18-NBioID-pENRTR1A plasmid	Second Reaction	Figure 2.1
9	SS18-NBioID-PLEX307 plasmid	Third Reaction (LR Cloning)	Figure 2.1
10	NBioID-SS18L1 plasmid	First Reaction	Figure 2.1
11	SS18L1-NBioID-pENRTR1A plasmid	First Reaction	Figure 2.1
12	SS18L1-NBioID-PLEX307 plasmid	Third Reaction (LR Cloning)	Figure 2.1
13	BirA*-NBioID-pENTR1A plasmid	First Reaction	Figure 2.2
14	BirA*-NBioID-PLEX307 plasmid	Final Reaction (LR Cloning)	Figure 2.2
15	CBioID-SS18 plasmid	First Reaction	Figure 2.1

3.1.2 Overexpression and the Validation of the SS18L2-BirA Fusion Protein*

After all overexpression (OE) plasmids were cloned into a lentiviral backbone, MDA-MB-231 cells were transduced with the OE plasmids SS18L2-NBioID- PLEX307 and SS18L2-CBioID- PLEX307. Hereafter, the OE plasmids will be denoted as "SS18L2-NBioID" and "SS18L2-CBioID" for the purpose of simplification. To test whether the cloning procedure was successful, western blot experiment from the cell pellets of the SS18L2-NBioID and SS18L2-CBioID was conducted. The cell pellets were fractionated as Total, Cytosolic and Nuclear fractionations. The SS18L2-NBioID fusion protein is 48 kDa while SS18L2-CBioID fusion protein is 46 kDa. Non-transfected MDA-MB-231 cell line (MDA Parental) was utilized as a control group, since the cell line does not endogenously express the engineered fusion protein. The cloning and overexpression were validated by western blot in **Figure 3.1**.

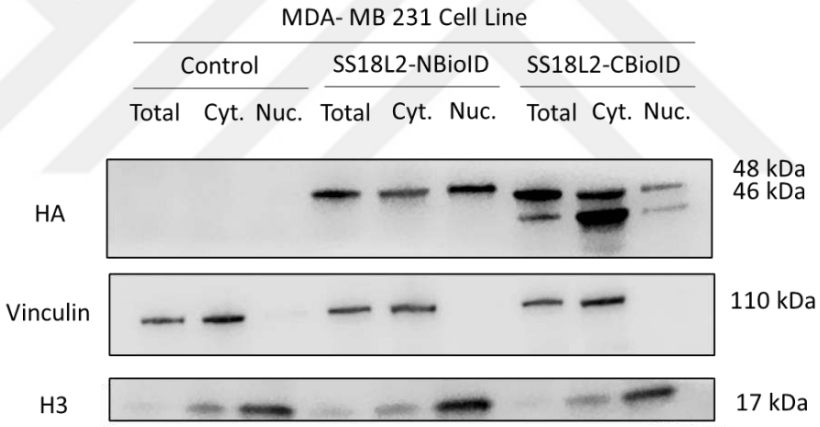

Figure 3.1. Western blot confirmation for SS18L2-NBioID and SS18L2-CBioID overexpressed cells. Anti-HA Tag antibody is used to detect the SS18L2-BirA* fusion proteins. Cyt. refers to cytosolic fraction and Nuc. refers to nuclear fraction. Vinculin (110 kDa) was used as a cytosolic control, H3 (17 kDa) was used as a nuclear control. 40 µg protein loaded.

MDA-MB-231 parental cell line did not give a signal at the sizes of the fusion proteins as expected. Cytosolic control can only be seen in total and cytosolic fractions while the nuclear control can be seen abundantly in the nuclear fraction. The blot showed bands at 48 kDa for SS18L2-NBioID fusion protein, confirming the cloning of the fusion. Bands were visible in all total, cytosolic, and nuclear fractionations. However, there were multiple bands visible for SS18L2-CBioID fusion protein, at 46 kDa and below. Moreover, the results do not correlate with SS18L2-NBioID in terms of the total,

cytosolic and nuclear levels. Different bands and the differences in the blots led to the conclusion that SS18L2-CBioID construct might be degrading, which will further be discussed. Therefore, the further experiments were continued with the N-BioID constructs for all the fusion proteins.

Cloning of the BirA* only control SS18-NBioID and SS18L1-NBioID fusion was also confirmed with the OE of MDA-MB-231 cell line. The western blots for the confirmations are indicated in **Figure 3.2** and **Figure 3.3**.

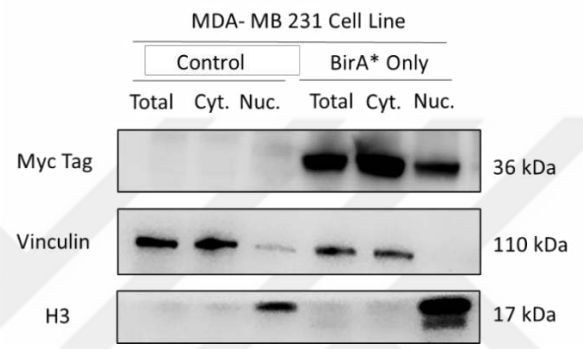

Figure 3.2. Western blot confirmation for BirA*-NBioID overexpressed cells. Anti-Myc Tag antibody is used to detect the BirA* protein. Cyt. refers to cytosolic fraction and Nuc. refers to nuclear fraction. Vinculin (110 kDa) was used as a cytosolic control, H3 (17 kDa) was used as a nuclear control. Fusion proteins were blotted with anti HA-Tag antibody. 40 µg protein loaded.

The blot in **Figure 3.2** showed bands at 36 kDa for BirA*-NBioID fusion protein, confirming the cloning of BirA* only control group. Bands were visible in all total, cytosolic, and nuclear fractionations.

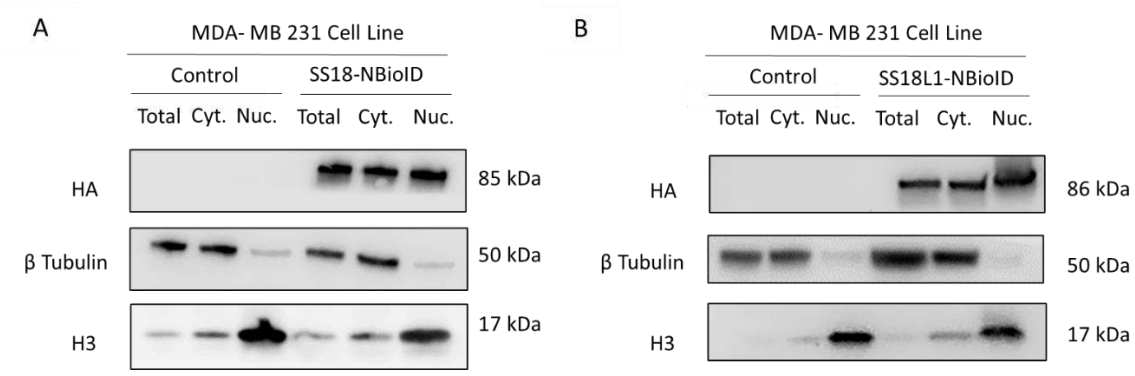

Figure 3.3. A, B. Western blot confirmation for SS18-NBioID and SS18L1-NBioID overexpressed cells. Anti-HA Tag antibody is used to detect the SS18-BirA* and SS18L1-BirA* fusion proteins. Cyt. refers to cytosolic fraction and Nuc. refers to

nuclear fraction. β Tubulin (50 kDa) was used as a cytosolic control, H3 (17 kDa) was used as a nuclear control. 40 µg protein loaded.

Similar to the experiment conducted for the validation of SS18L2-BirA* fusion protein, the blots in **Figure 3.3 A** and **B** demonstrated visible bands at 85 kDa for SS18-NBioID fusion protein, and at 86 kDa for SS18L1-NBioID fusion protein confirming the cloning of the fusion. Bands were visible in all total, cytosolic, and nuclear fractionations as expected.

All the blots confirm that the generation of the fusion proteins and the expression of SS18L2-NBioID, SS18-NBioID and SS18L1-NBioID in MDA-MB-231 cells were successful. BirA* only OE plasmid was also cloned, and the protein can be demonstrated in the western blot.

3.1.3 The Localization of SS18L2 to Nucleus

Since the localization of SS18L2 was not reported in literature before, immunofluorescence staining was conducted to examine SS18L2 localization in the cell. First, the localizations of SS18L2-NBioID and SS18L2-CBioID were illustrated under fluorescence microscopy with anti-HA Tag antibody (with the help of Nareg Pınarbaşı-Değirmenci) (**Figure 3.4**). The fluorescence microscopy image exposures were normalized to control, parental MDA-MB-231 cell line stained with anti- HA Tag antibody. It was indicated that SS18L2-NBioID fusion protein localizes mainly in the nucleus, as the green fluorescence for anti-HA Tag antibody can be seen in the nucleus abundantly. Moreover, there were no fluorescence signals in the nucleolus and only lower signals were present in the cytosol. These results suggested that SS18L2 fusion protein mainly localizes to the nucleus, excluding nucleolus; and based on the low cytosolic signals, SS18L2 might have a role in the cytosol as well. Similarly, SS18L2-CBioID resulted in the same pattern with SS18L2-NBioID, being mainly localized in the nucleus but not in the nucleolus (**Figure 3.4**). The immunofluorescence results of SS18L2-CBioID and SS18L2-NBioID indicated similar localization of these two fusions. Although their localizations were similar, there were major differences in the proteins biotinylation, where possible degradation products of the SS18L2-CBioID construct were observed in western blots (**Figure 3.1**). Therefore, we chose to continue mass

spectrometry experiments with the N-BioID constructs, which displayed proper subcellular localization, efficient biotinylation and pull down.

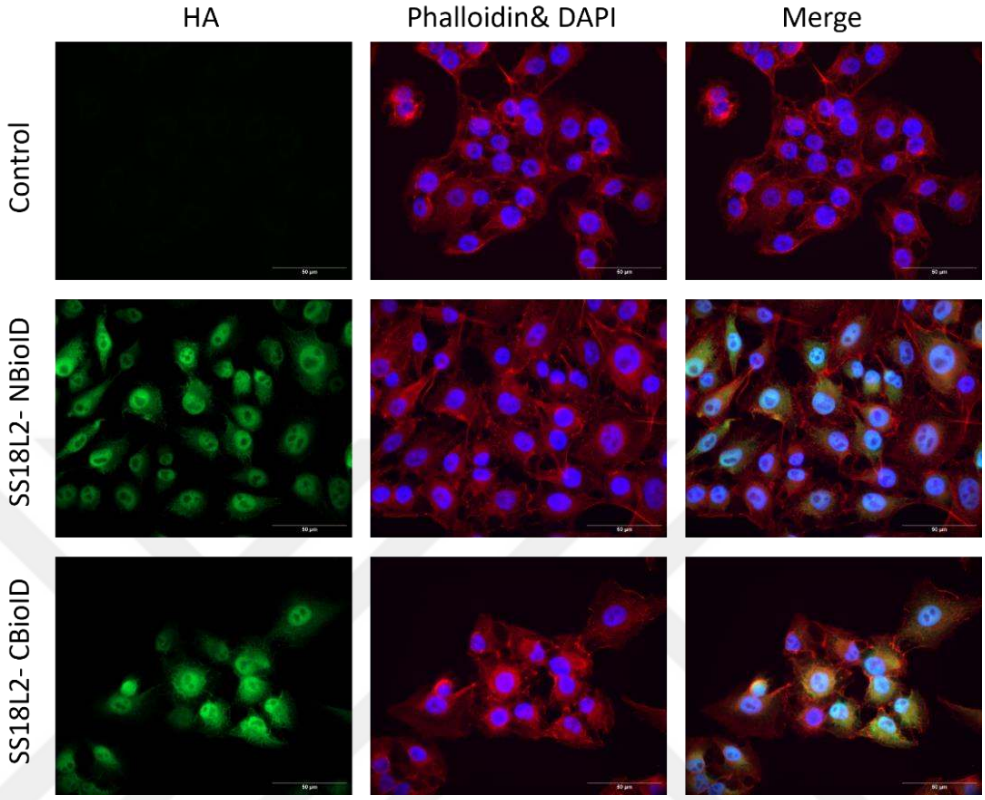

Figure 3.4. Immunofluorescence staining indicating the nuclear localization of SS18L2. MDA-MB-231 cell overexpressing SS18L2-NBioID and SS18L2-CBioID fusion constructs were stained with anti- HA Tag antibody (green), Phalloidin (red staining actin) and DAPI (blue staining nucleus). Un-transfected cells were used as control. Scale bar 50 μm, normalized to control.

After demonstrating the localization of SS18L2, the localization of BirA* was examined with immunofluorescence staining. Since BirA* biotinylates every protein that the bait protein interacts with, BirA* alone will promiscuously biotinylate any protein that is localized within 10 nm. Therefore, the staining was expected to be seen scattered in the cell, localizing both in the cytoplasm and the nucleus. The staining of the cells with anti-Myc Tag antibody resulted in the expected staining pattern, encompassing every part of the cell (**Figure 3.5).** The experiments were conducted in different times and with different antibodies; therefore, signals of SS18L2-NBioID (**Figure 3.4**) and BirA*-NBioID (**Figure 3.5)** are not comparable.

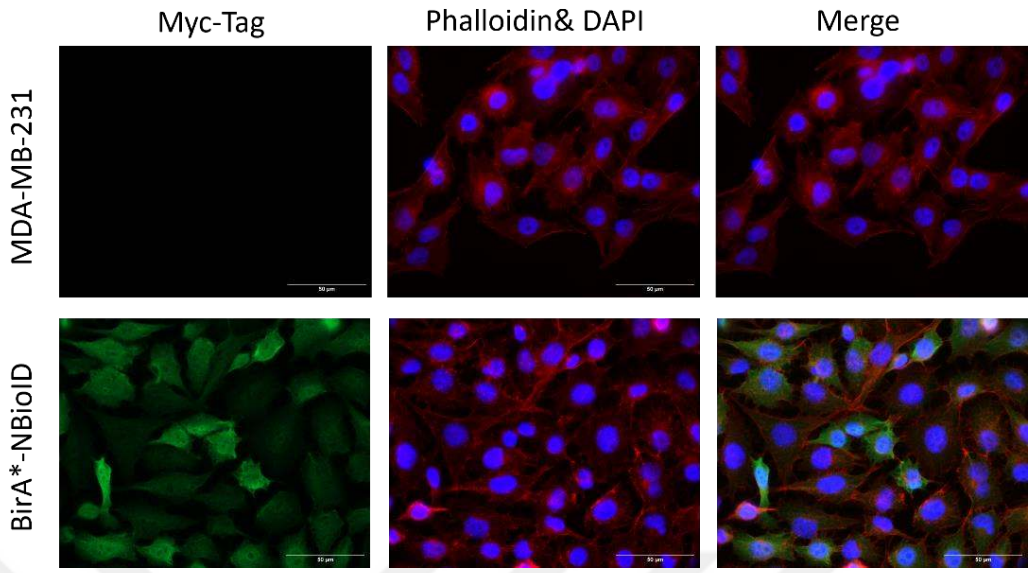

Figure 3.5. Immunofluorescence staining indicating the localization of BirA* scattered in the cell. MDA-MB-231 cell overexpressing BirA*-NBioID construct were stained with anti- Myc Tag antibody (green), Phalloidin (red staining actin) and DAPI (blue staining nucleus). Un-transfected cells were used as control. Scale bar 50 μm, normalized to control.

After the demonstration of the localization of SS18L2 into nucleus and confirming the localization of the BirA* only control group, the localization of SS18-NBioID and SS18L1-NBioID were examined under fluorescence microscopy with anti-HA Tag antibody (**Figure 3.6**). The fluorescence microscopy images were normalized to control, parental MDA-MB-231 cells were stained with anti- HA Tag antibody. As expected, and known from the literature before, SS18 and SS18L1 both showed localizations in the nucleus, with anti- HA Tag antibody giving high nuclear signals.

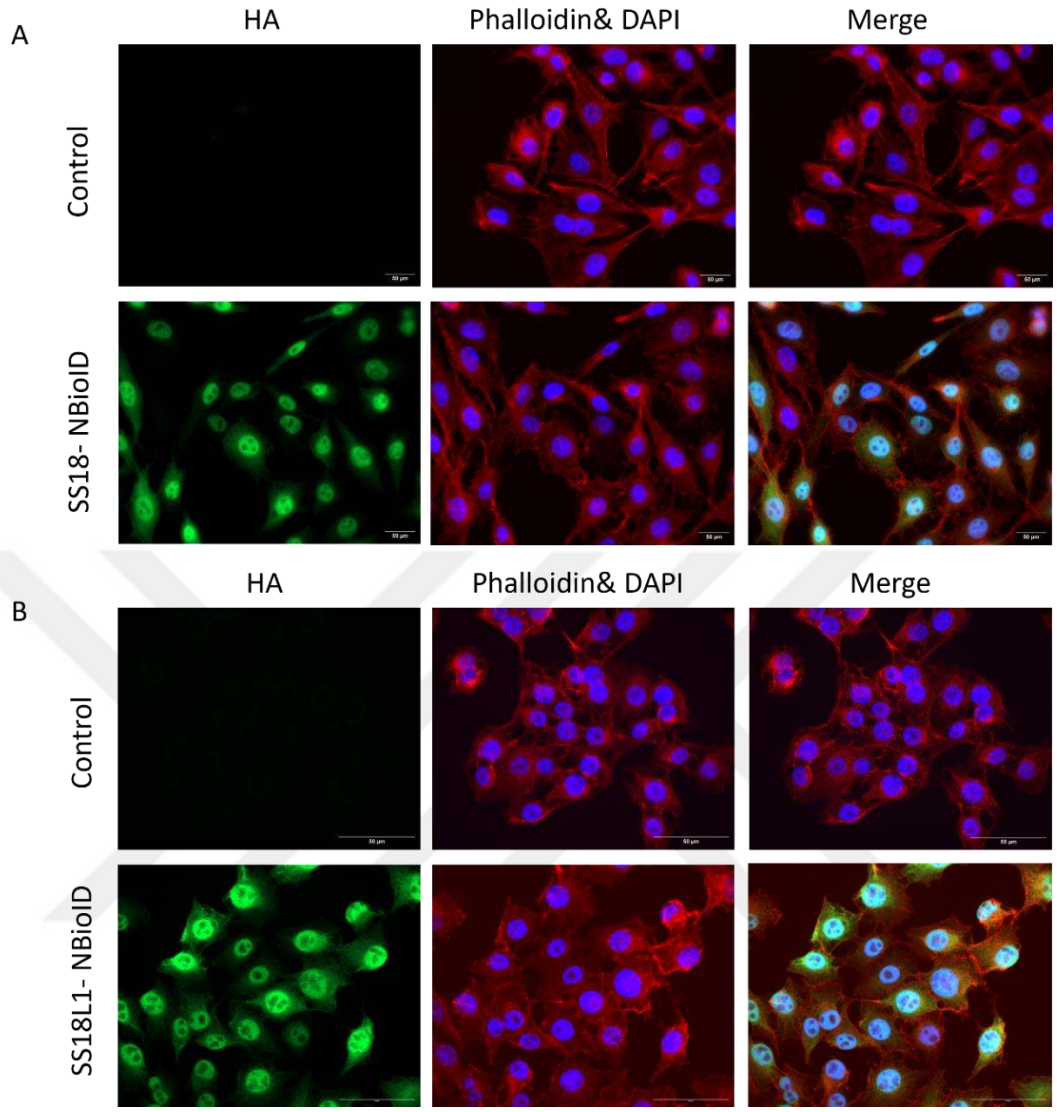

Figure 3.6. A, B. Immunofluorescence staining indicating the nuclear localization of SS18 and SS18L1. MDA-MB-231 cell overexpressing SS18-NBioID and SS18L1-NBioID fusion constructs were stained with anti- HA Tag antibody (green), Phalloidin (red staining actin) and DAPI (blue staining nucleus). Un-transfected cells were used as control. Scale bar 50 µm, normalized to control.

3.2 Validation of the Biotinylation and Streptavidin Pull Down

3.2.1 Biotinylation Optimizations and Validations of the Overexpressed Cells

After it was shown that all fusion proteins were constructed, the next step was to confirm that the overexpressed cells can be biotinylated in the presence of a biotin containing medium. Biotinylation was first determined with transfection in HEK 293T cells (**Figure 3.7A**). With streptavidin HRP, un-transfected HEK 293T cells did not give a dark blotted signal, while SS18L2-NBioID overexpressed HEK 293T cells with

biotinylation resulted in a high signal, indicating the cells were biotinyled and streptavidin HRP can bind to the biotinylated cells. SS18L1-NBioID OE HEK 293T cells also gave a high signal. Moreover, PLK4-BioID OE cells, kind gift from TO lab as a control group, for which biotinylation was confirmed before, gave a high signal as well. The results indicated that HEK 293T OE cells can be successfully biotinylated with transfection.

After confirming biotinylation with transfection in HEK 293T cells, biotinylation in MDA-MB-231 cells with transduction was conducted. Cells were transduced with SS18L2-NBioID and SS18L2-CBioID, to see if there are any differences in biotinylation as well (**Figure 3.7B**). The transfected HEK293T cells with SS18L2-NBioID were used as a control. The blot resulted in a high signal for the positive control group, which suppresses the signals of other samples. However, biotinylation can still be identified from the blot, showing many of the unknown proteins were biotinylated compared to the negative control. Since the interacting proteins of SS18L2 are not known, another control protein, such as a loading control, could not be utilized in the experiments.

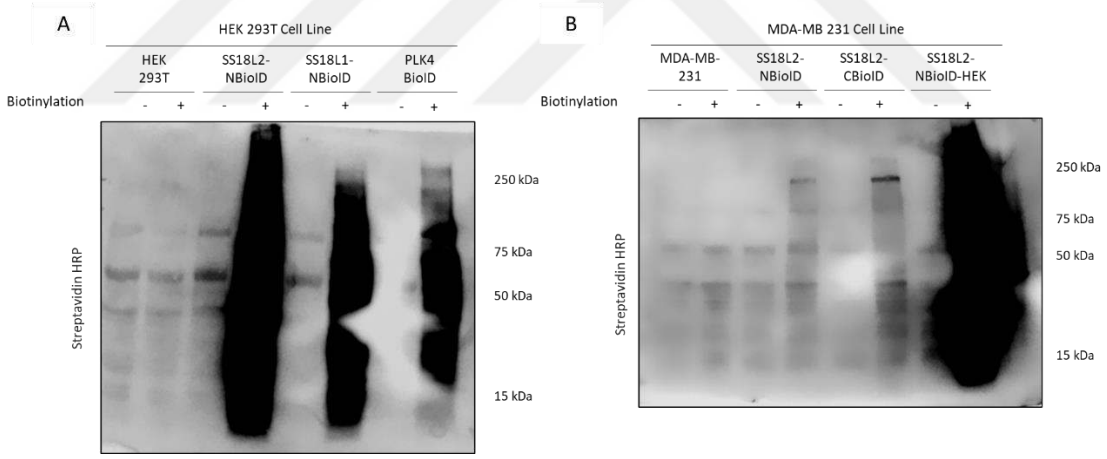

Figure 3.7. Western blot for biotinylation trials with HEK 293T and MDA-MB-231 cell lines. A. Biotinylation after transfection of SS18L2-NBioID and SS18L1-NBioID constructs to HEK 293T cells. Total lysis. Plk4 is utilized as a positive control as a kind gift from TO Lab. Un- transfected HEK 293T cells are utilized as control. 80 µg protein loaded. Blotted with streptavidin HRP antibody. **B.** Biotinylation after transduction of SS18L2-NBioID and SS18L2-CBioID constructs to MDA-MB-231 cells. Total lysis. Transfection of SS18L2-NBioID to HEK 293T cells were utilized as control. Un-transfected MDA-MB-231 cells were used as control. 60 µg protein loaded. Blotted with streptavidin HRP antibody. Biotinylated cells were incubated in 50 µM biotin medium. "+", "-" indicates the presence of biotinylation.

The biotinylation rates with transduction were lower than they were with transfection. This can be expected, because of the higher expression levels of the plasmid when it is introduced into the cell directly. However, the biotinylation rates in the western blot in **Figure 3.10B** still seemed lower, as many of the proteins were expected to be biotinylated when SS18L2 fusion is overexpressed. To test whether the slight bands occur because of the high signaled control or if SS18L2 has fewer interacting proteins, which will result in lower biotinylation levels, same procedure was applied to in MDA-MB-231 cells by overexpressing SS18L2-NBioID and SS18L2-CBioID without the transfection control (**Figure 3.8**). The blot indicated that biotinylation does occur in both OE cells, leading to a conclusion that biotinylation with transduction can be achieved. However, SS18L2-CBioID was still checked in the blot for the expression levels of biotinylation. SS18L2-CBioID cells seem to be biotinylating more, this may be the cause of the degraded proteins biotinylating various other proteins than the exact interactome of SS18L2. Therefore, the blot supported the decision to continue with SS18L2-NBioID and other NBioID constructs for further experiments.

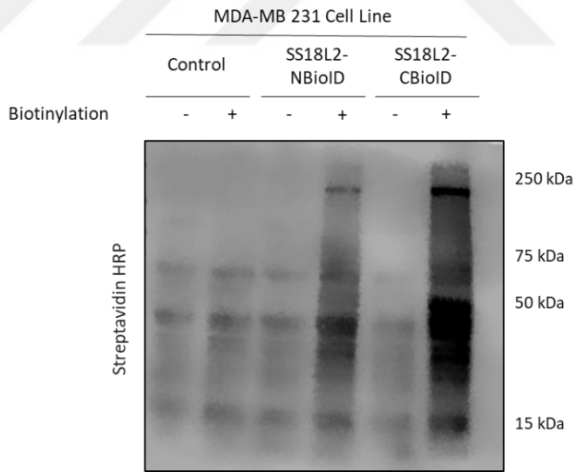

Figure 3.8. Western blot for biotinylation trials with only MDA-MB-231 cell line cytosolic lysis. Biotinylation after transduction of SS18L2-NBioID and SS18L2-CBioID constructs to MDA-MB-231 cells. Cytosolic fractionation. Un-transfected cells were used as control. 60 µg protein loaded. Blotted with streptavidin HRP antibody. Biotinylated cells were incubated in 50 µM biotin medium. "+", "-" indicates the presence of biotinylation.

To check if the biotinylation rates can be increased with double infection, and to further confirm the degree of biotinylation of SS18L2, SS18L2-NBioID and SS18L2-

CBioID, OE cells were infected twice before adding biotin to the medium (**Figure 3.9**). The results were in line with the previous blot (**Figure 3.8**), in terms of the differences between SS18L2-NBioID and SS18L2-CBioID. The level of biotinylation in SS18L2-NBioID was again similar to what was demonstrated before, suggesting that the interaction partners of SS18L2 might be low. In addition, the blot was utilized to try a freshly prepared biotin for further experiments, which is shown to be biotinylating the cells and can be utilized.

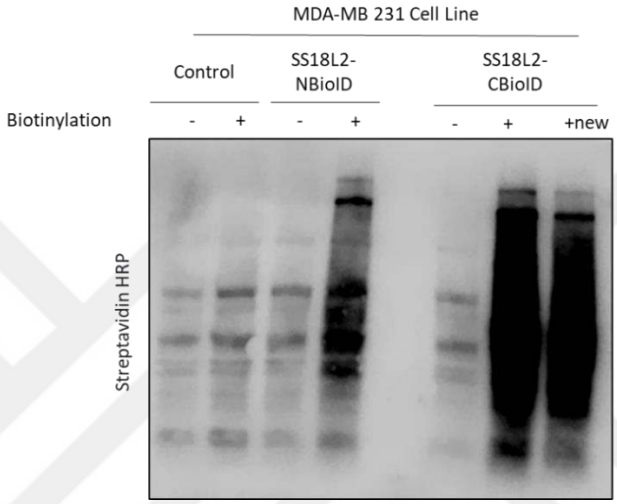

Figure 3.9. Western blot for biotinylation trials with double infection. Biotinylation after double infection with transduction of SS18L2-NBioID and SS18L2-CBioID constructs to MDA-MB-231 cells. Total lysis. Un-transfected cells were used as control. Second infection was conducted 24 hours after first infection. 60 µg protein loaded. Blotted with streptavidin HRP antibody. Biotinylated cells were incubated in 50 µM biotin medium. "+", "-" indicates the presence of biotinylation. "+ new" labeled lane indicates freshly prepared biotin.

After the optimization of biotinylation, MDA-MB-231 cells were transduced with SS18L2-NBioID and SS18L2-CBioID. First mass spectrometry was planned to be a pilot experiment; therefore, both cytosolic and nuclear fractions of overexpressed cells were examined to decide which fraction will be needed later on. Therefore, the pilot mass spectrometry analyses were conducted as cytosolic and nuclear fractionations (**Figure 3.10A, B**). The blots both showed a high biotinylation rate in BirA*-NBioID OE cells, which was expected due to the promiscuous biotinylation of BirA*. Moreover, it is possible to conclude that SS18L2-NBioID OE cells were biotinylated by showing a smeared band in the blot. BirA* only OE non-biotinylated cells also show a slight smear,

due to the fact that BirA* alone can biotinylate other proteins even biotin is not present because of its high biotin ligase activity.

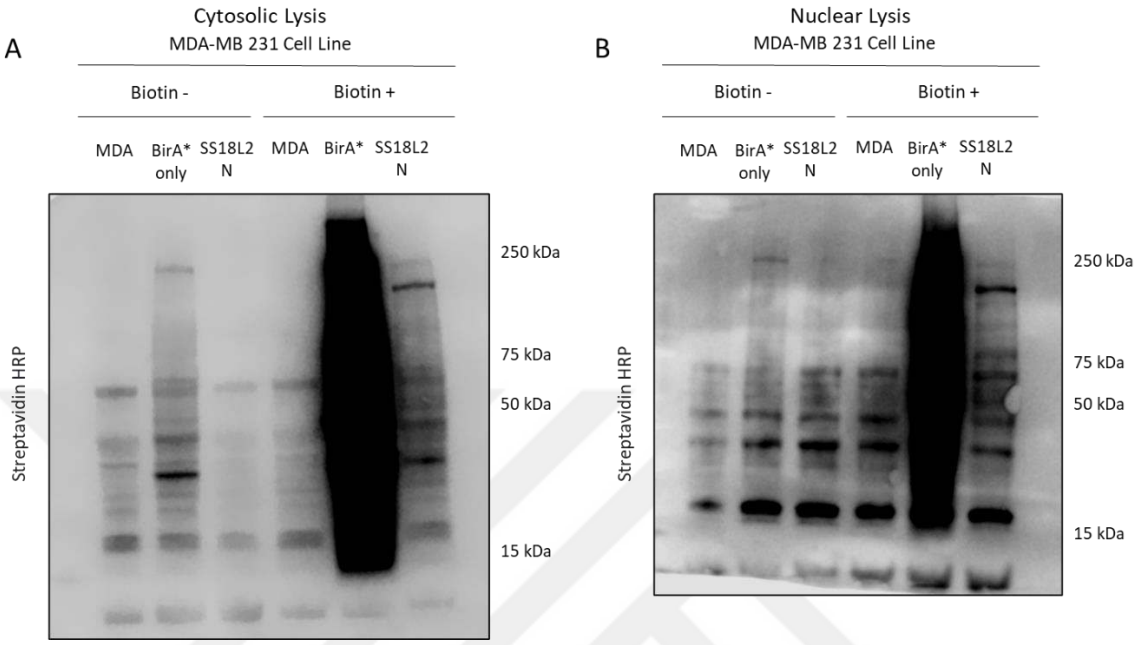

Figure 3.10. Western blot for biotinylation validations for SS18L2-NBioID and BirA* overexpressed cells. A. Biotinylation after transduction of SS18L2-NBioID and BirA*-NBioID constructs to MDA-MB-231 cells. Cytosolic fractionation. Un-transfected and non- biotinylated MDA-MB-231 cells, non-biotinylated BirA*-NBioID and SS18L2-NBioID overexpressed cells are used as control. 60 µg protein loaded. Blotted with streptavidin HRP antibody. Biotinylated cells were incubated in 50 µM biotin medium. **B.** Nuclear fractionation.

After the validations of the biotinylations of SS18L2-NBioID and BirA*-NBioID OE cells, the biotinylations of SS18-NBioID (**Figure 3.11**) and SS18L1-NBioID (**Figure 3.12**) OE cells were also validated. The figures indicate a smeared signal, showing the biotinylation of many unknown proteins. The western blots showed that in both cytosolic and nuclear fractions, SS18-NBioID and SS18L1-NBioID OE cells can also be biotinylated.

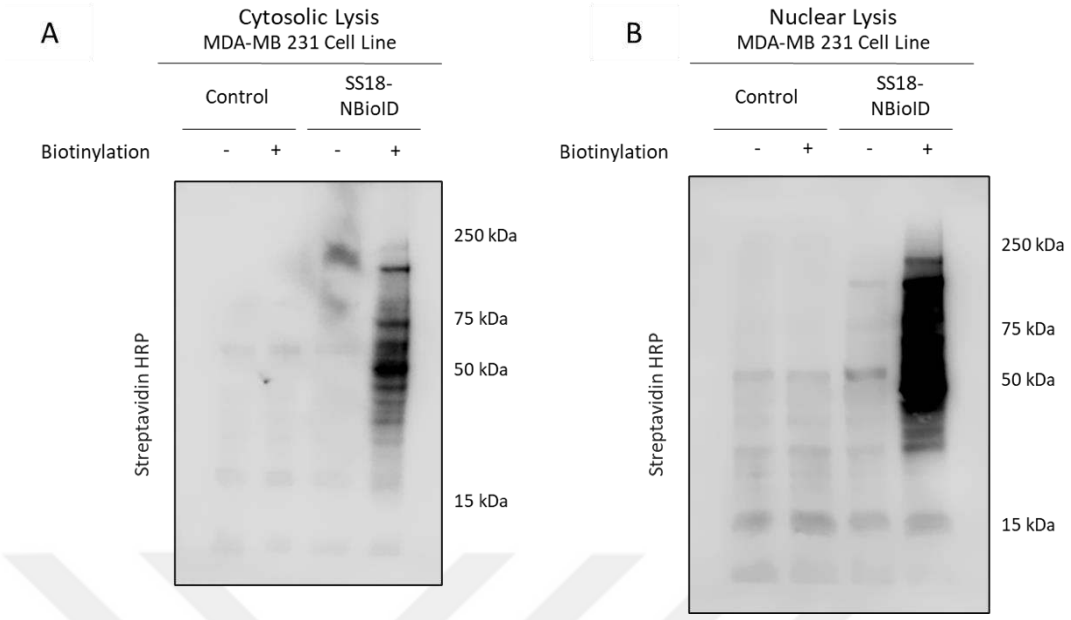

Figure 3.11. Western blot for biotinylation validations for SS18-NBioID overexpressed cells. A. Biotinylation after transduction of SS18-NBioID construct to MDA-MB-231 cells. Cytosolic fractionation. Un-transfected and non- biotinylated MDA-MB-231 cells and non-biotinylated SS18-NBioID overexpressed cells are used as control. 60 µg protein loaded. Blotted with streptavidin HRP antibody. Biotinylated cells were incubated in 50 µM biotin medium**.** "+", "-" indicates the presence of biotinylation **B.** Nuclear fractionation.

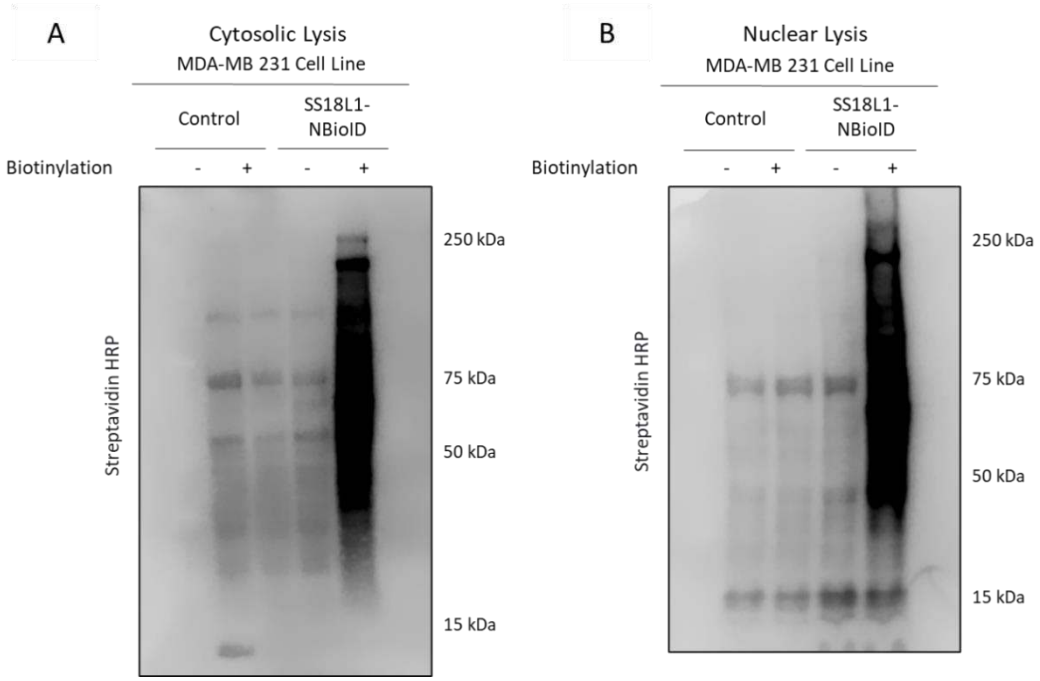

Figure 3.12. Western blot for biotinylation validations for SS18L1-NBioID overexpressed cells. A. Biotinylation after transduction of SS18L1-NBioID construct to MDA-MB-231 cells. Cytosolic fractionation. Un-transfected and non- biotinylated MDA-MB-231 cells and non-biotinylated SS18L1-NBioID overexpressed cells are used

as control. 60 µg protein loaded. Blotted with streptavidin HRP antibody. Biotinylated cells were incubated in 50 µM biotin medium. "+", "-" indicates the presence of biotinylation **B.** Nuclear fractionation.

3.2.2 Pull-Down validations

The biotinylated proteins are needed to be pulled down by Streptavidin Beads due to the high affinity of Streptavidin to Biotin. First, the pull-down procedure was optimized, and was tested with a total fraction of SS18L2-NBioID OE MDA-MB-231 cells (**Figure 3.13A**). The experiment was confirmed with a western blot, seeing a black smear similar to the biotinylation figures. The experiment was conducted with Streptavidin HRP antibody, and the biotin labelled proteins were detected. It was shown that the pull-down experiments were successful. Moreover, with the same samples, SS18L2 was blotted (**Figure 3.13B**), to see whether it was possible to pull down the protein of interest, which is expected be one of the most biotinylated proteins in the experiment due to its fusion to BirA*. A band in 48 kDa was observed, verifying that SS18L2 was also pulled down.

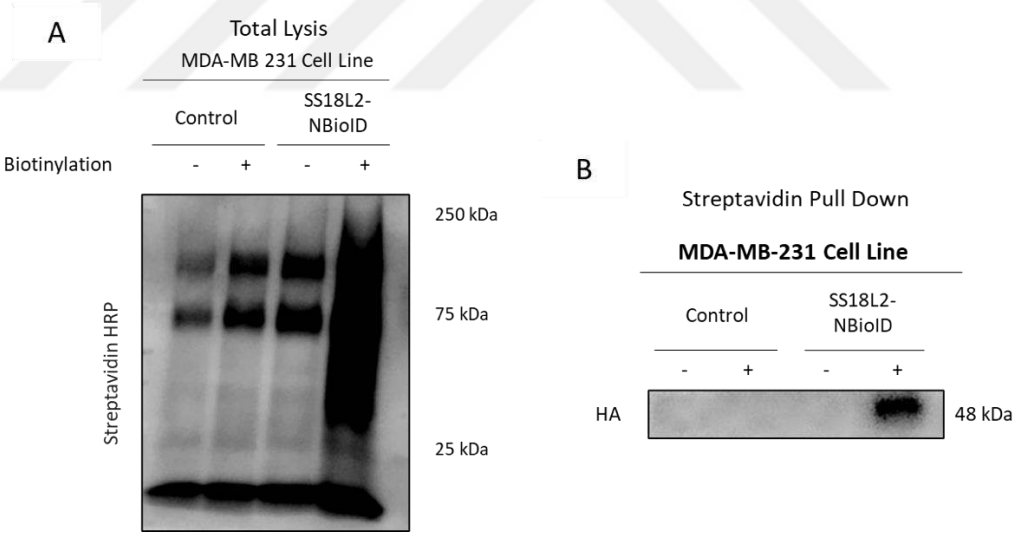

Figure 3.13. Western blot for pull-down trials for SS18L2-NBioID overexpressed cells. A. Pull-down trial with SS18L2- NBioID overexpressed cells. Total lysis. Un-transfected and non- biotinylated MDA-MB-231 cells and non-biotinylated SS18L2-NBioID overexpressed cells are used as control, 250 µg protein, pulled down with 10 µl of Streptavidin Beads. Blotted with streptavidin HRP antibody. "+", "-" indicates the presence of biotinylation. **B.** Western blot from the same samples at A. SS18L2-BirA* fusion protein at 48 kDa was blotted with HA antibody. Un-transfected and non- biotinylated MDA-MB-231 cells and non-biotinylated SS18L2-NBioID overexpressed cells are used as control.

Following the optimization of the experiments, all of the pull-downs of the fractionations of SS18L2-NBioID an BirA*-NBioID OE cells were confirmed with western blots (**Figure 3.14A, B**). It was demonstrated that BirA*-NBioID fractions gave a high intensity smeared band, similar to the biotinylation of the same OE cells. In addition, a smeared signal was observed in both fractions of SS18L2-NBioID in the presence of biotin. Furthermore, it was demonstrated that SS18L2 itself can be pulled down in both cytosolic and nuclear fractions (**Figure 3.14C**).

In the confirmations, 200 μg of protein was pulled down, which may not be efficient to pull down every protein in the SS18L2 interactome. The experiments were only conducted to validate the pull-down experiments. Higher protein levels would be needed to confirm a specific protein from the pull down.

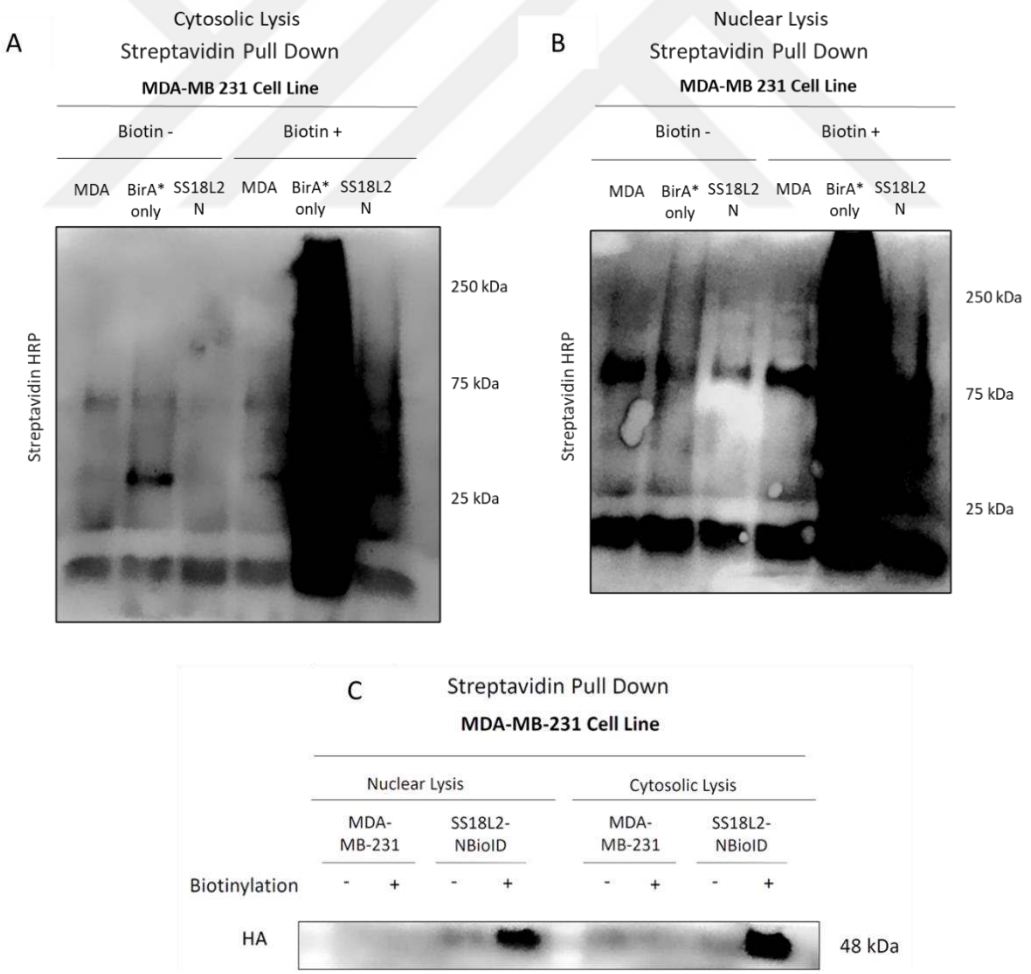

Figure 3.14. Western blot for pull down validations for SS18L2-NBioID overexpressed cells. A. Pull-down validations before mass spectrometry. With SS18L2-

NBioID and BirA*-NBioID overexpressed cells. Cytosolic fractionation. Un-transfected and non- biotinylated MDA-MB-231 cells and non-biotinylated SS18L2-NBioID overexpressed cells are used as control. 200µg protein was pulled down with 8 µl of Streptavidin Beads for all samples. Blotted with streptavidin HRP antibody. **B.** Nuclear fractionation. **C.** SS18L2-BirA* fusion protein at 48 kDa was blotted with HA antibody from pull down samples. Nuclear and cytosolic fractionations. Un-transfected and non-biotinylated MDA-MB-231 cells and non-biotinylated SS18L2-NBioID overexpressed cells are used as control.

Additionally, the same confirmation experiments were applied to SS18-NBioID (**Figure 3.15A, B**) and SS18L1-NBioID (**Figure 3.16A, B**) OE cells for both cytosolic and nuclear fractions. A smeared signal was observed in both fractions of SS18-NBioID and SS18L1-NBioID in the presence of biotin. All results demonstrated that the pull-down procedures were successful and can be utilized in a larger scale to prepare samples before mass-spectrometry.

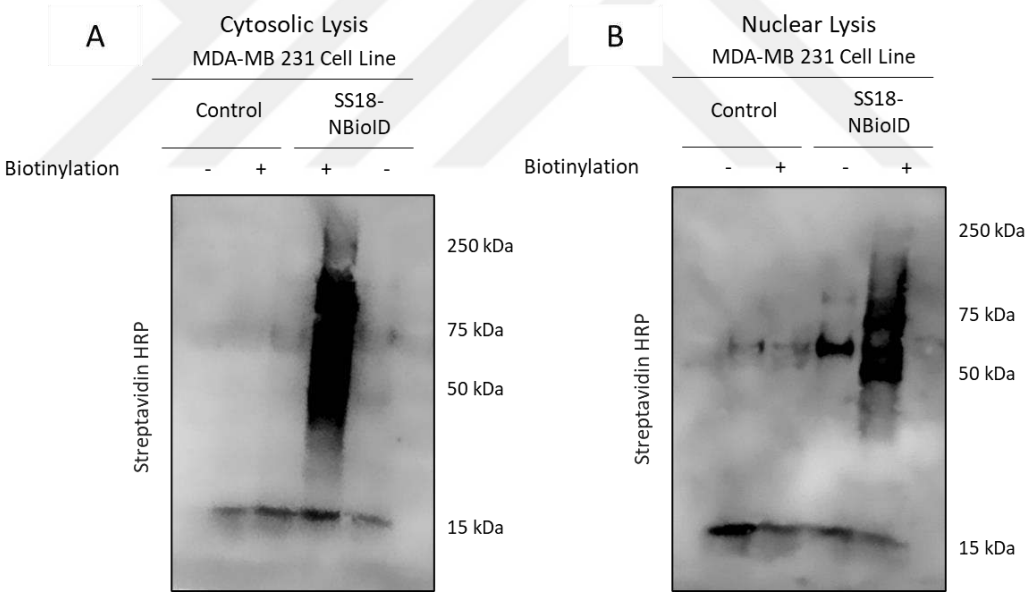

Figure 3.15. Western blot for pull down validations for SS18-NBioID overexpressed cells. A. Pull down validations before mass spectrometry. With SS18-NBioID overexpressed cells. Cytosolic fractionation. Un-transfected and non-biotinylated MDA-MB-231 cells and non-biotinylated SS18L2-NBioID overexpressed cells are used as control. 200µg protein was pulled down with 7.5 µl of Streptavidin Beads for all samples. Blotted with streptavidin HRP antibody. **B.** Nuclear fractionation. 100µg protein was pulled down with 5 µl of Streptavidin Beads.

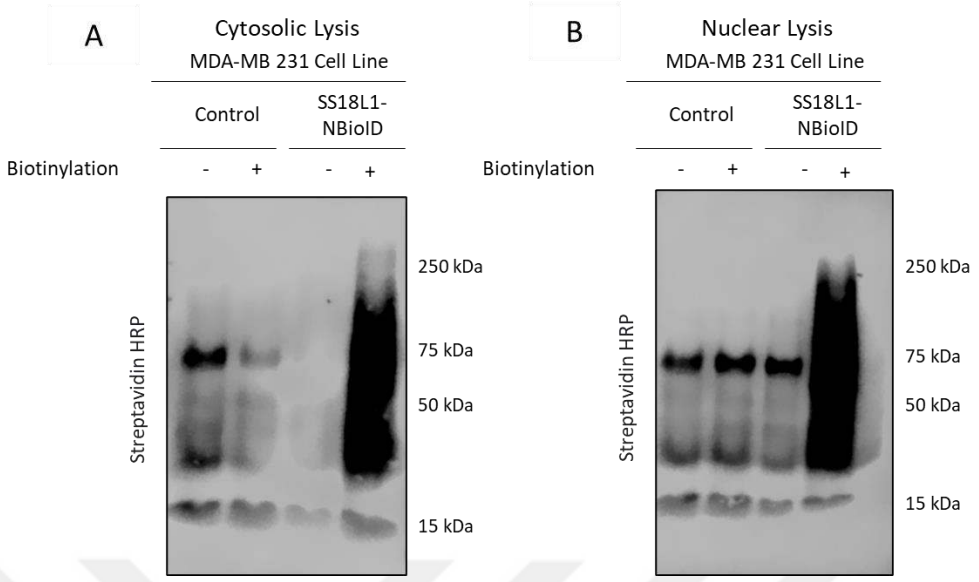

Figure 3.16. Western blot for pull down validations for SS18L1-NBioID overexpressed cells. A. Pull down validations before mass spectrometry. With SS18L1-NBioID overexpressed cells. Cytosolic fractionation. Un-transfected and non-biotinylated MDA-MB-231 cells and non-biotinylated SS18L2-NBioID overexpressed cells are used as control. 200µg protein was pulled down with 7.5 µl of Streptavidin Beads for all samples. Blotted with streptavidin HRP antibody. **B.** Nuclear fractionation. 200µg protein was pulled down with 7.5 µl of Streptavidin Beads for all samples.

3.3 Identification of SS18L2 Interactome with Mass Spectrometry

3.3.1 Analysis of SS18L2 Mass Spectrometry Results

After the confirmations of biotinylation and pull-down efficiency, a pilot experiment to be sent to mass spectrometry was conducted. As previously demonstrated, the localization of SS8L2 was identified to be in the nucleus (**Figure 3.4**). To be certain of which fractionation will be utilized for the identification of SS18L2 interactome, both cytosolic and nuclear fractionated and biotinylated proteins were analyzed in the mass spectrometry in the pilot experiment. Moreover, to determine which control will be applied to the analysis, both proteins from biotinylated BirA*-NBioID OE cells and biotinylated un-transfected MDA-MB-231 cells were utilized. The analyzed samples can be listed as: SS18L2-NBioID cytosolic, BirA*-NBioID cytosolic, MDA-MB-231 cytosolic, SS18L2-NBioID nuclear, BirA*-NBioID nuclear and MDA-MB-231 nuclear.

All samples were biotinylated proteins, non- biotinylated proteins are only used as controls in the confirmation experiments.

The pilot experiment demonstrated that the analysis of SS8L2-NBioID cytosolic fraction with BirA*-NBioID cytosolic fraction as control has a low LogFC (0.21 LogFC) value for SS18L2, while it was expected for SS18L2 to have higher log fold change values, because SS18L2 fusion would be biotinylated itself (**Figure 3.17A).** MA plot demonstrates SS18L2 has a lower level of expression intensity and log fold change in the cytosolic fractionation. On the contrary, the analysis of SS8L2-NBioID nuclear fraction with BirA*-NBioID nuclear fractionation as control resulted the highest LogFC (9.12 LogFC) value to be SS18L2 with a high expression intensity (**Figure 3.17B**). Moreover, some of the proteins that were known to be in the BAF Complex were demonstrated in both cytosolic and nuclear analysis. The proteins ARID1A, SMARCA4, SAMARCB1, SMARCC1, SMARCD1 and GLTSCR1 all resulted in higher LogFC and expression values in the nuclear analysis than the cytosolic analysis. The logFC values of the respective proteins were nearly close to 0 or negative.

In addition, the analysis of SS8L2-NBioID cytosolic fraction with biotinylated un-transfected MDA-MB-231 cytosolic fractionation as control gave a similar result related to the LogFC of SS18L2 (2.95 LogFC) and the demonstrated BAF complex members (**Figure 3.18A**). LogFC of SS18L2 was slightly higher in the analysis with BirA*-NBioID as control. However, the MA plot also indicates SS18L2 is not expressed highly in the cytosolic fractions. In the MA plot with MDA-MB-231 nuclear fractionation as control (**Figure 3.18B**), SS18L2 and GLTSCR1 had the highest LogFC values (11.39 and 11.57 LogFC). In addition, all the illustrated BAF complex members have positive LogFC values. Since the proteins are known to be interacting in the nucleus, the results attested to the success of the pilot experiments.

The MA plots for different control groups led to similar results; however, the MA plots with biotinylated un-transfected MDA-MB-231 as control displayed a more uniform distribution of the protein expressions than the MA plots, where BirA*-NBioID is the control group. Moreover, due to the high expression of BirA* and its elevated biotinylation rate, the analysis in **Figure 3.17** resulted in proteins with a cellular function. Selected proteins can be demonstrated best at the nuclear plot in **Figure 3.18B.** Therefore, nuclear fractionation and un-transfected MDA-MB-231 as control were utilized in the further experiments.

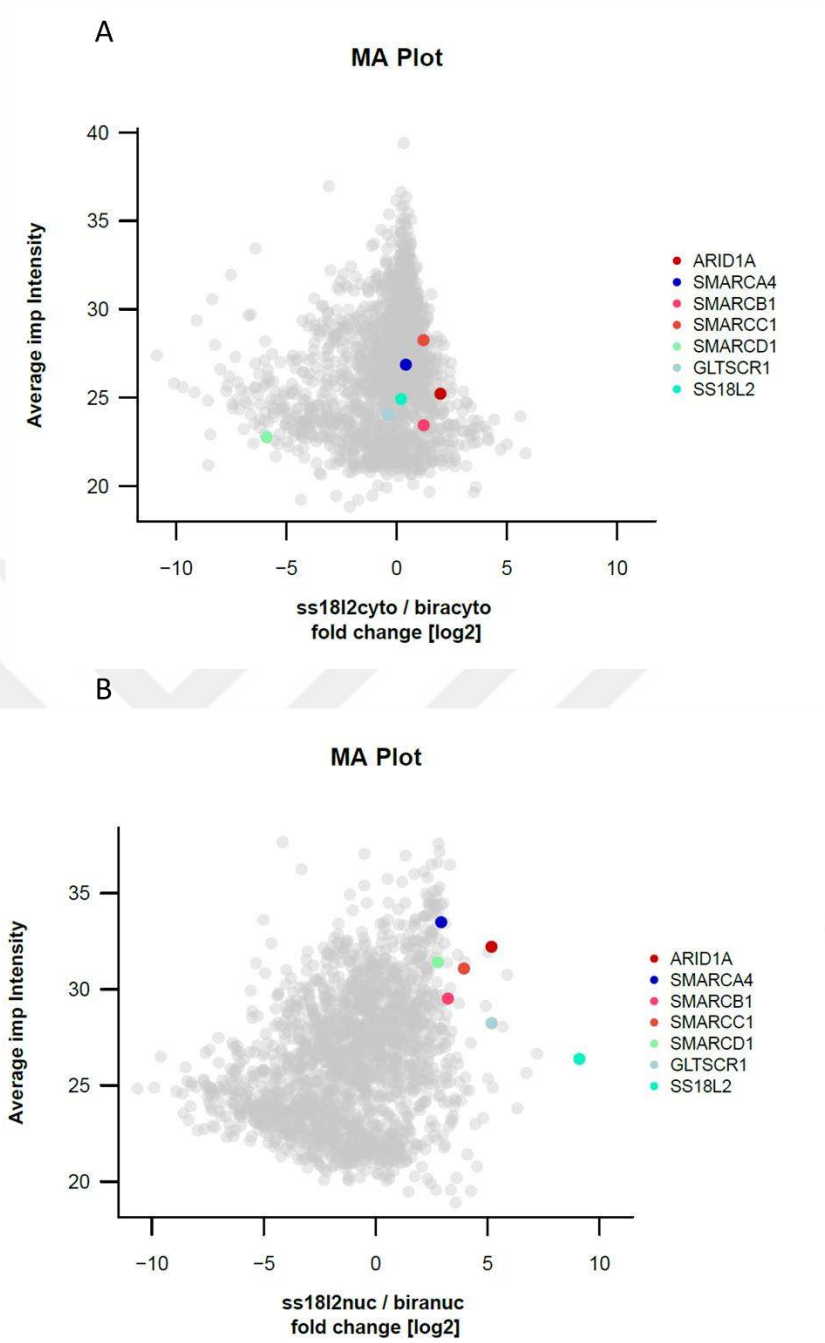

Figure 3.17. SS18L2 pilot mass spectrometry analysis with BirA*-NBioID control. A. MA plot of the interactome of SS18L2-NBioID with BirA*-NBioID as control. Cytosolic fractionation. Plot indicates the average expression intensities, and Log2 fold change values of SS18L2. Analysis is conducted with MaxQuant and Cassiopeia. Indicated proteins are handpicked to be shown in the graph. **B.** MA plot of the interactome of SS18L2-NBioID with BirA*-NBioID as control. Nuclear fractionation. Analysis is conducted with MaxQuant and Cassiopeia. Selected proteins are indicated in the graph with color codes.

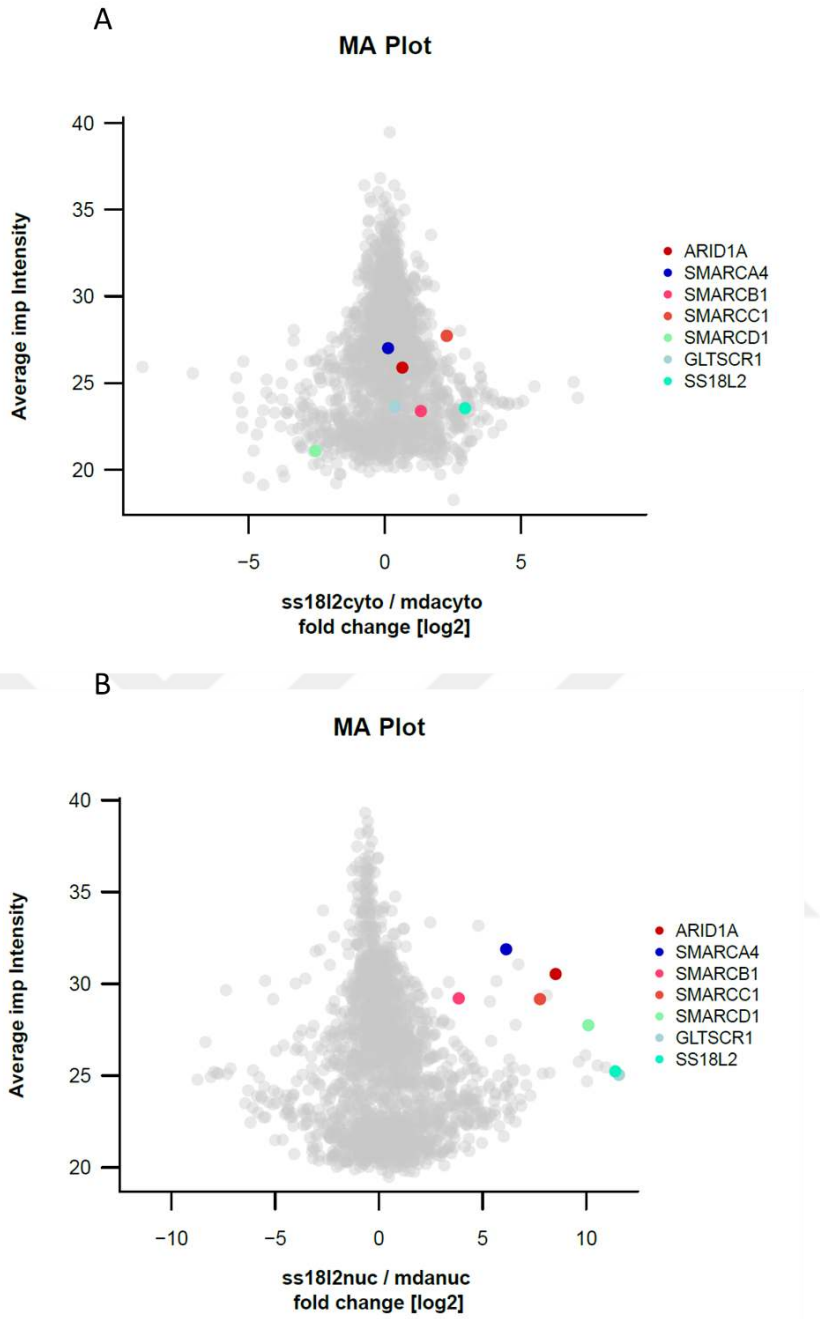

Figure 3.18. SS18L2 pilot mass spectrometry analysis biotinylated MDA-MB-231 control. A. MA plot of the interactome of SS18L2-NBioID with biotinylated MDA-MB-231 cells as control. Nuclear fractionation. Plot indicates the average expression intensities, and Log2 fold change values of SS18L2. Analysis is conducted with MaxQuant and Cassiopeia. Indicated proteins are handpicked to be shown in the graph. **B.** MA plot of the interactome of SS18L2-NBioID with biotinylated MDA-MB-231 cells as control. Nuclear fractionation. Analysis is conducted with MaxQuant and Cassiopeia. Selected proteins are indicated in the graph with color codes.

Following the pilot mass spectrometry analysis, all OE cells with the engineered fusion proteins SS18L2-BirA*, SS18-BirA* and SS18L1-BirA* were sent to mass spectrometry for final n=2 biological replicates. 2 biological replicates were additionally prepared and included SS8L2 and its homologs. Un-transfected and biotinylated MDA-MB-231 cells were utilized as the control group through all the experiments. The analyzed samples can be listed as: SS18L2-NBioID nuclear, SS18-NBioID nuclear, SS18L1-NBioID nuclear and MDA-MB-231 nuclear.

The interactome of SS18L2 is illustrated with the volcano plot and MA plot in **Figure 3.19.** 169 proteins were identified with LogFC higher than 2. The top 10 significantly enriched proteins are shown with color codes in the volcano plot in **Figure 3.19A.** Most significantly enriched proteins and the log fold changes related to the protein expression levels are demonstrated in the MA plot in **Figure 3.19B.** The findings suggested the top 10 significantly expressed proteins as follows: SMARCC1, SMARCD1, SMARCA4, ARID1B, HAUS6, GIGYF2, GLTSCR1, DPF2 and ARID1A. In the analysis, SS18L2 is also significantly enriched with a 7.2 LogFC. The top 15 most significantly expressed proteins and their LogFC values are shown in **Table 3.2.** Among the highly expressed proteins 8 out of 10 proteins were observed to be BAF complex members. Both canonical and non-canonical BAF complex members were observed to have a high expression rate. While SMARCB1 is not found in the non-canonical subtype, DPF2 is specific to canonical BAF and GLTSCR1 is specific to non-canonical BAF. Other canonical and non-canonical BAF complex members besides ACTB and SS18 were found to have log fold changes greater than 3, showing a significance in the relation of SS18L2 to BAF complex. In addition, pBAF specific member PBRM1 and BRD7 were expressed with a LogFC greater than 3.

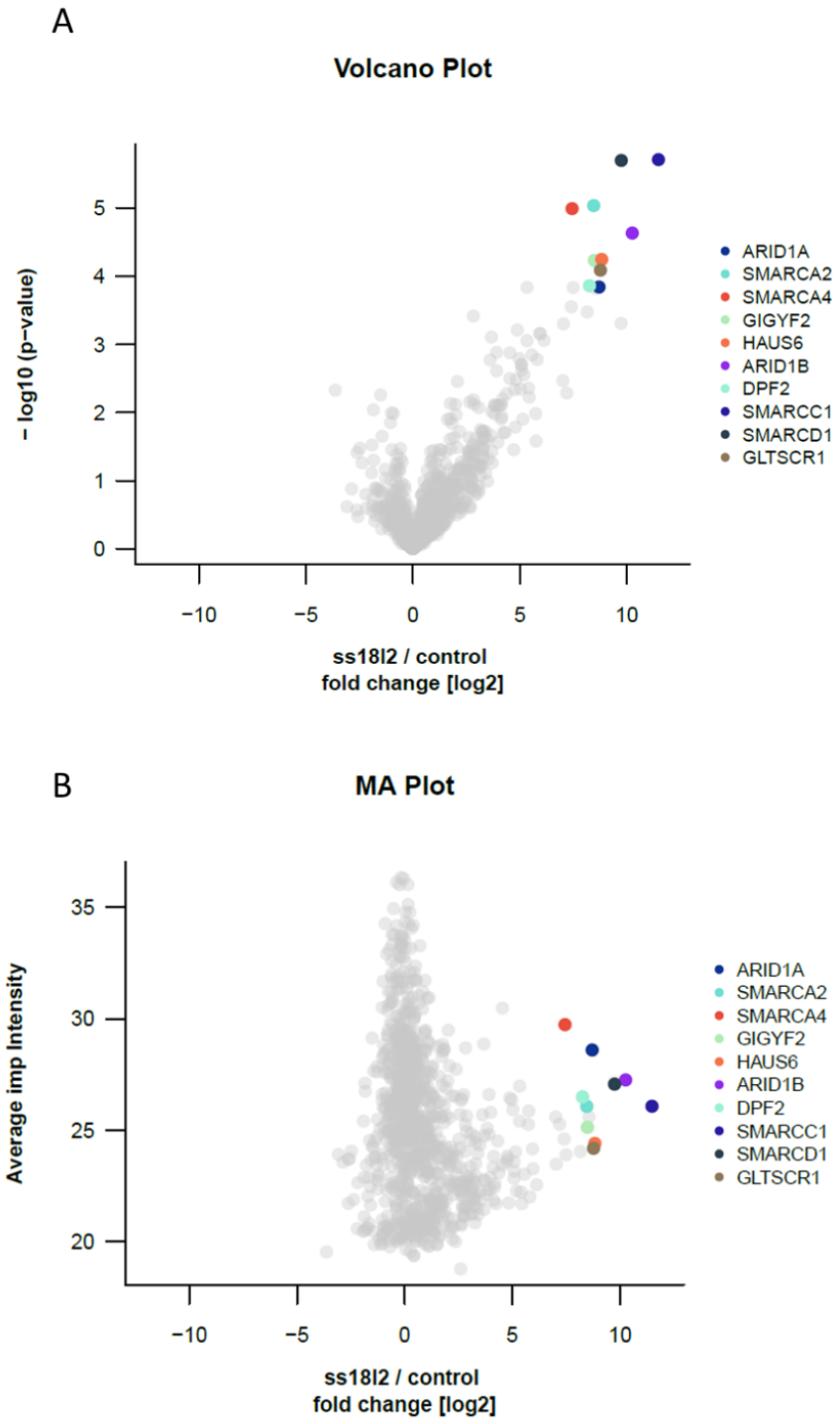

Figure 3.19. SS18l2 mass spectrometry analysis. A. Volcano plot of the interactome of SS18L2-NBioID with biotinylated MDA-MB-231 cells as control. Nuclear fractionation. Analysis is conducted with MaxQuant and Cassiopeia. Top 10 most significant proteins are indicated with color codes. p-value<0.001. **B.** MA plot of the interactome of SS18L2-NBioID with biotinylated MDA-MB-231 cells as control. Plot indicates the average expression intensities, and Log2 fold change values of SS18L2 interactome. Analysis is conducted with MaxQuant and Cassiopeia. Top 10 most

significant proteins are indicated with color codes. p-value<0.001. n=2 biological replicates.

Table 3.2. Top 15 most significant proteins of SS18L2 interactome analysis. With MaxQuant and Cassiopeia. Proteins are sorted as their significance related to their P. Value. BAF complex members are indicated with orange.

PROTEIN	LOGFC	P.VALUE
SMARCC1	11.48	1.95E-06
SMARCD1	9.75	2.00E-06
SMARCA2	8.45	9.22E-06
SMARCA4	7.45	1.02E-05
ARID1B	10.26	2.33E-05
HAUS6	8.83	5.68E-05
GIGYF2	8.49	5.87E-05
GLTSCR1	8.77	8.17E-05
DPF2	8.29	0.000138108
ARID1A	8.71	0.000144455
SMARCB1	5.34	0.000145858
HAUS8	7.5	0.00014684
SMARCC2	8.56	0.000152386
ZBTB33	7.41	0.000281588
BCL7C	8.15	0.000335279

Gene ontology analysis of the interactome of SS18L2 resulted in biological analysis such as chromatin organization, remodeling, or ATP-dependent chromatin remodeling with significantly small FDR values (**Figure 3.20A**). Moreover, processes such as cell cycle regulation, mitotic cell cycle related pathways and regulation of G2/M phase transition were among the significant processes. To further confirm the mass spectrometry and the related hits, one of the hits GLTSCR1 was validated to be pulled down in the biotinylated SS18L2-NBioID OE cells (**Figure 3.20B**). The result is a confirmation for the pull down of mass spectrometry; however, does not give an exact notion related to the interaction of SS18L2 to GLTSCR1.

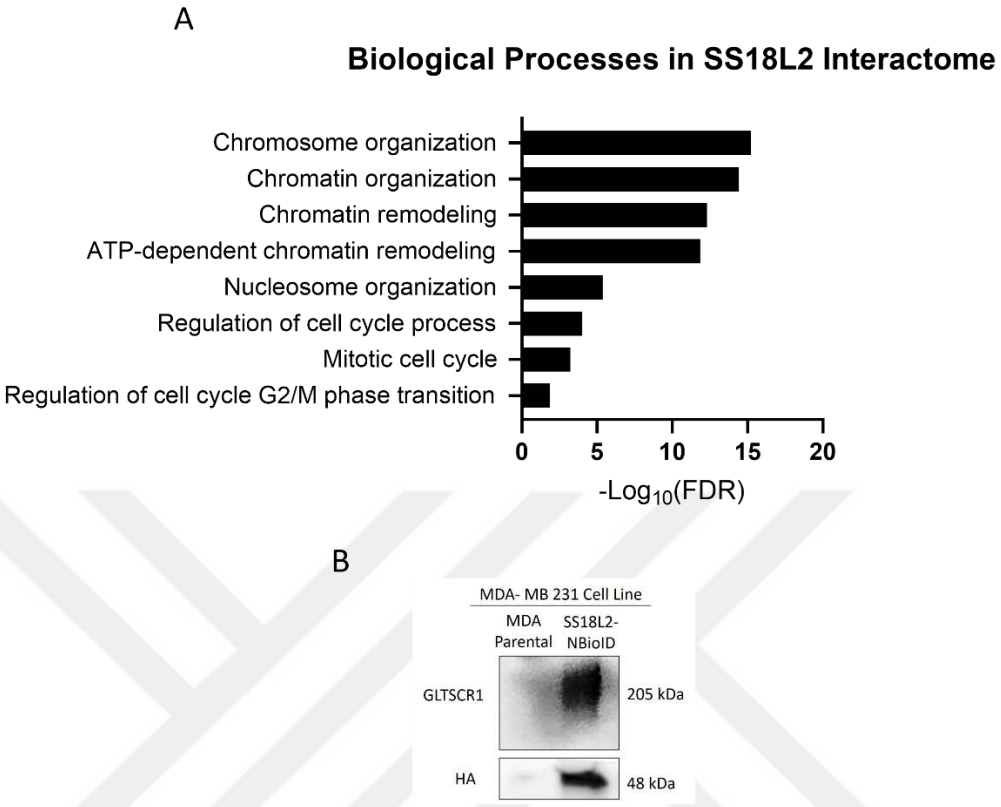

Figure 3.20. Further investigation of the SS18L2 interactome. A. Biological processes analysis of the interactome of SS18L2. Data from Gene Ontology Analysis, STRING platform, LogFC>1.5 as cut off. **B.** Confirmation of mass spectrometry. Western blot with anti**-** GLTSCR1 antibody is blotted after pull-down of the biotinylated SS18L2-NBioID OE nuclear proteins. Un-transfected biotinylated MDA-MB-231 cells were used as control.

3.3.2 Analysis of SS18 and SS18L1 Mass Spectrometry Results

The interactome of SS18 is demonstrated with the volcano plot and MA plot in (**Figure 3.21**). 374 proteins were identified with LogFC higher than 2. Top 10 most significantly enriched proteins are illustrated in (**Figure 3.21A**) with color codes. Moreover, significantly enriched proteins related to their LogFC and average expression are shown in **Figure 3.21B.** The findings suggested the top 10 significantly expressed proteins as follows: CPSF6, SMARCC1, SMARCD1, SMARCA2, YLPM1, SMARCA4, ARID1B, TCF20, KHSRP and RBM17. SS18 itself was also significantly enriched with 6.09 LogFC. In addition, the top 15 most significantly expressed proteins and their LogFC values are shown in **Table 3.3.** Five out of ten proteins between the top 10 most significant proteins were found to be BAF complex members. The interactome of SS18 was found to be comprised of BAF complex members, as expected, and known from the literature. In addition, PBRM1 and BRD7, members specific to pBAF complex were also expressed with a fold change greater than 3.

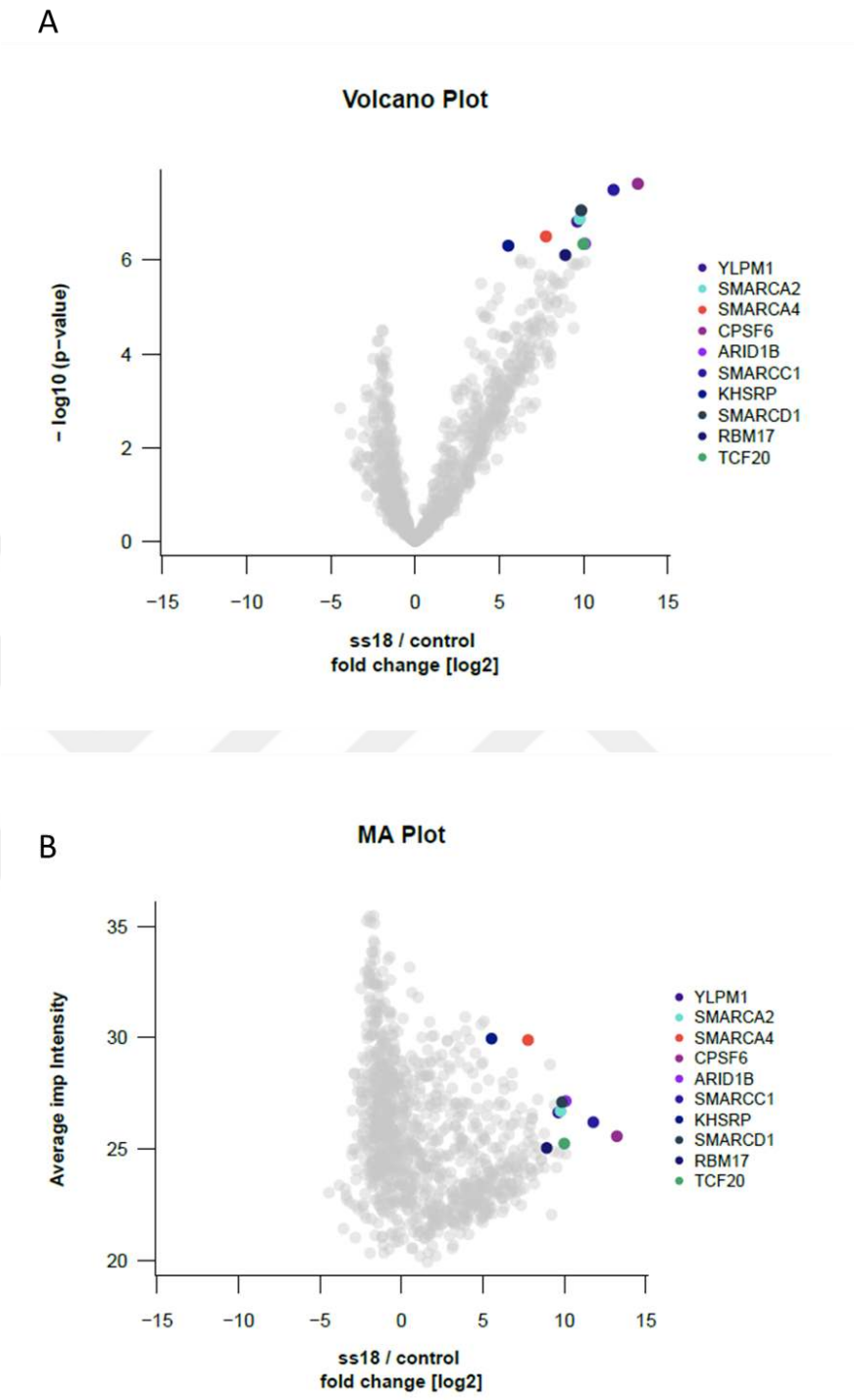

Figure 3.21. SS18 mass spectrometry analysis. A. Volcano plot of the interactome of SS18-NBioID with biotinylated MDA-MB-231 cells as control. Nuclear fractionation. Analysis is conducted with MaxQuant and Cassiopeia. Top 10 most significant proteins are indicated with color codes. p-value<0.001. **B.** MA plot of the interactome of SS18-NBioID with biotinylated MDA-MB-231 cells as control. Plot indicates the average expression intensities, and Log2 fold change values of SS18 interactome. Analysis is conducted with MaxQuant and Cassiopeia. Top 10 most significant proteins are indicated with color codes. p-value<0.001. n=2 biological replicates.

Table 3.3. Top 15 most significant proteins of SS18 interactome analysis. With MaxQuant and Cassiopeia. Proteins are sorted as their significance related to their P. Value. BAF complex members are indicated with orange.

PROTEINS	LOGFC	P. VALUE
CPSF6	13.21	2.35E-08
SMARCC1	11.77	3.17E-08
SMARCD1	9.85	8.69E-08
SMARCA2	9.77	1.31E-07
YLPM1	9.62	1.50E-07
SMARCA4	7.76	3.13E-07
ARID1B	10.08	4.48E-07
TCF20	9.98	4.53E-07
KHSRP	5.53	4.93E-07
RBM17	8.91	7.80E-07
ZFR	6.25	9.83E-07
NUDT21	10.07	1.09E-06
FUBP1	6.30	1.14E-06
SF3B2	6.79	1.16E-06
CSTF2	9.35	1.16E-06

In addition to SS18, the interactome of SS18L1 was illustrated with the plots in **Figure 3.22**. 267 proteins were identified with LogFC higher than 2. The plots demonstrate top 10 most significantly enriched proteins as follows: SMARCA2, SMARCD1, SMARCC1, SMARCA4, GLTSCR1, CPSF6, SS18L1, ARID1B, YLPM1 and RBM17. Moreover, the top 15 proteins according to their significance is indicated in **Table 3.4** below. Seven out of ten most significant proteins in the interactome were BAF complex members. Furthermore, PBRM1 and BRD7, members specific to pBAF complex were also expressed with a LogFC greater than 3.5. The results were similar to SS18 in terms of their relation to BAF complex as previously known in the literature. Moreover, other significantly expressed proteins in both interactomes were highly similar, indicating the similarity between SS18 and SS18L1, again as expected.

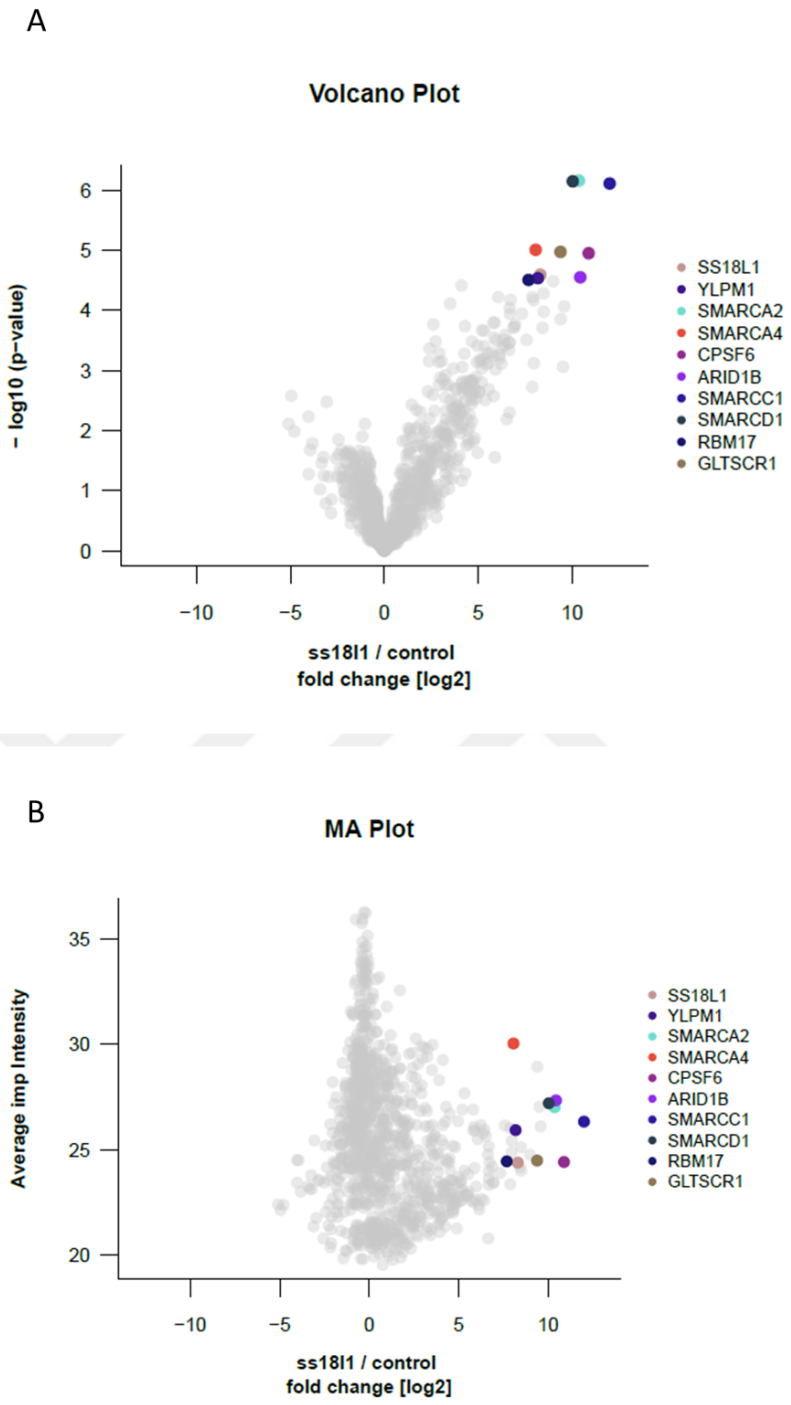

Figure 3.22. SS18L1 mass spectrometry analysis. A. Volcano plot of the interactome of SS18L1-NBioID with biotinylated MDA-MB-231 cells as control. Nuclear fractionation. Analysis is conducted with MaxQuant and Cassiopeia. Top 10 most significant proteins are indicated with color codes. p-value<0.001. **B.** MA plot of the interactome of SS18L1-NBioID with biotinylated MDA-MB-231 cells as control. Plot indicates the average expression intensities, and Log2 fold change values of SS18L1 interactome. Analysis is conducted with MaxQuant and Cassiopeia. Top 10 most significant proteins are indicated with color codes. p-value<0.001. n=2 biological replicates.

Table 3.4. Top 15 most significant proteins of SS18L1 interactome analysis. With MaxQuant and Cassiopeia. Proteins are sorted as their significance related to their P. Value. BAF complex members are indicated with orange.

PROTEINS	LOGFC	P. VALUE
SMARCA2	10.36	6.85E-07
SMARCD1	10.04	7.05E-07
SMARCC1	12	7.66E-07
SMARCA4	8.06	9.75E-06
GLTSCR1	9.38	1.06E-05
CPSF6	10.88	1.11E-05
SS18L1	8.30	2.53E-05
ARID1B	10.44	2.79E-05
YLPM1	8.18	2.89E-05
RBM17	7.68	3.09E-05
TCF20	8.99	3.25E-05
KHSRP	4.11	3.83E-05
GIGYF2	8.48	5.17E-05
PCF11	6.08	5.98E-05
FIP1L1	7.88	6.00E-05

3.3.3 Comparison between SS18L2 and its homologs

Following the identification of the possible interactomes of SS18L2, SS18 and SS18L1, the difference or the similarity of the homologs were examined. First, proteins that are unique to SS18L2 were depicted. The differences were detected by distinguishing the differential proteins specific to SS18, SS18L1 and SS18L2, and the ones that are not expressed in the other groups as shown in the venn diagram (**Figure 3.23**). There were 17 proteins that were found to be specific to SS18L2. The differential proteins are indicated in the table below (**Table 3.5**).

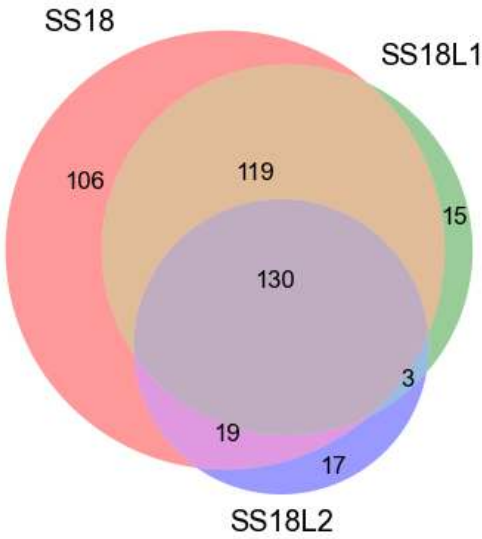

Figure 3.23. Venn diagram for the distribution of the proteins. LogFC>2 as cut off.

Specific proteins of SS18L2 different from its homologs did not have a high log fold change and none of the proteins were significant in terms of p value according to the MaxQuant analysis. NELFA was identified as an essential factor related to transcriptional elongation, but due to the high p value, it was not chosen as a hit. Furthermore, some of the SS18L2-unique interactors were related to mitochondria. Mitochondrial proteins were not abundant in the 130 proteins that are common for all SS18L2, SS18 and SS18L1. However, a correlation was not made due to low significance.

Table 3.5. The 17 differential proteins for SS18L2. LogFC values are indicated in the graph. LogFC>2 as cut off.

	Proteins	Protein Names	LogFC
1	SS18L2	SS18 Like Protein 2	7.2
2	AIFM1	Apoptosis Inducing Factor 1, Mitochondrial	3.71
3	MPG	DNA-3-Methyladenine Glycosylase	3.3
4	DDX28	Probable ATP-Dependent RNA Helicase DDX28	3.16
5	MAFG	Transcription Factor MafG	3.1
6	MAFK	Transcription Factor MafK	3.1
7	HM13	Minor Histocompatibility Antigen H13	3.04
8	MAP4	Microtubule Associated Protein 4	2.83
9	ZW10	Centromere/Kinetochore Protein ZW10 Homolog	2.42
10	NDUFA4	Cytochrome C Oxidase Subunit NDUFA4	2.39

11	TIMM23	Mitochondrial Import Inner Membrane Translocase Subunit Tim23	2.26
12	NDUFB5	NADH: Ubiquinone Oxidoreductase Subunit B5	2.23
13	SACM1L	Phosphatidylinositol-3-phosphatase SAC1	2.16
14	MRPL15	39S Ribosomal Protein L15, Mitochondrial	2.12
15	NELFA	Negative Elongation Factor A	2.09
16	ACACA	Acetyl-CoA Carboxylase Alpha	2.08
17	APMAP	Adipocyte Plasma Membrane Associated Protein	2.06

From the comparison between SS18L2 and the homologs, no specific conclusion could be made due to significance cut-offs. Therefore, we decided to conduct another analysis taking the expression rates into account as well.

The relationship of SS18L2 with its homologs was further demonstrated in the volcano plots below (**Figure 3.24**). The plots have biotinylated MDA-MB-231 cells as the control group. In **Figure 3.24A,** the volcano plot from the suggested interactome of SS18L2, SS18 and SS18L1 were indicated to have a LogFC near zero with a small significance. SS18L2 can be seen to have a high expression rate as expected in its own interactome. Moreover, in **Figure 3.24B** the volcano plot from the suggested interactome of SS18, SS18 was indicated to have a LogFC positively near zero and SS18L2 was indicated to have a LogFC negatively near zero. SS18 itself can be shown to have a high expression level. Finally, in **Figure 3.24C**, the plot suggests SS18 and SS18L2 had a negatively enriched LogFC with small significance in the interactome of SS18L1. Together, these results show that the homologs do not have high expression rates within each other's interactomes.

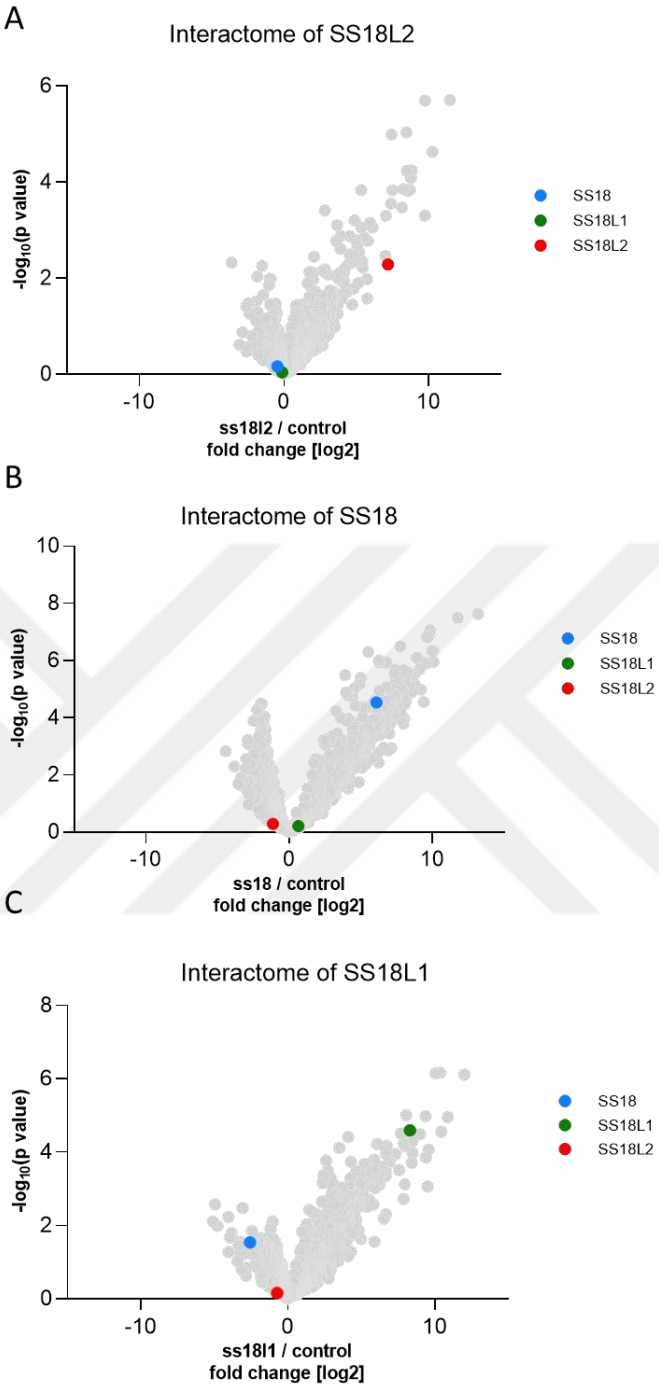

Figure 3.24. Volcano plots showing the differences between SS18L2 and its homologs in the initial suggested interactomes. A. Volcano plot of the interactome of SS18L2, demonstrating SS18, SS18L1 and SS18L2 regarding the Log2 fold change and -Log10 (p values). **B.** Volcano plot of the interactome of SS18, demonstrating SS18, SS18L1 and SS18L2 regarding the Log2 fold change and -Log10 (p values). **C.** Volcano plot of the interactome of SS18L1, demonstrating SS18, SS18L1 and SS18L2 regarding the Log2 fold change and -Log10 (p values). Un-transfected and biotinylated MDA-MB-231 cells used as control.

The differences between the expression rates of the interacting proteins were shown with SS18L2 over SS18 as control (**Figure 3.25**) and SS18L2 over SS18L1 as the control group (**Figure 3.26**). Proteins ARID1A, SMARCA4, SMARCB1, SMARCC1, GLTSCR1 and BRD9 were selected to demonstrate the BAF complex members in all of the groups, including SS18, SS18L1; and SS18L2 were visualized in the volcano plot and the MA plots in both figures.

When SS18 interactome was assigned as the control group, the BAF complex members had a negative or close to zero LogFC values with no significance (**Figure 3.25A**). The top 15 proteins with highest LogFC values are shown in (**Table 3.6**). The proteins in the table were sorted related to the LogFC values because negatively enriched proteins in the analysis had the highest significance. In the case of determining interacting partners, positively enriched proteins were taken into consideration. It was also essential to determine the LogFC values and p values of the hit proteins in the comparison analysis with the original interactome analysis of SS18L2. HAUS6 and HAUS8 proteins were found to be significantly enriched in the interactome analysis of SS18L2, which was previously shown (8.83 and 7.5 LogFC values respectively) (**Table 3.2**). Moreover, other top 15 proteins in the analysis with SS18 control had LogFC values lower than 3 and none of them were significant in SS18L2 interactome, besides MPG (**Table 3.6**). SS18 was not present with SS18L2 overexpression, while the major different protein was SS18L2 in the analysis. The high expression in the comparison could be expected due to the high biotinylation of SS18L2 when compared to the control group. However, the distinct difference between SS18 and SS18L2 might be related to both homologs not interacting with each other as well. Even SS18L1 had a negative LogFC with a low significance, indicating SS18 and SS18L1 were not possible interacting partners. In the MA plot, there were low number of proteins that were positively enriched compared to negatively enriched proteins. (**Figure 3.25B**).

Along with the analysis with SS18 as control, the comparison made with SS18L1 as the control group resulted in similar outcomes. Selected BAF complex members all had a negative LogFC with a small -Log10 (p value) (**Figure 3.26A**). Moreover, the overall positively enriched proteins based on the intensity and LogFC values are higher in SS18L2 over SS18L1 comparison than the SS18L2 over SS18 (**Figure 3.26B**). The top 15 proteins with highest LogFC values are shown in (**Table 3.7**). SDHB and MPG proteins were also found to be significantly enriched in the interactome analysis of

SS18L2 with p<0.5 for all analysis (**Table 3.7**). SS18 was indicated to be positively expressed in the SS18L2 over SS18L1 comparison; however, in both interactomes with MDA control SS18 has negative LogFC values. SS18L2 interactome had a higher but still negatively enriched LogFC for SS18; therefore, in the graphs SS18 seems to be expressed with a LogFC smaller than 3 (**Figure 3.26**). In both plots, again SS18L2 was the highest enriched protein and SS18L1 was among the least interacting proteins with SS18L2.

Based on the findings, HAUS6 and HAUS8 proteins were both determined to be enriched significantly with p value smaller than 0.01 in the comparison analysis with SS18 control and smaller than 0.001 in the original interactome analysis of SS18L2 with higher LogFC values. Considering the LogFC values with the p values from all the analysis, HAUS6 and HAUS8 were determined as essential proteins to be further investigated. Alternatively, it could be suggested that the homologs may not be interacting with each other while all of them exhibit a potential interaction with the BAF complex.

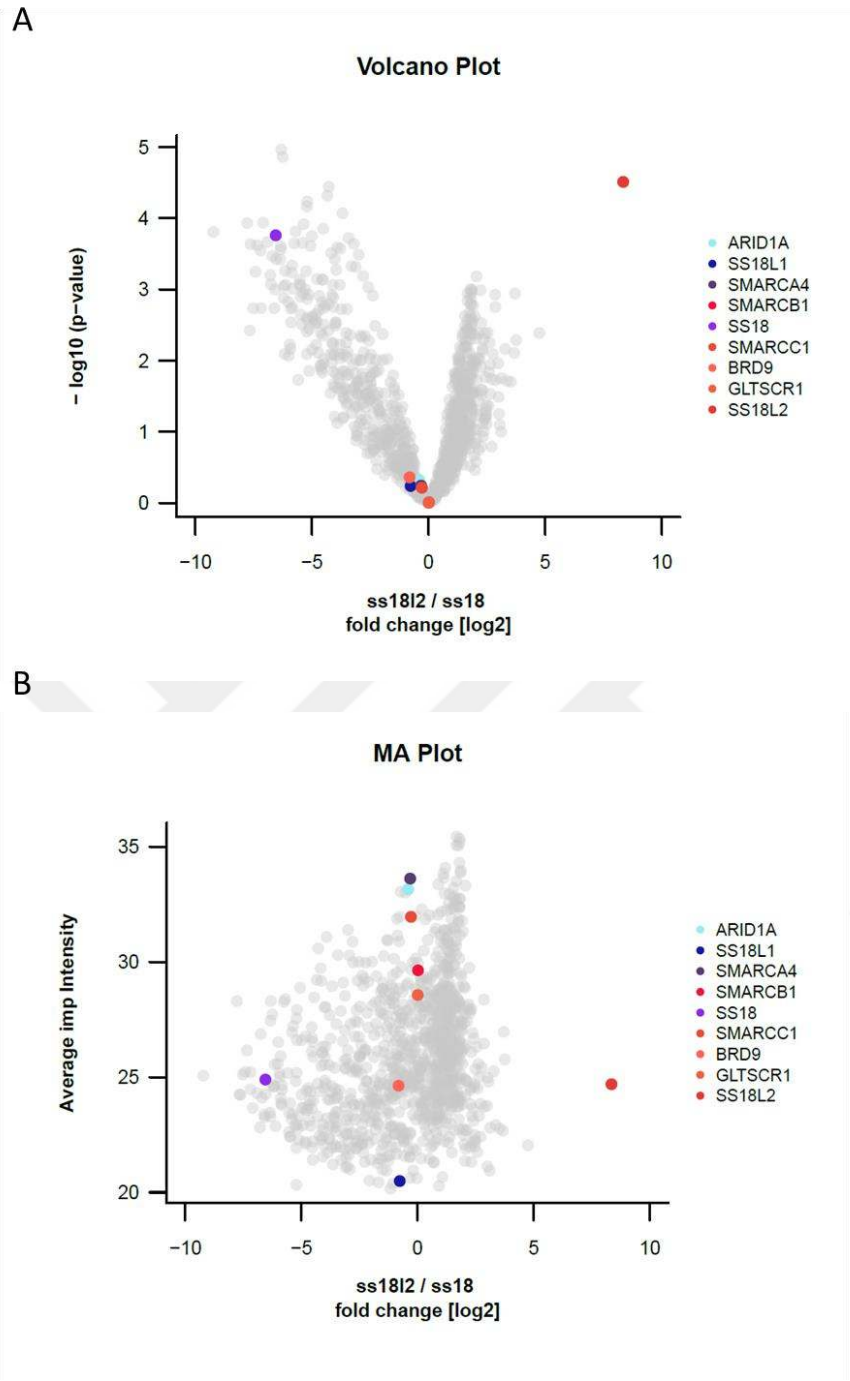

Figure 3.25. Comparison analysis of mass spectrometry results of SS18L2 with SS18 as control group. A. Volcano plot of the comparison of SS18L2-NBioID with SS18-NBioID as control. Nuclear fractionation. Analysis is conducted with MaxQuant and Cassiopeia. **B.** MA plot of the comparison. Plot indicates the average expression intensities, and Log2 fold change values of the expression differences between SS18L2 over SS18. Analysis is conducted with MaxQuant and Cassiopeia. Selected proteins are indicated in the graph with color codes. n=2 biological replicates.

Table 3.6. Top 15 proteins with the highest LogFC values for SS18L2 with SS18 as control group. With MaxQuant and Cassiopeia. Proteins are sorted according to their LogFC values. LogFC values and p values of the proteins in the SS18L2 interactome with un-transfected biotinylated MDA-MB-231 control are indicated. Significant proteins with p value<0.5 in all analysis are indicated with green.

PROTEINS	LOGFC	P.VALUE	LOGFC IN SS18L2	P. VALUE IN SS18L2
SS18L2	8.35	3.10E-05	7.2	0.005183013
HAUS6	3.71	0.001144636	8.83	5.68E-05
IBA57	4.75	0.004099677	1.19	0.291968379
HAUS8	3.75	0.005177547	7.5	0.00014684
EIF2S2	3.28	0.006876318	0.43	0.59860258
PNO1	3.68	0.00783193	-0.76	0.36422946
STK10	2.99	0.010899007	0.2	0.786623869
SH3BP4	3.12	0.017656411	-0.26	0.787569813
NDUFB5	2.87	0.018831212	2.23	0.292152374
MPG	3.36	0.019328206	3.3	0.034829653
ATP5B	3.53	0.019868295	0.87	0.363886844
UTP18	3.01	0.032238462	0.18	0.873906478
PARL	3.02	0.055901598	0.01	0.983948234
APOO	3.05	0.080048194	1.39	0.136873202
PHGDH	3.12	0.100137923	1.17	0.47048197

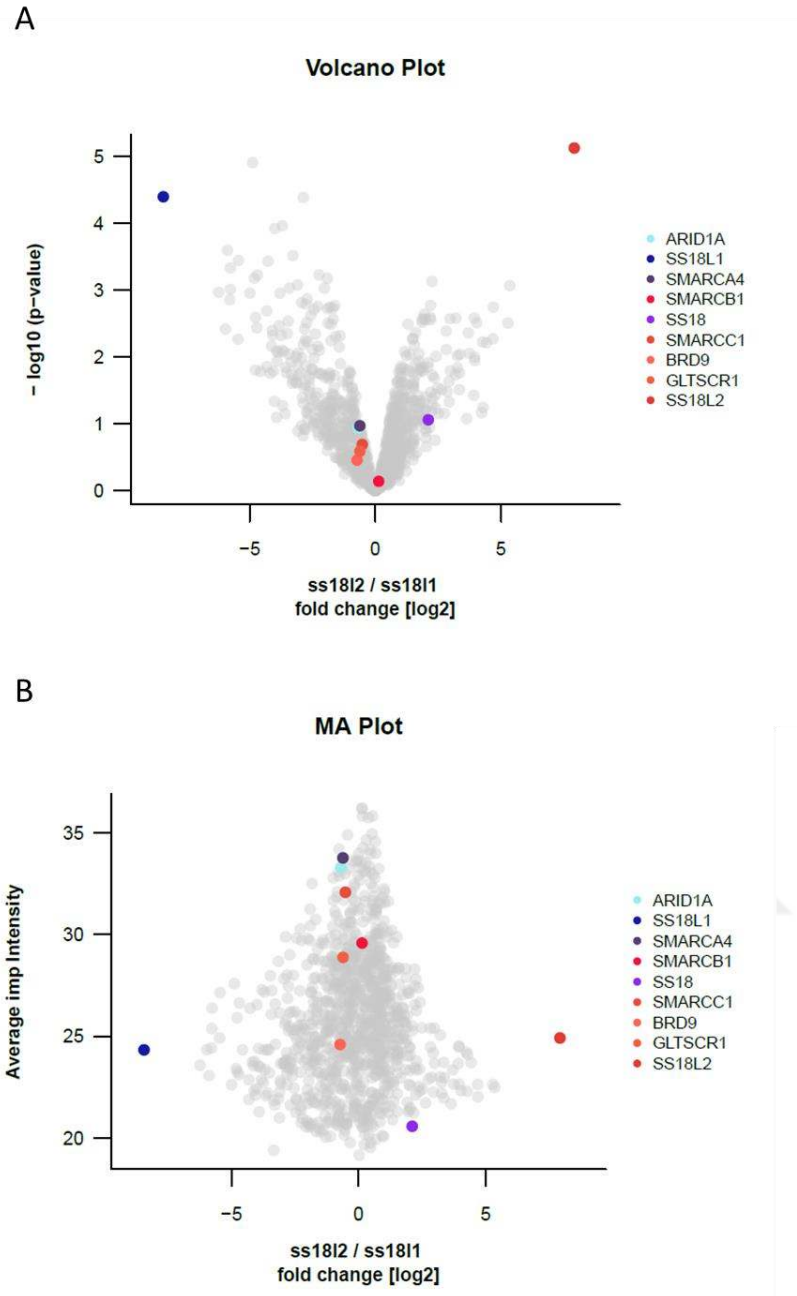

Figure 3.26 Comparison analysis of mass spectrometry results of SS18L2 with SS18L1 as control group. A. Volcano plot of the comparison of SS18L2-NBioID with SS18L1-NBioID as control. Nuclear fractionation. Analysis is conducted with MaxQuant and Cassiopeia. **B.** MA plot of the comparison. Plot indicates the average expression intensities, and Log2 fold change values of the expression differences between SS18L2 over SS18L1. Analysis is conducted with MaxQuant and Cassiopeia. Selected proteins are indicated in the graph with color codes. n=2 biological replicates.

Table 3.7. Top 15 proteins with the highest LogFC values for SS18L2 with SS18L1 as control group. With MaxQuant and Cassiopeia. Proteins are sorted according to their LogFC values. LogFC values and p values of the proteins in the SS18L2 interactome with un-transfected biotinylated MDA-MB-231 control are indicated. Significant proteins with p value<0.5 in all analysis are indicated with green.

PROTEINS	LOGFC	P.VALUE	LOGFC IN SS18L2	P. VALUE IN SS18L2
SS18L2	7.92	7.49E-06	7.2	0.005183013
NDUFB5	5.35	0.00085076	2.23	0.292152374
RAB27A	5.28	0.00309669	0.49	0.597707702
SH3BP4	4.7	0.00178827	-0.26	0.787569813
LGALS1	4.68	0.00532974	1.73	0.103700228
TAP2	4.43	0.00609926	1.09	0.274249277
ABCB10	4.43	0.00609926	1.09	0.274249277
EEF1D	4.29	0.05622459	5.96	0.000706697
MRPL15	4.27	0.00590683	2.13	0.067124345
APOO	4.23	0.06776443	1.39	0.136873202
DDX28	4.04	0.00305371	3.16	0.085539894
PARL	3.96	0.01123729	0.01	0.983948234
MPG	3.95	0.00999217	3.3	0.034829653
STT3A	3.94	0.00261014	-0.1	0.924412554
SDHB	3.92	0.03728313	6.13	0.000873726

Previously, **Figure 3.19** indicated an interaction of SS18L2 to most of the BAF complex members. To demonstrate the similarity and the significance of the BAF complex members that were depicted as “hits”, two plots were constructed. The BAF complex members from the previous findings, as shown in the graphs, are as listed: ARID1A, ARID1B, BCL7C, BRD9, DPF2, GLTSCR1, SMARCB1, SMARCA2, SMARCA4, SMARCC1, SMARCC2, SMARCD1 and SMARCD2 (**Table 3.8**). In **Figure 3.27A**, the proteins in the suggested interactomes of SS18L2 and SS18 are shown. Moreover, in **Figure 3.27B**, the proteins in the suggested interactomes of SS18L2 and SS18L1 are shown. Both graphs demonstrate that the BAF complex members that had the high expression rates in SS18L2 were the highest interacting partners when compared

to SS18 and SS18L1. As SS18 and SS18L1 were previously shown to be interacting with BAF complex, these results suggest an interaction of SS18L2 to BAF (SWI/ SNF) complex.

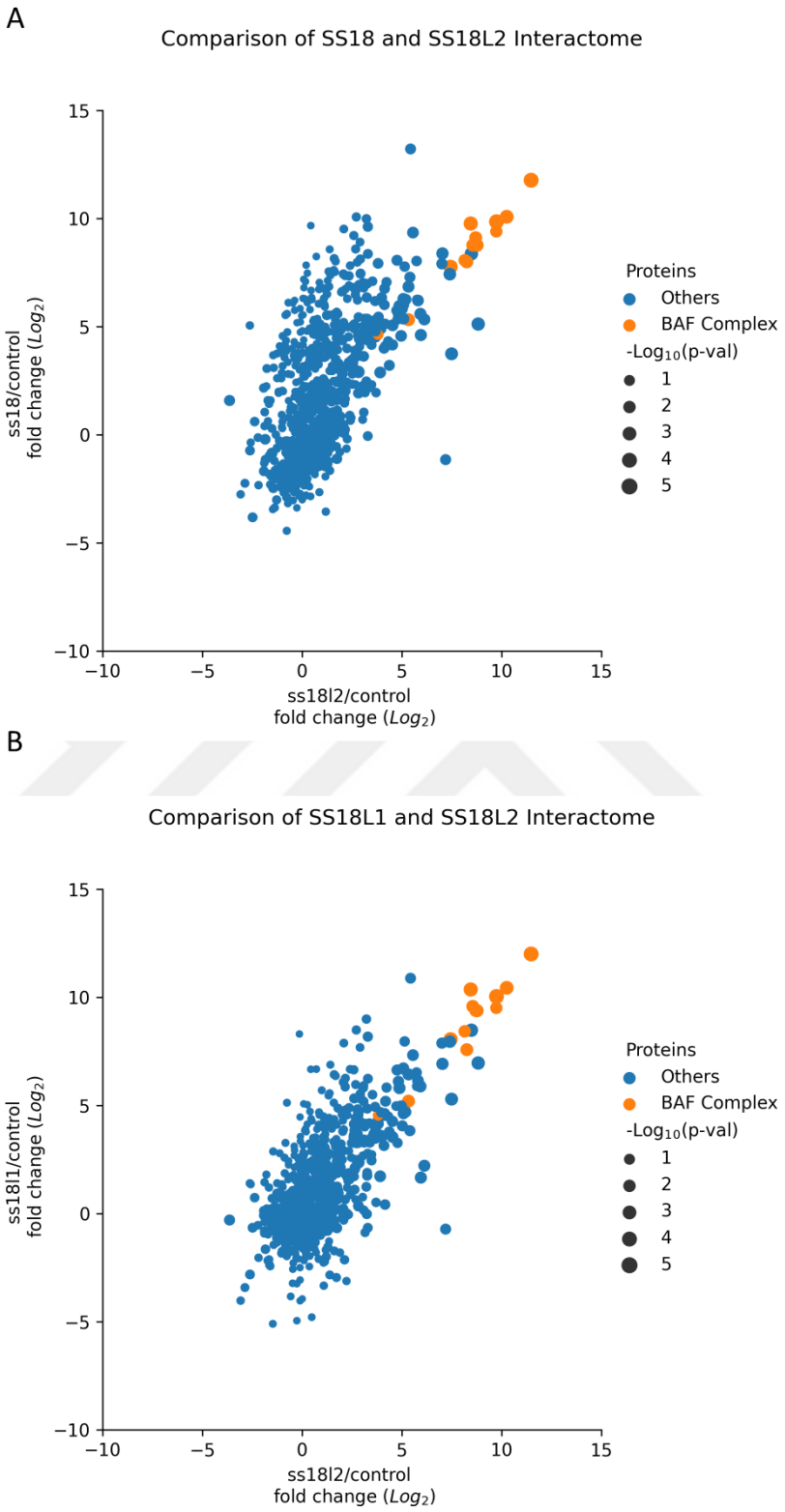

Figure 3.27. Common proteins in the comparison of SS18L2 and its homologs. **A.** Graph showing the distribution of the interactome proteins of SS18L2 and SS18. Comparison was shown with log fold change (Log2) and -Log10 (p values) are indicated by size, taking the p values of SS18L2 interactome into consideration. BAF complex members in the comparison are indicated with orange on the graph. Analysis is conducted

with MaxQuant and Cassiopeia. **B.** Graph showing the distribution of the interactome proteins of SS18L2 and SS18L1. Comparison was shown with log fold change (Log2) and -Log10 (p values) are indicated by size, taking the p values of SS18L2 interactome into consideration. BAF complex members in the comparison are indicated with orange on the graph. Analysis is conducted with MaxQuant and Cassiopeia. Un-transfected and biotinylated MDA-MB-231 cells used as control.

Table 3.8. LogFC values of the highest expressed BAF complex proteins in the interactomes of SS18L2, SS18 and SS18L1.

Proteins	LogFC SS18L2	LogFC SS18	LogFC SS18L1
ARID1A	8.71	9.11	9.39
ARID1B	10.26	10.08	10.44
BCL7C	8.15	8.07	8.42
BRD9	3.81	4.63	4.54
DPF2	8.26	7.99	7.58
GLTSCR1	8.77	8.76	9.38
SMARCA2	8.45	9.77	10.36
SMARCA4	7.45	7.76	8.06
SMARCB1	5.34	5.32	5.20
SMARCC1	11.48	11.77	12.00
SMARCC2	8.56	8.77	9.58
SMARCD1	9.75	9.85	10.04
SMARCD2	9.74	9.40	9.51

Chapter 4:
DISCUSSION

Triple negative breast cancer is the most aggressive subtype of breast cancers that occur with the absence of Estrogen, Progesterone and Her2 receptors. Due to the lack of specific hormone receptors, conventional targeted therapies fail to work (Yin et al., 2020). Because of the low efficiency of conventional therapies, poor prognosis and the complexity of the cancer type, we decided to study the epigenetics of TNBC (Almansour, 2022). An uncharacterized epigenetic modifier from our previous studies, SS18L2, was indicated to decrease TNBC cell viability and arresting the cells at G2/M phase upon its knockout. In addition, the study resulted in a difference where SS18L2 KO significantly decreased TNBC cell viability when the KO of its homologs SS18 and SS18L1 did not have an effect on cell survival (Yedier-Bayram et al., 2022). The limited knowledge in the literature about SS18L2 evoked the question of "how does SS18L2 affect the TNBC cell fitness and viability?". As an initial approach to the question, the interactome of SS18L2 was investigated to further understand the characteristics of this epigenetic modifier.

The interacting partners of SS18L2 and its homologs was analyzed by proximity labelling in this study. BioID approach was utilized to further characterize SS18L2 due to a longer biotinylation period that detects the whole interactome of the protein, even the weaker and transient interactions (Sears et al., 2019).

Generating and confirming insert- BirA* BioID constructs

As the first steps in the BioID procedure, the SS18L2-BirA*, SS18-BirA* and SS18L1-BirA* fusion proteins were successfully engineered into a lentiviral vector with three stepped cloning, resulting in 15 engineered plasmids (**Appendix A: Maps Generated for Fusion Plasmids**). The N or C terminal plasmids are both considered, engineered, and tested before further experiments (Sears et al., 2019). In the case of SS18L2 lacking a DNA binding (QPGY) domain, a specific terminus for the interactions of SS18L2 is not known in the literature. Therefore, both plasmids that have the BirA* mutant in their different termini was utilized.

MDA-MB-231 cells were transduced with the OE plasmids SS18L2-NBioID-PLEX307 and SS18L2-CBioID-PLEX307; moreover, with SS18-NBioID-PLEX307 and SS18L1-NBioID-PLEX307. The overexpression of the fusion proteins was confirmed by western blotting of the SS18L2 constructs first. SS18L2-NBioID was effectively overexpressed and was visible in both cytosolic and nuclear fractions with relative controls (**Figure 3.1**). However, the repeated western blot experiments all resulted in two bands of the SS18L2- CBioID in all the fractionations, different from the expected size 46 kDa. This could be due to the BioID fusion perturbating the bait protein (Mehus et al., 2016). The disruption alters the protein structure, causing the lysine residues to be unavailable for the normal interactions or for various post-translational modifications. As a result of the degradation occurring in one of the constructs, it is not reliable to proceed the experiments involving SS18L2- CBioID construct. Therefore, the experiments were conducted with the N-BioID constructs for all of the inserts. The fusion plasmids SS18-NBioID and SS18L1-NBioID were also effectively overexpressed.

Localization of SS18L2 to nucleus

The examination of BioID fusion constructs also included the assessment of the localization of the fusion proteins. The control group BirA* only was expectedly expressed everywhere in the cell due to a promiscuous biotinylation through the cell without a bait protein (**Figure 3.5**). SS18 (dos Santos et al., 1997) and SS18L1 (Pradhan & Liu, 2005) are shown to localize to nucleus in the literature. We illustrated that the fusion proteins SS18-NBioID and SS18L1-NBioID were localized to nucleus in accordance with our expectations (**Figure 3.6**). However, the localization of SS18L2 was not investigated by any scientific studies. Our studies have identified a localization of SS18L2 being mainly in nucleus, for both N and C-BioID constructs (**Figure 3.4**). These results all the validate the successful cloning and overexpression of our constructs.

Further, the same localization to nucleus in both SS18L2-NBioID and SS18L2-CBioID lead to a conclusion that there are not any significant differences between these fusion proteins. Because of the observed degradation of the C-BioID construct and no extreme difference in localization, the decision to proceed with the N-BioID constructs was supported.

Validating biotinylation and streptavidin pull down

Following the BioID procedure, biotinylation and pull down with streptavidin must be validated before sending samples to mass spectrometry analysis. The biotinylation was optimized by trials with the biotinylation of SS18L2-NBioID OE in HEK293T and MDA-MB-231 cells. Moreover, biotinylation was successfully conducted for MDA-MB-231 SS18L2-NBioID and BirA*-NBioID OE cells in the cytosolic and nuclear fractionations (**Figure 3.10**). Biotinylation was conducted for MDA-MB-231 SS18L2-CBioID cells as well, resulting in a thicker blot indicating more biotinylated proteins in the cell. The confirmational change of the protein structure due to degradation might be causing excess interaction of the SS18L2 protein, leading to higher levels of biotinylation. Moreover, with the validation of the biotinylation of SS18L2-NBioID and BirA*-NBioID OE cells, SS18-NBioID OE (**Figure 3.11**) and SS18L1-NBioID OE (**Figure 3.12**) cells were also successfully biotinylated and confirmed with western blot experiments.

After biotinylation, the pull-down efficiency of the biotinylated proteins was tested. The streptavidin beads and protein correlation were optimized. Moreover, we have validated the pull-down experiments for all of the OE cells: SS18L2-NBioID, BirA*-NBioID (**Figure 3.14**), SS18-NBioID (**Figure 3.15**) and SS18L1-NBioID (**Figure 3.16**) for both cytosolic and nuclear fractions. The biotinylation of the bait protein, which is expected to be one of the highest interactions due to the fusion with the mutated BirA protein (Roux et al., 2018), was also tested with western blot. It was possible to pull down the bait protein, SS18L2 in both cytosolic and nuclear fractions (**Figure 3.14**).

Mass Spectrometry analysis revealing an interactome of SS18L2 related to BAF complex

A pilot experiment was conducted with biotinylated MDA-MB-231 cells and BirA* only overexpressed and biotinylated cells as control groups. Both cytosolic and nuclear fractionations of SS18L2-NBioID were analyzed by mass spectrometry. The analysis of the cytosolic fraction resulted in lower LogFC values, even for the log fold change of SS18L2 itself. On the other hand, the nuclear fractions resulted in a high LogFC value for SS18L2, indicating that the fusion protein could biotinylate itself and the hits related are

reliable. This shows that the experiment was successfully conducted. Moreover, we have shown the localization of SS8L2 to be nuclear, as expected. In addition, the control groups for further biological replicates were decided to be nuclear fractions. There were more contaminant cellular proteins detected in the analysis of SS18L2 interactome when BirA* was expressed as a control group. Differently, when biotinylated MDA-MB-231 cells was the control group, there was an equal distribution of the positive and negatively enriched proteins, with a smaller number of unrelated proteins with high log fold change values. Therefore, biotinylated MDA-MB-231 cells were chosen to be the control group for further experiments.

After conducting the pilot experiment sufficiently, all biotinylated and nuclear fractions collected from cells overexpressing SS18L2-BirA*, SS18-BirA* and SS18L1-BirA* fusions were analyzed by mass spectrometry. The suggested interactome from the mass spectrometry analysis of SS18L2 had BAF (SWI/SNF) complex, chromatin remodeling complex (Mashtalir et al., 2018), members as the highest hits with a high log fold change values and significance (**Table 3.2)**. The top interacting proteins included common BAF complex member proteins as well as proteins specific to canonical and non-canonical BAF. Canonical BAF complex members were ARID1A and ARID1B, and non-canonical BAF complex member was GLTSCR1. There were also PBAF specific members expressed in the proposed interactome. All of the interactions could indicate an interaction of SS18L2 with all of the BAF complex subtypes, especially canonical and non-canonical BAF.

Analyzing the interactome further with gene ontology analysis, processes such as chromatin organization or G/M phase transition were observed to be significant. The results are consistent with our previous findings, indicating that SS18L2 knockout arrests the cells at the G2/M phase (Yedier-Bayram et al., 2022). Therefore, a relation of SS18L2 to cell cycle could be inferred. Moreover, the chromatin and chromosome organization related pathways indicate and support the relation of SS18L2 to BAF complex in TNBC.

Mass spectrometry analysis of SS18 and SS18L1 also resulted in BAF complex members as the highest hits, with a high log fold change value and significance (**Table 3.3, Table 3.4**). Both proteins were also seen to be interacting with pBAF members as well. The highest hits from SS18 and SS18L1 were most significantly related to canonical and non-canonical BAF complex, as shown by tandem affinity purification before, supporting the reliability of our mass spectrometry analysis (Middeljans et al., 2012). The

article by Middeljans et al. demonstrates the interacting proteins of SS18 and SS18L1 in human cells to be related to all subtypes of the BAF complex, similar to our suggested interactome proteins.

Identifying SS18L2-specific interactions, different from its homologs SS18 and SS18L1

To find an interacting partner specific to SS18L2 and find a lead to interpret the role of SS18L2 in TNBC cell viability, comparisons between SS18L2 and its homologs were conducted. Since we have seen functional differences between the homologs in our previous study, a different interacting partner was expected to be found (**Figure 1.6**).

First, we checked the hits that are unique to SS18L2, not being expressed in SS18 or SS18L1. The comparison resulted in 17 different proteins unique to SS18L2, including transcriptional factors, a kinetochore protein, and cellular non- significant proteins (**Table 3.5**). The protein that was interesting to identify was NELFA, a transcriptional elongation factor which was shown to be induced by ERK activation causing tumor formation and progression in a study (Ohe et al., 2022). The interaction therefore could have been essential for explaining the role of SS18L2 in tumor cell viability. However, no connection was made due to the slightly lower expressions and non-significance, and the unique 17 proteins were not followed up.

Second, the comparisons were analyzed regarding the expression rates of the proteins, to find out an interaction partner that interacts with SS18L2 higher than its homologs. The interacting partners indicated in the suggested interactome of SS18L2, which are related to BAF complex were mostly negatively enriched when SS18 or SS18L1 was the control group. The highest expression with significance was SS18L2, which could be due to the higher biotinylation levels experimental and cannot result in a conclusion that SS18L2 is not interacting with its homologs. However, the high difference in expression for all proteins could be a lead for future experiments like co-immunoprecipitation and could also be interesting to determine whether these homologs interact with SS18L2 directly.

Identifying HAUS6 and HAUS8 as possible interacting partners of SS18L2

The comparison analysis resulted in proteins HAUS6 and HAUS8, which were also depicted as the top 15 highest hits in the proposed interactome of SS18L2 (**Table 3.2**). Moreover, the proteins were identified as the highest hits in the comparison analysis with SS18 as the control group (**Table 3.6**). The proteins are also significant in the analysis with SS18L1 as the control group but with a lower LogFC value. The proteins identified in the analysis were all compared with the previously established SS18L2 interactome to select the significant proteins with consistently high expressions across all analyses for SS18L2. Therefore, proteins with lower LogFC values and lacking significance in the SS18L2 interactome were not followed up.

HAUS proteins are subunits of the Augmin complex, also known as the homologous to augmin subunits (HAUS) complex with 8 different subunits (Gabel et al., 2022). The HAUS/augmin complex was shown to be required for mitotic spindle integrity in cell division with live cell imaging (Lawo et al., 2009). A study by Cunha-Ferreira et al. suggests the augmin complex subunits colocalized with γ-TuRC, which is a γ-tubulin subunit that mediates microtubule nucleation (Cunha-Ferreira et al., 2018). The article proposes that HAUS/augmin complex regulates microtubule nucleation in neuronal development. HAUS6 subunit was also shown to contribute to the microtubule binding by cryo-EM studies (Gabel et al., 2022). In addition, the oncogenic role of HAUS6 was demonstrated, where its knock-down reduced tumor growth in colorectal cancer by downregulating p21 and activating the p53 pathway (Shen et al., 2021). Moreover, Kraus et. al. has shown that augmin complex is regulated by Ran-GTP and binds to importin α and importin β. Further, two NLS sites on HAUS8 that binds to microtubules were found (Kraus et al., 2023). The definite binding sites of HAUS6 are still unknown; however, the study indicates that the microtubule binding sites of HAUS6 and HAUS8 are both T-II (tetramer II) augmin subcomplexes (Kraus et al., 2023). The similarity could correlate to the transportation of HAUS6 into to nucleus as well. The nuclear localization signal is a supportive finding to our results and answer the question of how HAUS proteins may interact with SS18L2 in the nucleus. Further, the role of HAUS/Augmin complex related to microtubule formation could correlate with the role of SS18L2 in the cell cycle (**Figure 1.6**, **Figure 3.20**). Considering not only the oncogenic role of HAUS6 but also the

significance of HAUS6 and HAUS8 on mitotic spindle formation could be crucial to further investigate the interactions of SS18L2 with HAUS6 and HAUS8.

Unresolved questions, new hypotheses, and future directions

Suggesting that the interactome of SS18L2 comprised of BAF complex members in MDA-MB-231 cell line, which is similar to its homologs, raised several questions. The study suggested an interactome to characterize SS18L2; however, the question of the difference of SS18L2 from its homologs still needs further understanding. There are several approaches we can apply to resolve the remaining questions.

If the interactomes are highly similar and SS18L2 is only interacting with the BAF complex members, we might be seeing a difference in TNBC cell viability due to the expressions of the genes in the cells. To test this, QPCR (Real-time quantitative PCR) from genomic DNA could be conducted to find out the gene copy numbers of *SS18L2*, *SS18* and *SS18L1*. Another approach in determining the amplification levels is FISH (Fluorescence in situ hybridization), which is a technique utilized in determining the location of specific sequences, where probes can be utilized to detect gene amplifications (Bayani & Squire, 2004). For example, in a study conducted in 2006, the *HER-2/neu* gene amplification in breast cancer was evaluated by QPCR (Kulka et al., 2006). The study compares immunohistochemistry results to QPCR results and FISH results. The article suggests that detecting amplifications of a specific gene can be as accurate as FISH experiments. Therefore, the gene copy numbers can be verified by QPCR or with FISH experiments to detect if there are any differences between the homologs. High amplification levels of SS18L2 in TNBC cell line compared to a non-malignant cell line might indicate a significance in further searching a difference for the characterization of SS18L2. In addition to investigating the gene copy numbers, rescue experiments could be conducted, by overexpressing SS18L2 KO cells with SS18 or SS18L1 overexpression. If the cells are able to rescue the knockout, it suggests that the roles of SS18L2 to its homologs could be similar.

The structural differences between SS18L2 and its homologs is a fundamental area of investigation in order to determine the distinct role of SS18L2 in TNBC. It was previously known that SS18L2 lacks the QPGY domain (de Bruijn & Geurts van Kessel, 2006). The AlphaFold predictions also indicate the differences between the structures of

the three homologs (**Figure 4.1**) (*AlphaFold Protein Structure Database*, 2021). Furthermore, the amino acid sequences of SS18L2 are similar to the N terminus of SS18 and SS18L1 (**Figure 4.2**). A rescue experiment involving SS18L2 KO cells can be conducted by introducing only the N terminal homologous sites of SS18 and SS18L1, along with their full length. If N terminal sites of SS18 and SS18L1 can rescue SS18L2, whereas the full-length forms cannot, the results might suggest the different roles of the homologs, despite their similar SNH domain sites. The differential domain of SS18, which is the QPGY domain found in the C terminus of the protein, is reported to induce LLPS (Liquid-liquid phase separation) in the cells, inducing the formation of BAF complex (Cheng et al., 2022; Kuang et al., 2021). LLPS is a novel mechanism, which explains the compartmentalization of proteins into membranelles structures (Tong et al., 2022). In all of the homologs, N terminus SNH domain is the domain that interacts with the BAF complex (Kuang et al., 2021). Therefore, the lack of the QPGY domain in SS18L2 could result in aberrant levels of phase separation in the SS18L2 KO cells with the presence of SS18, which could be a difference to investigate further with microscopy techniques such as immunofluorescence or live cell imaging.

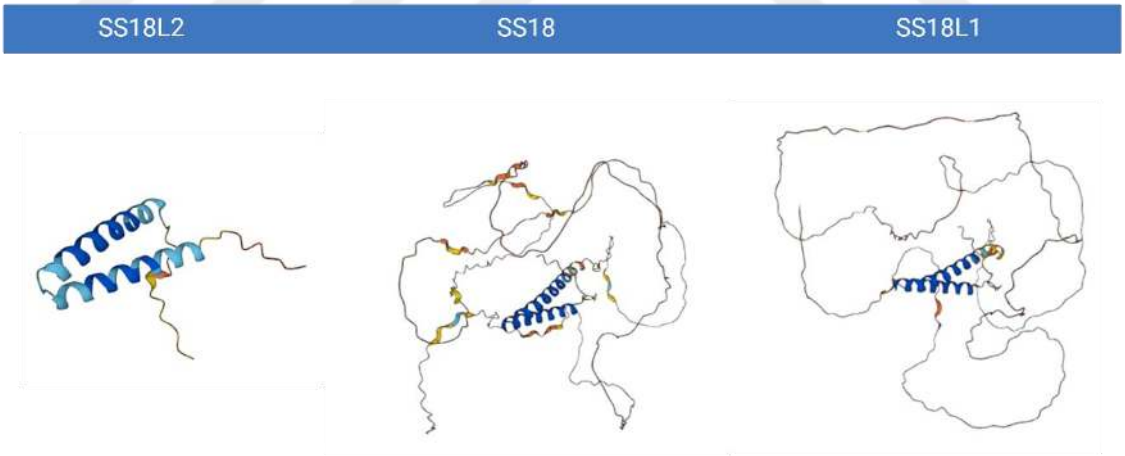

Figure 4.1. AlphaFold predictions of SS18L2, SS18 and SS18L1 (*AlphaFold Protein Structure Database*, 2021). Figure created with BioRender.com.

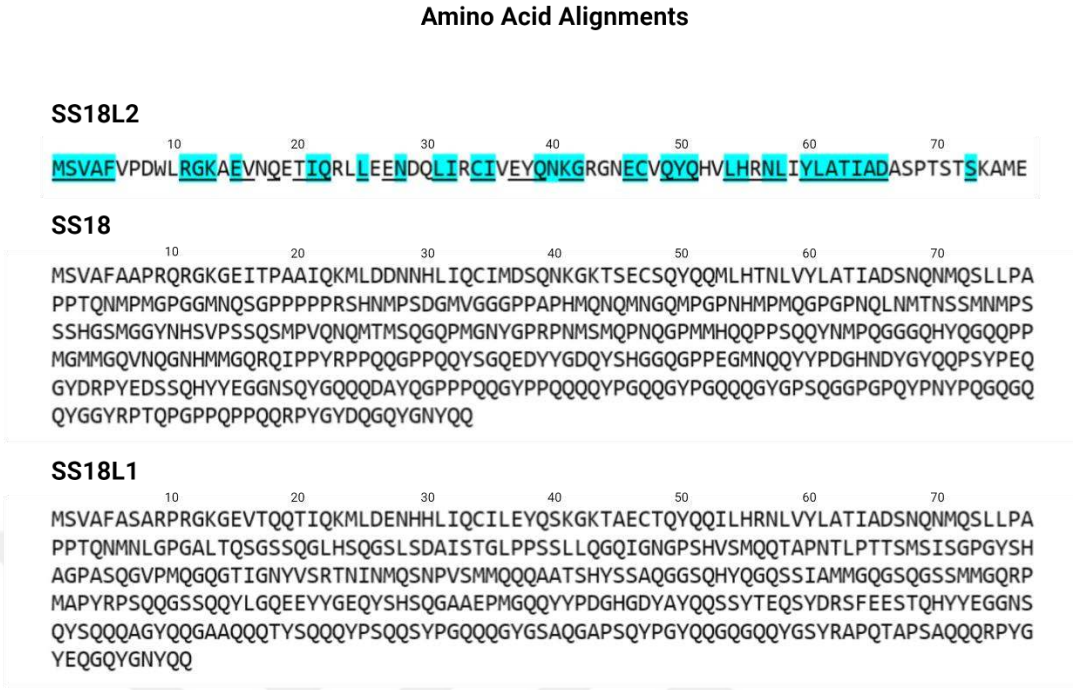

Figure 4.2. Amino acid sequence alignments of SS18L2, SS18 and SS18L1 (*AlphaFold Protein Structure Database*, 2021). Same amino acids in SS18L2 aligned with SS18 are highlighted with blue, aligned with SS18L1 are underlined. Figure created with BioRender.com.

Additionally, a correlation can be made with SS18 and its fusion with SSX proteins, forming synovial sarcoma as a major role in cancer. The fusion of SS18-SSX causes the loss of SMARCB1, by replacing the wild type SS18 in the BAF complex (McBride et al., 2018). Similarly, SS18L2 was speculated to compete with SS18 in the BAF complex because it lacks the transcriptional binding domain (de Bruijn & Geurts van Kessel, 2006). The competition and the interactions of SS18L2 with BAF complex members might be causing the BAF complex to activate the broad polycomb domains, similar in synovial sarcoma, and cause proliferation for TNBC cells (**Figure 1.5**). However, further functional assays need to be conducted to support this hypothesis.

Finally, the difference of SS18L2 from its homologs can be explained by database research. In Cancer Dependency Map (DepMap), SS18L2 was shown to be a common essential gene in the 1084 CRISPR screens out of 1095 while SS18 was strongly selective in 13 out of 1095 CRISPR screens and SS18L1 was strongly selective in 0 out of 1095 CRISPR screens (*Cancer Dependency Map*, 2021). The data might be indicating the differential role of SS18L2. In addition, the differential role of SS18L2 might be related

to the amplifications on the chromosome that are caused in TNBC. An article by Kawazu et. al. focuses on the genetic alterations of TNBC by whole genome sequencing. The study suggested structural variations in chromosome 19, chromosome 8 and chromosome 3 in TNBC (Kawazu et al., 2017). SS18L2 was shown to be located at 3p22.1 in DepMap, therefore, a correlation between the genomic location of SS18L2 to the amplifications in TNBC could lead for further experimentation as well. Besides CRISPR, SS18L2 could be silenced with siRNA or shRNA while viral transduction or transfection can result in off- target mutations by knock-out (Guo et al., 2023). In addition, since SS18L2 is an epigenetic regulator, where we propose an interaction with the BAF complex, it might be essential to detect the epigenetic modifications of SS18L2 in the chromatin level by ChIP (chromatin immunoprecipitation) assay (Gade & Kalvakolanu, 2012).

In conclusion, we have aimed to find the interacting partners of SS18L2 in TNBC because of the very little knowledge about the protein in literature. Moreover, we investigated the differences between SS18L2 and its homologs as SS18L2 had a differential effect on TNBC cell viability. The fusion constructs were successfully generated and all of them were validated. SS18L2 fusion proteins were demonstrated to localize in the nucleus. Moreover, the biotinylation and pull-down experiments were successfully optimized and validated. With our mass spectrometry analysis, we suggest an interactome of SS18L2, highly interacting with all three BAF complex subtypes (**Figure 4.3**). The differential role of SS18L2 from SS18 and SS18L1 still needs further investigation and could deepen the understanding of the role of SS18L2 in TNBC cell fitness.

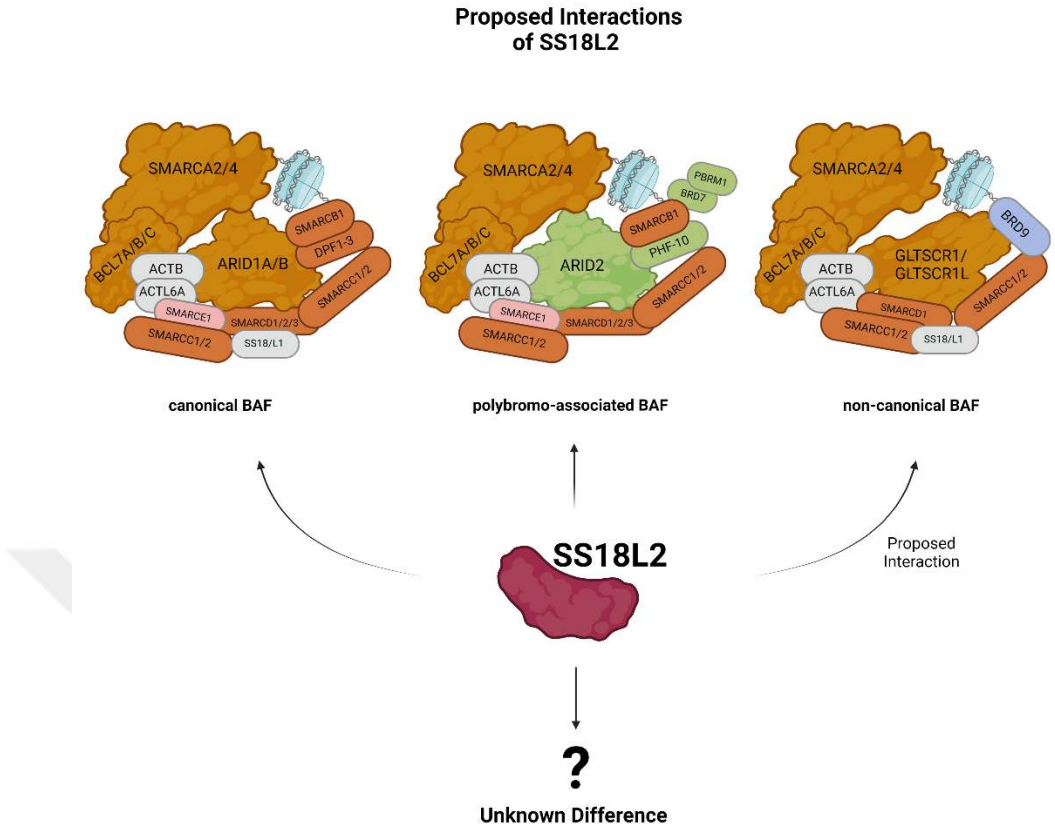

Figure 4.3. Proposed interactions of SS18L2 in MDA-MB-231 cells with BAF complex. The interaction partners that came up as the highest hit are shown in orange on the BAF complex subtypes. Weaker interactions are not shown. Shared subunits are in grey, subunits that are not found in ncBAF are red, subunits that are unique to pBAF are green and subunits that are unique to ncBAF are blue. Figure created with BioRender.com.

Chapter 5:

APPENDIX A: MAPS GENERATED FOR FUSION PLASMIDS

The plasmids generated in other to create lentiviral vectors with inser-BirA* fusion proteins are demonstrated with their maps created with SnapGene software. The plasmid maps correspond to the plasmids indicated in **Table 3.1.** Reaction numbers are indicated in **Figure 2.2.**

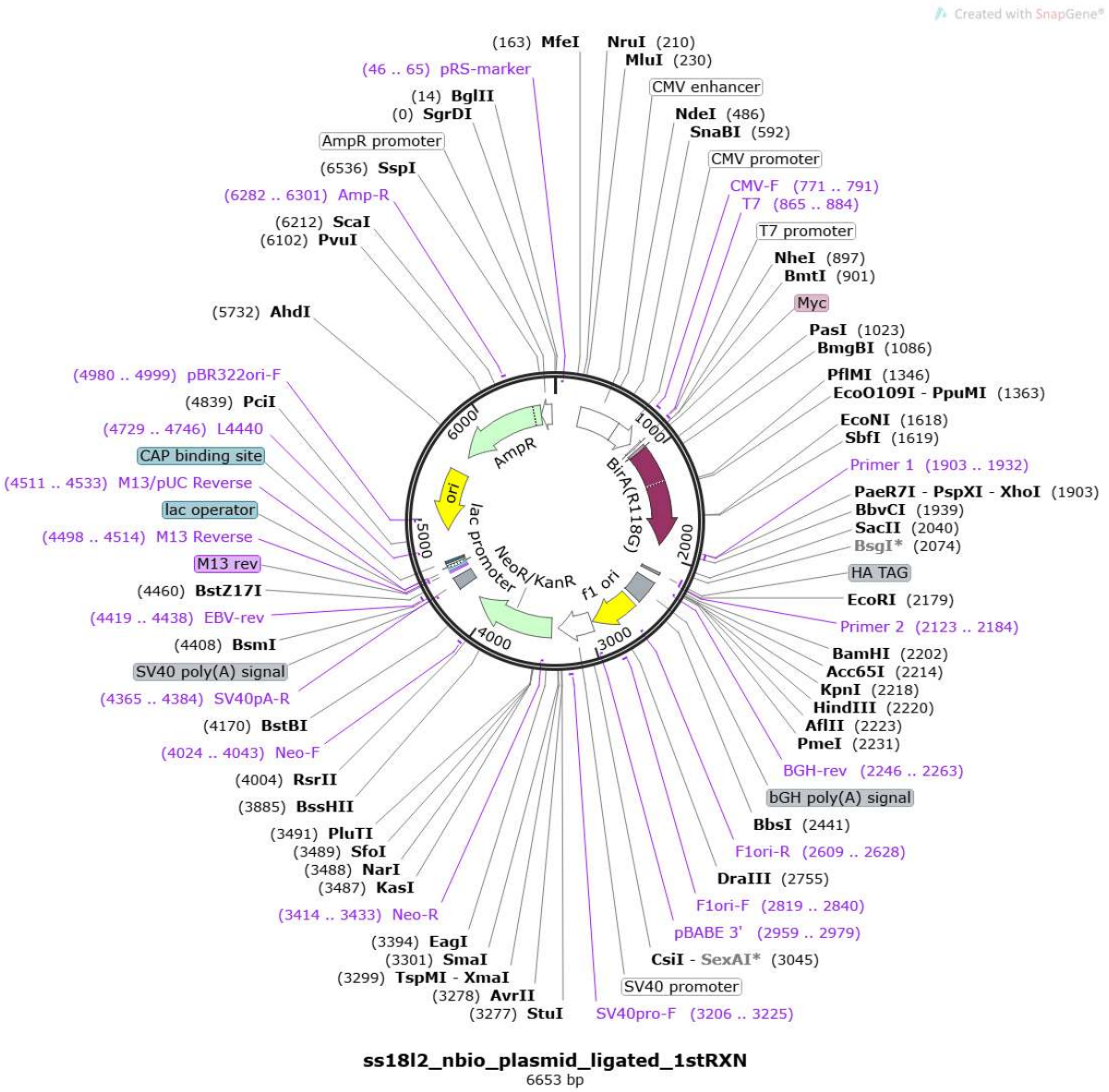

Figure 5.1. Plasmid map for NBioID-SS18L2 plasmid after first reaction.

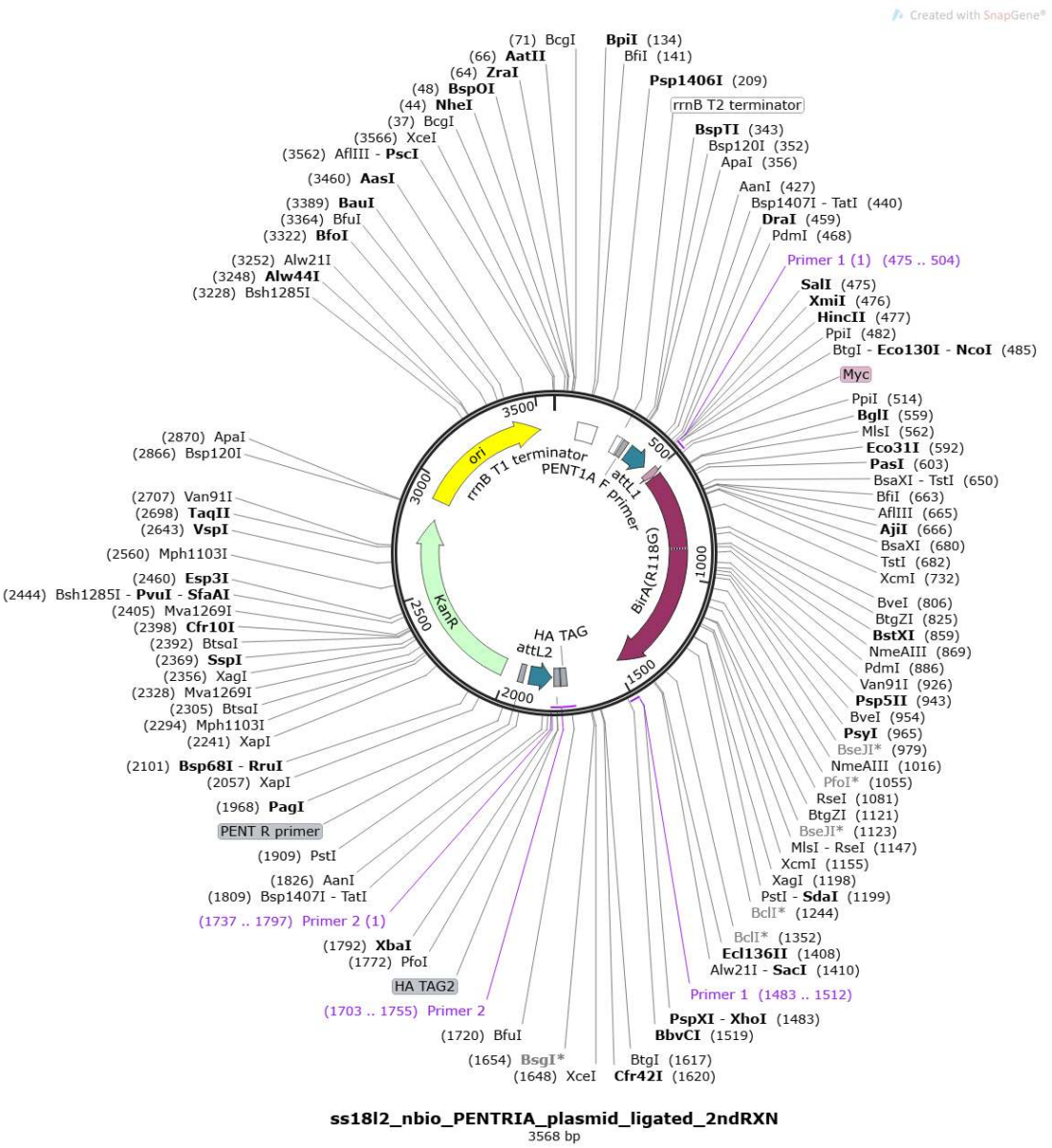

Figure 5.2. Plasmid map for SS18L2-NBioID-pENRTR1A plasmid after second reaction.

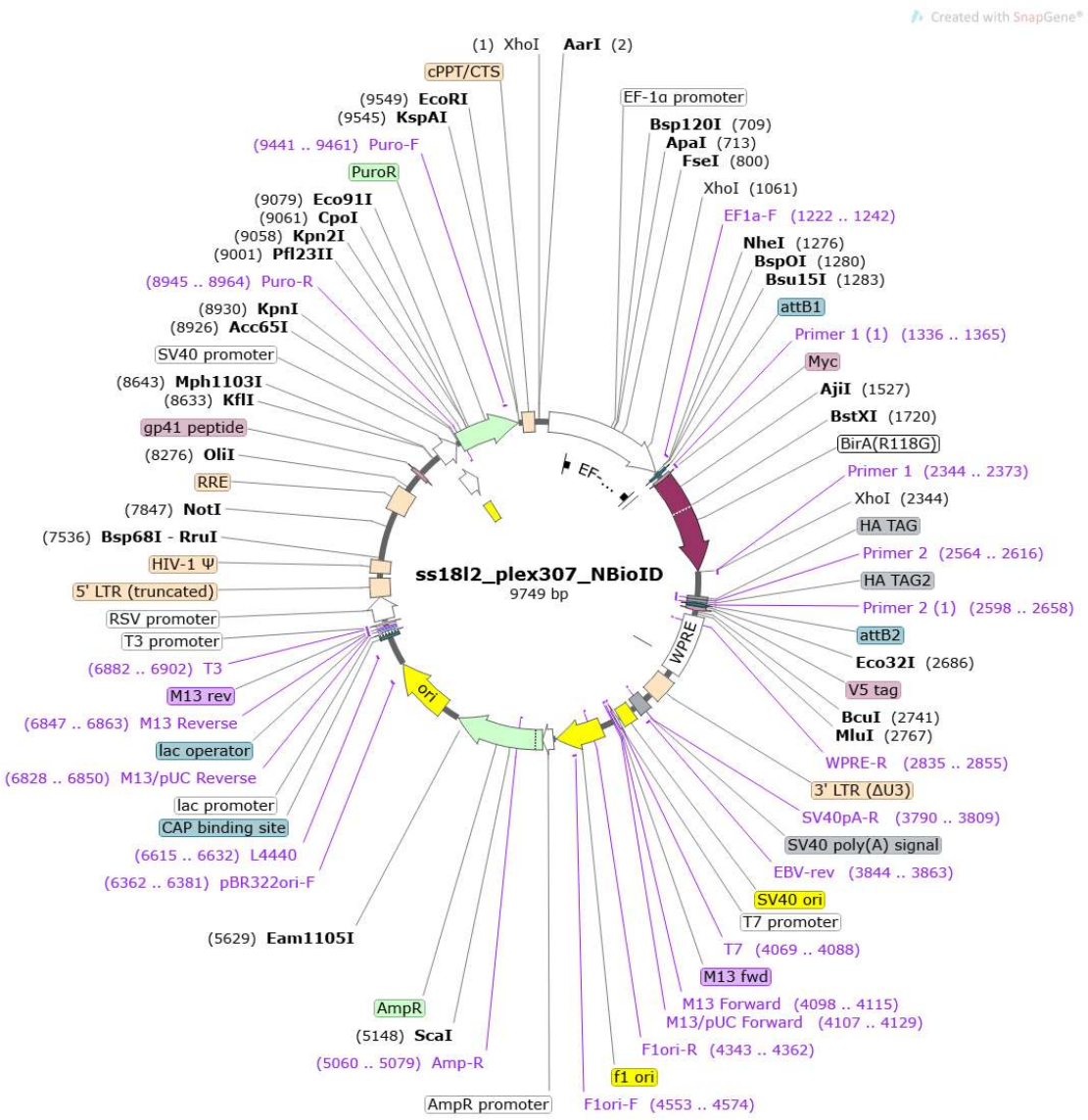

Figure 5.3. Plasmid map for SS18L2-NBioID-PLEX307 plasmid after third reaction.

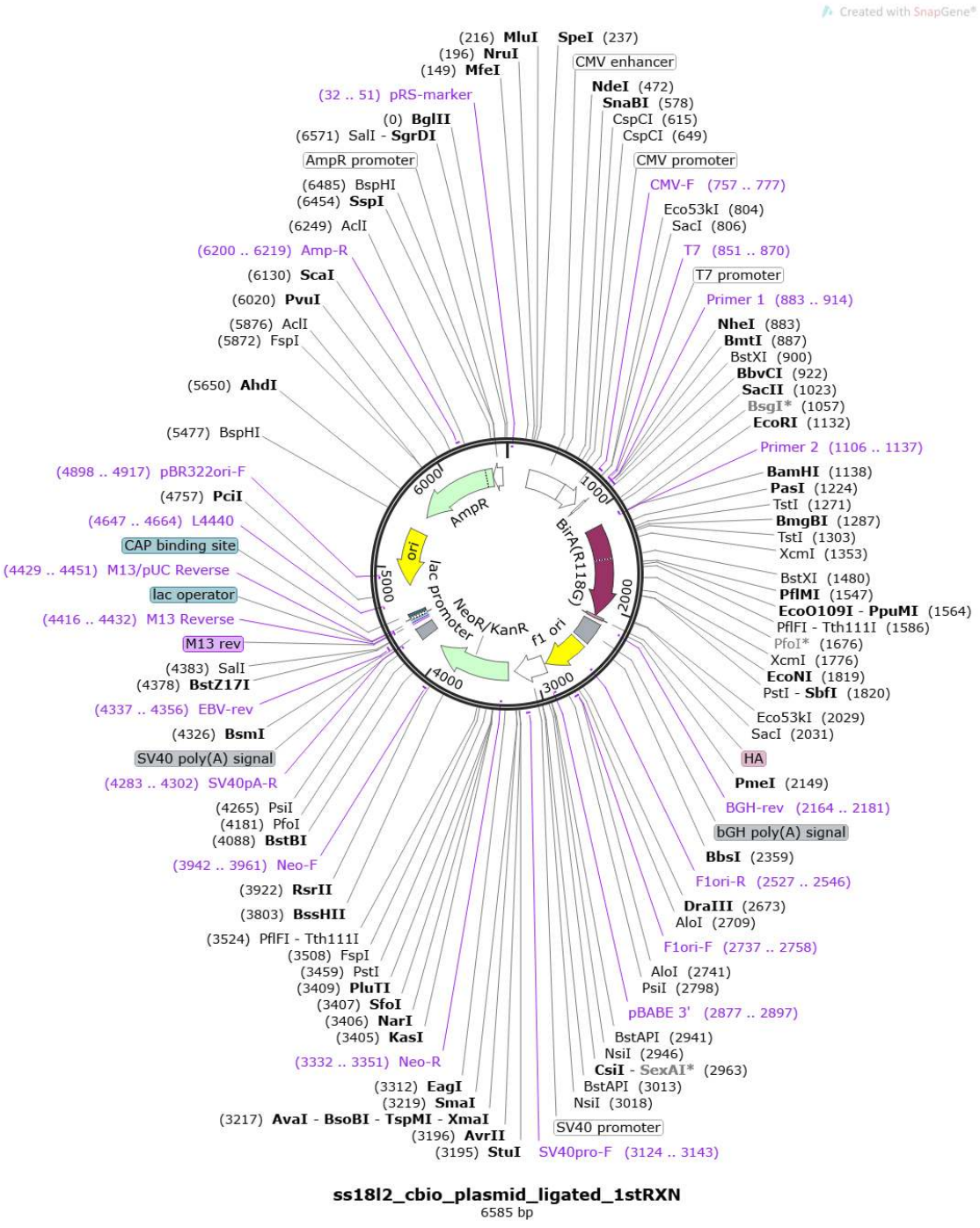

Figure 5.4. Plasmid map for CBioID-SS18L2 plasmid after first reaction.

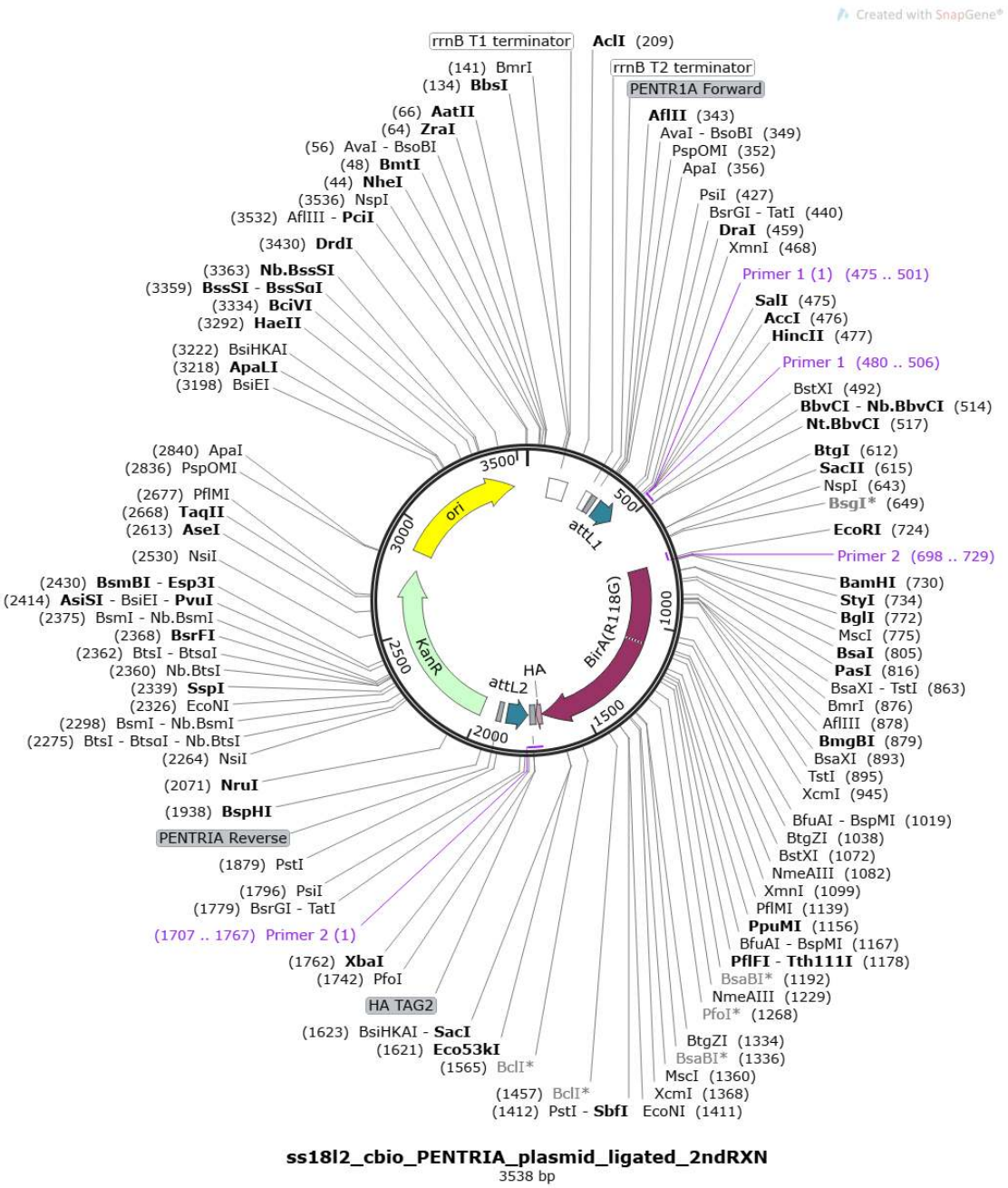

Figure 5.5. Plasmid map for SS18L2-CBioID-pENRTR1A plasmid after second reaction.

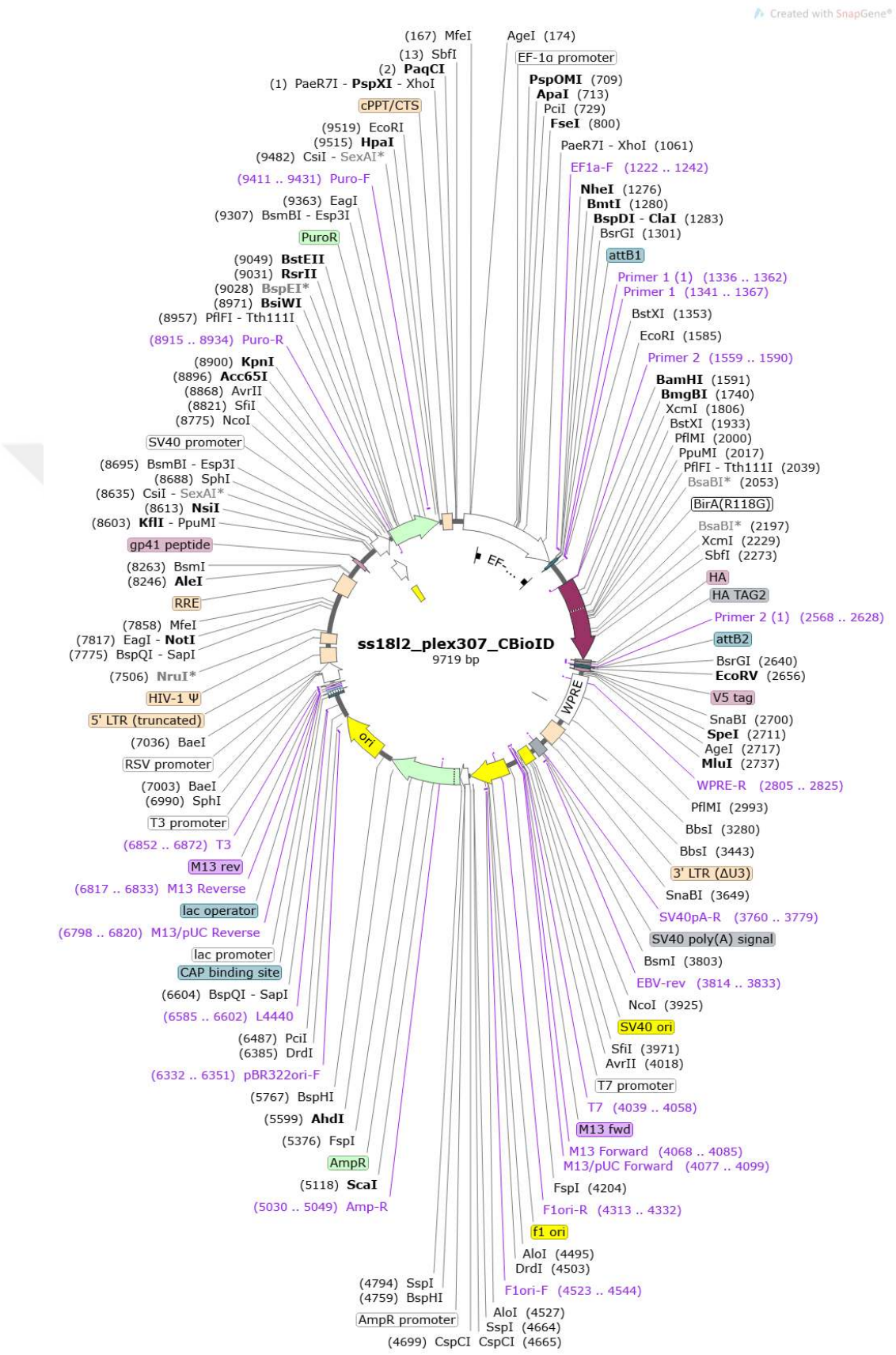

Figure 5.6. Plasmid map for SS18L2-CBioID-PLEX307 plasmid after third reaction.

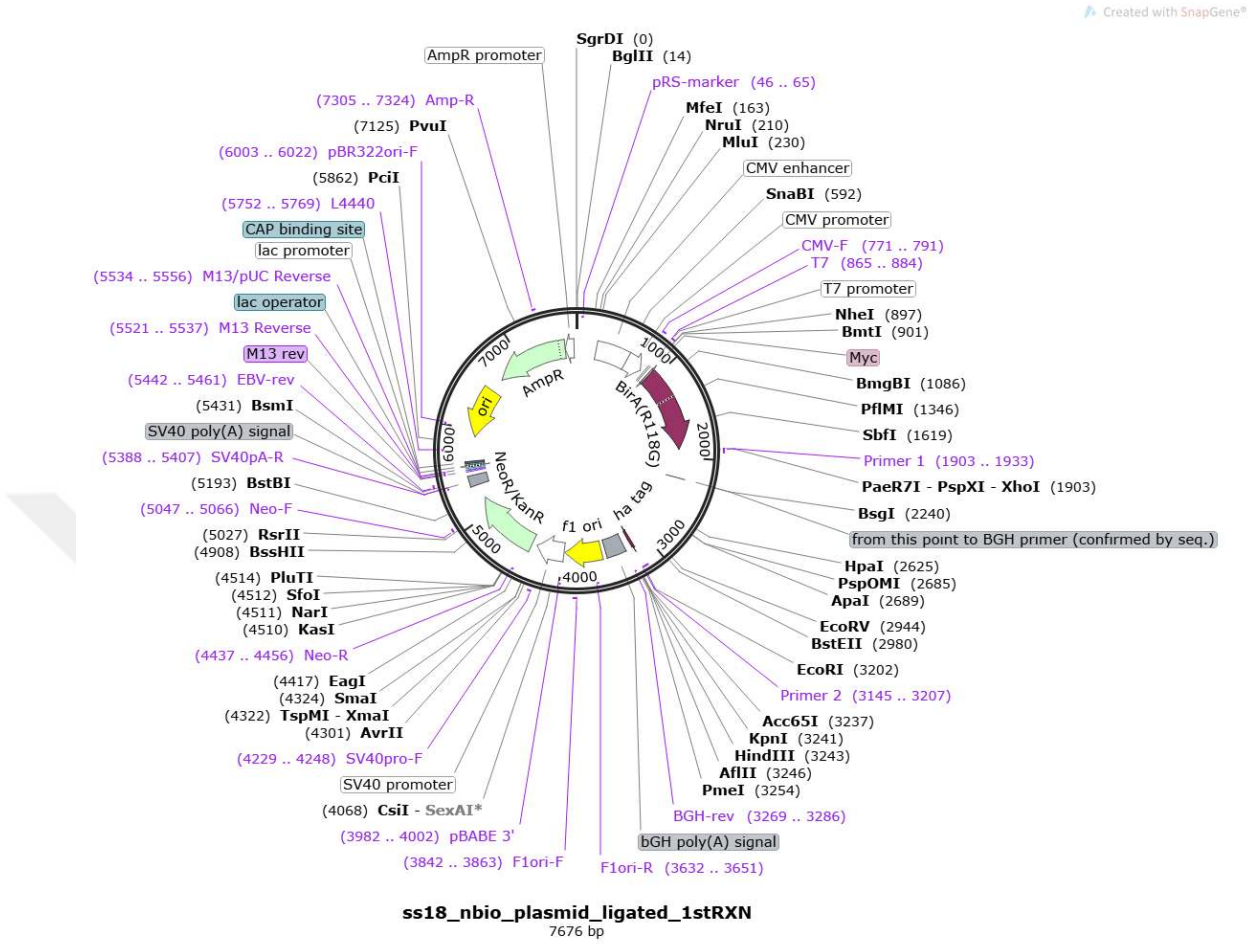

Figure 5.7. Plasmid map for NBioID-SS18 plasmid after first reaction.

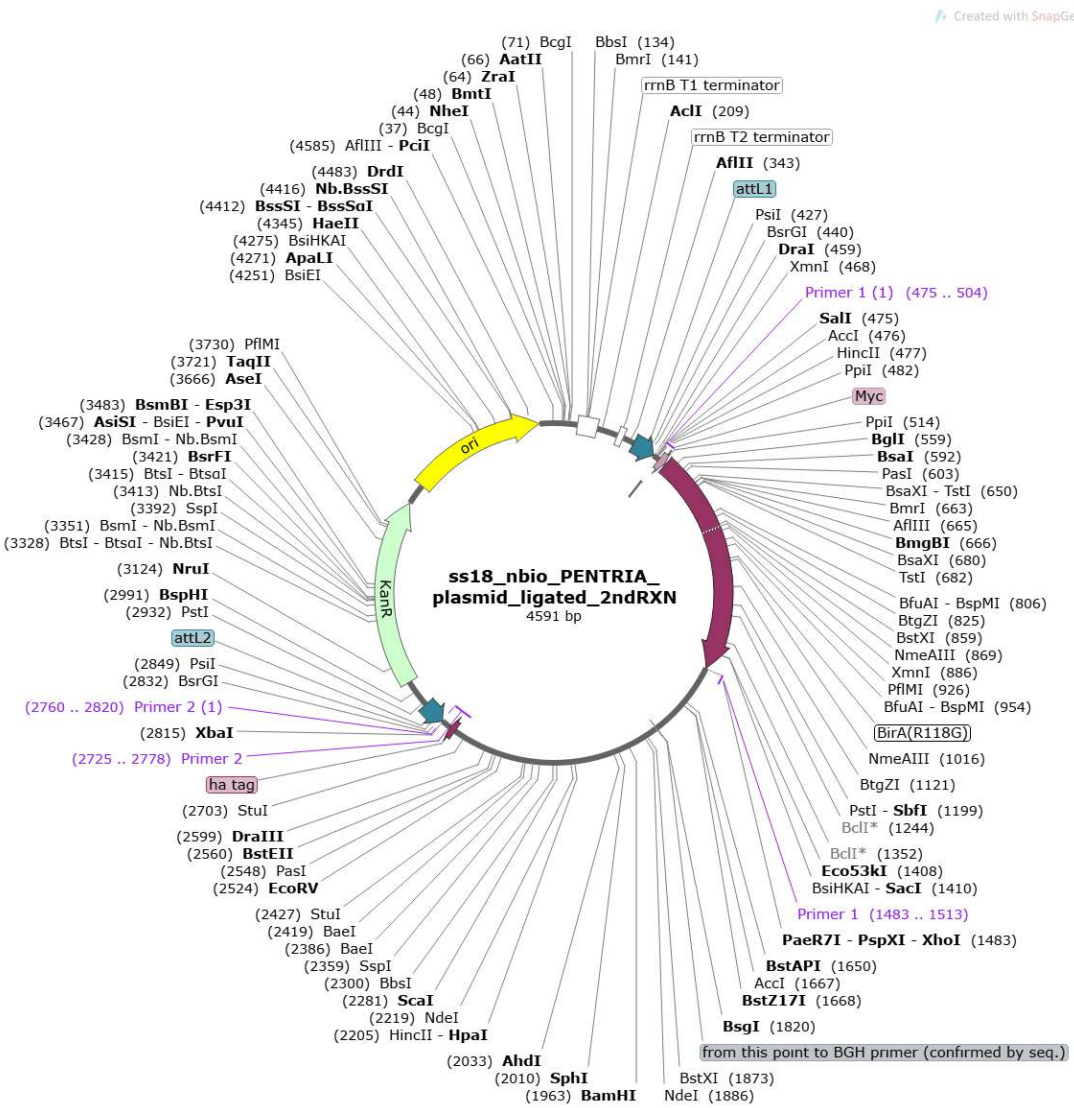

Figure 5.8. Plasmid map for SS18-NBioID-pENRTR1A plasmid after second reaction.

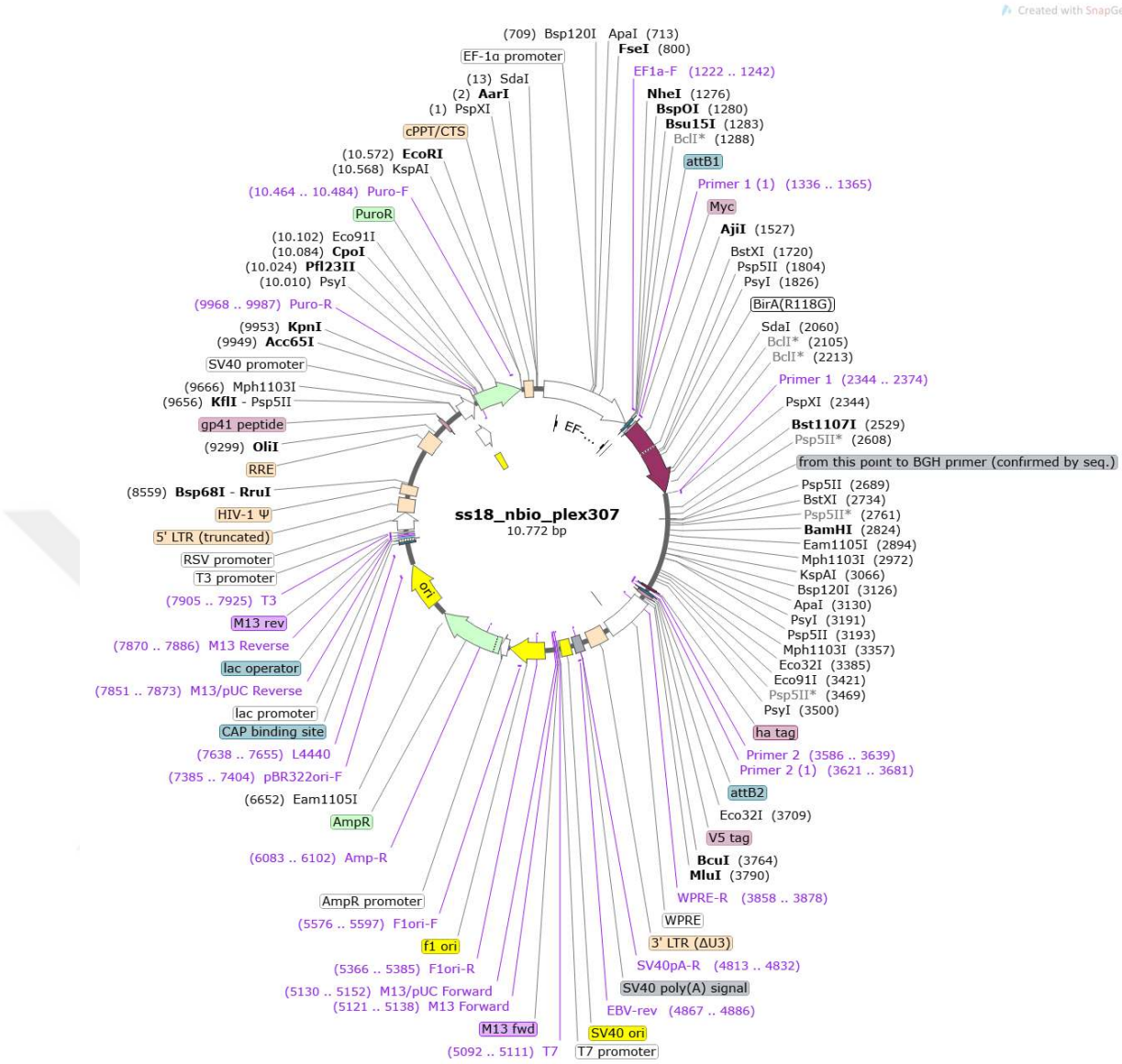

Figure 5.9. Plasmid map for SS18-NBioID-PLEX307 plasmid after third reaction.

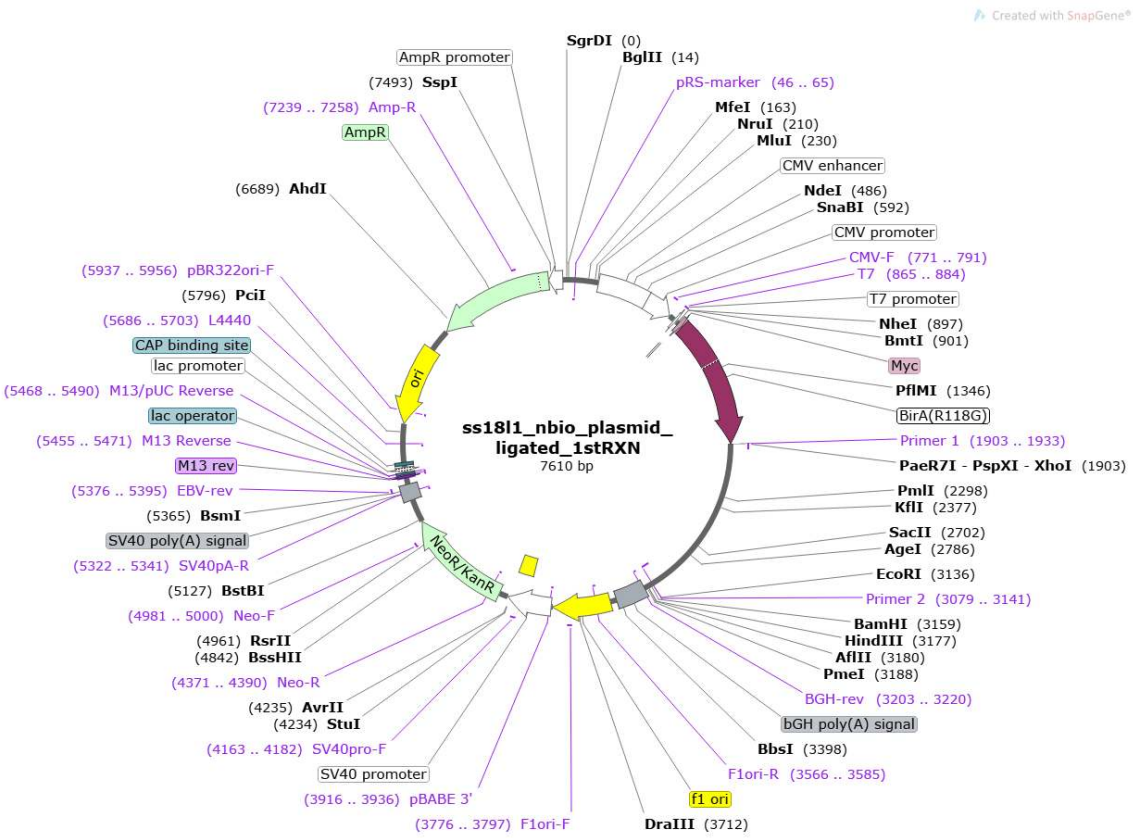

Figure 5.10. Plasmid map for NBioID-SS18L1 plasmid after third reaction.

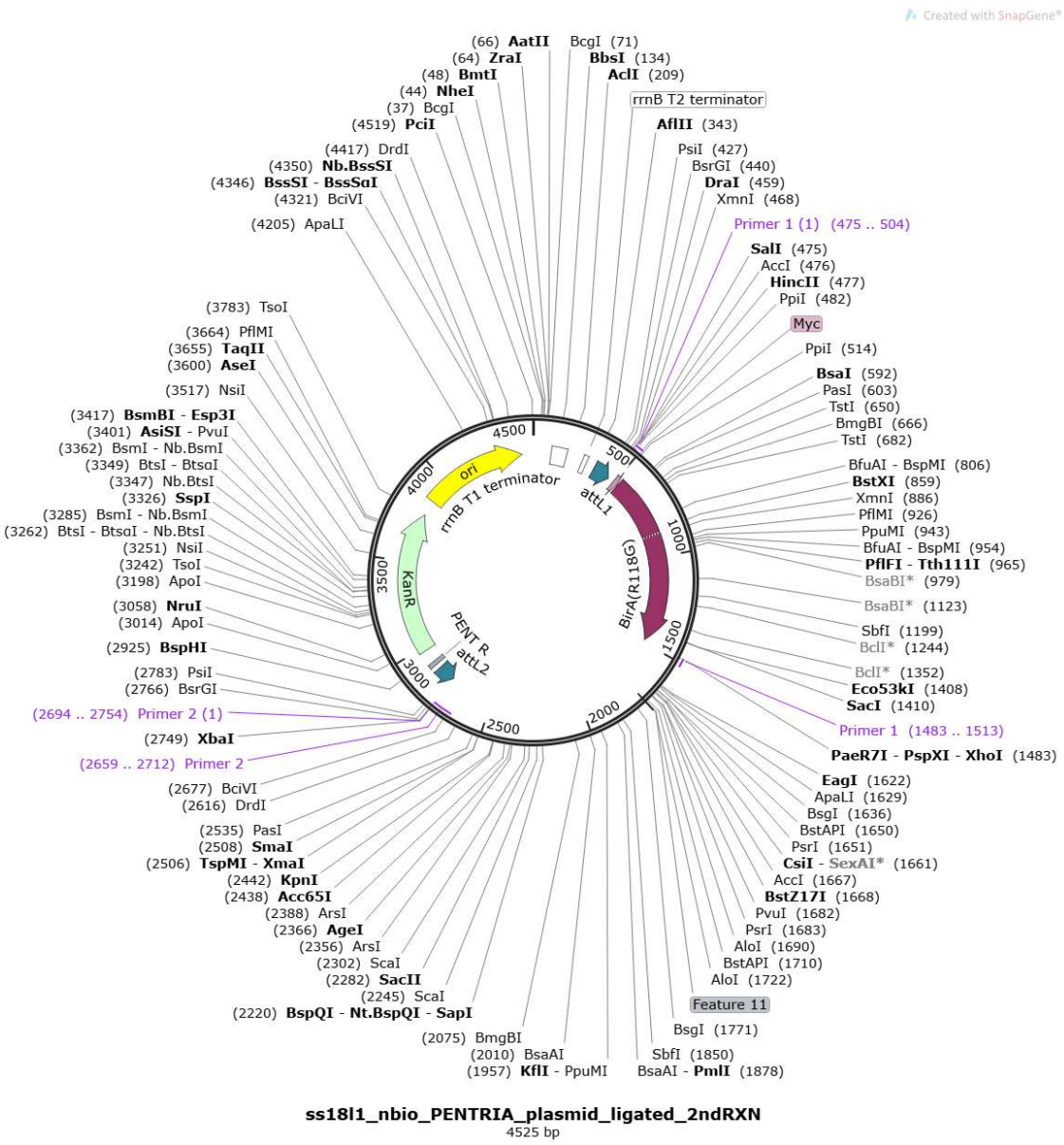

Figure 5.11. Plasmid map for SS18L1-NBioID-pENRTR1A plasmid after second reaction.

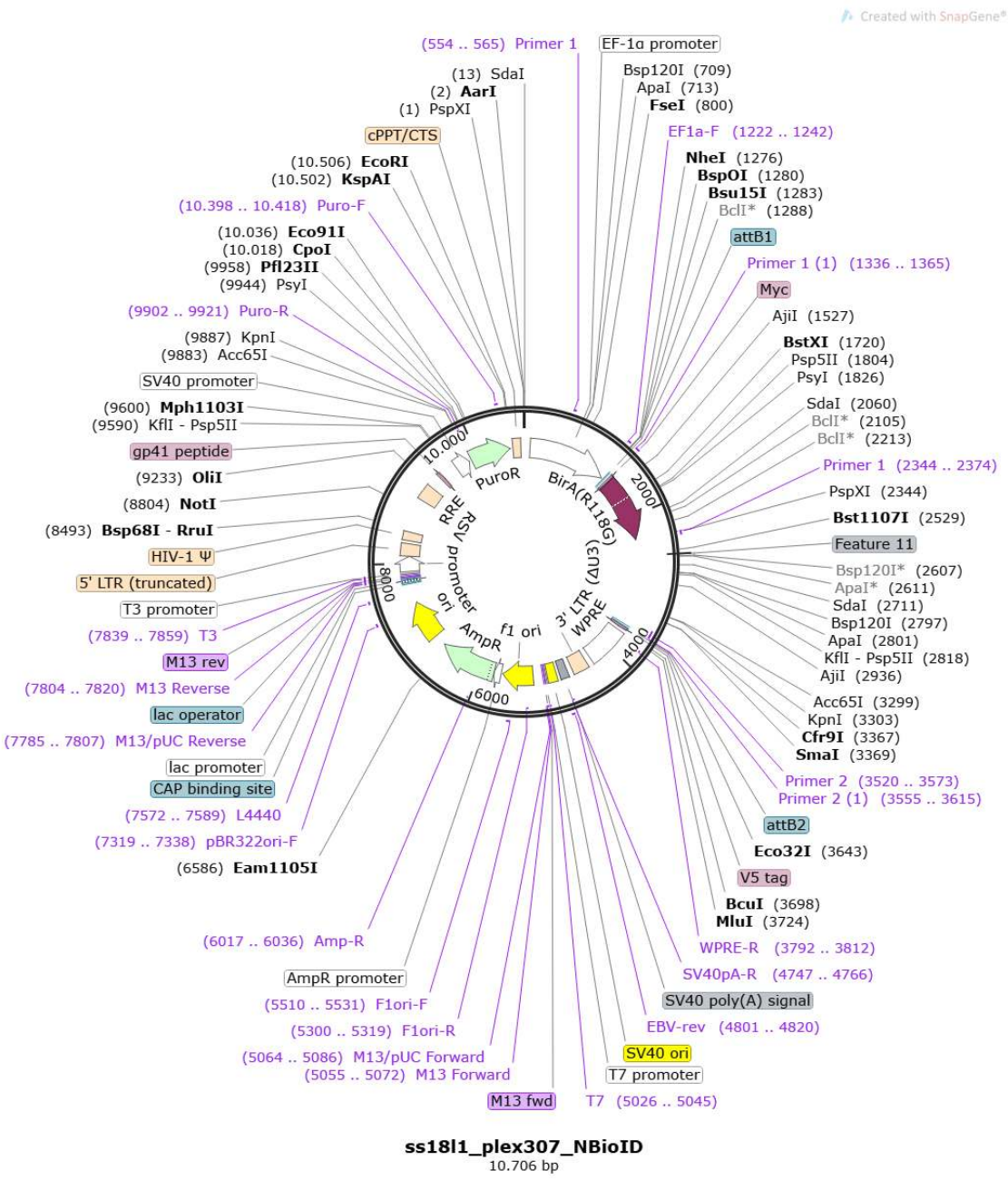

Figure 5.12. Plasmid map for SS18L1-NBioID-PLEX307 plasmid after second reaction.

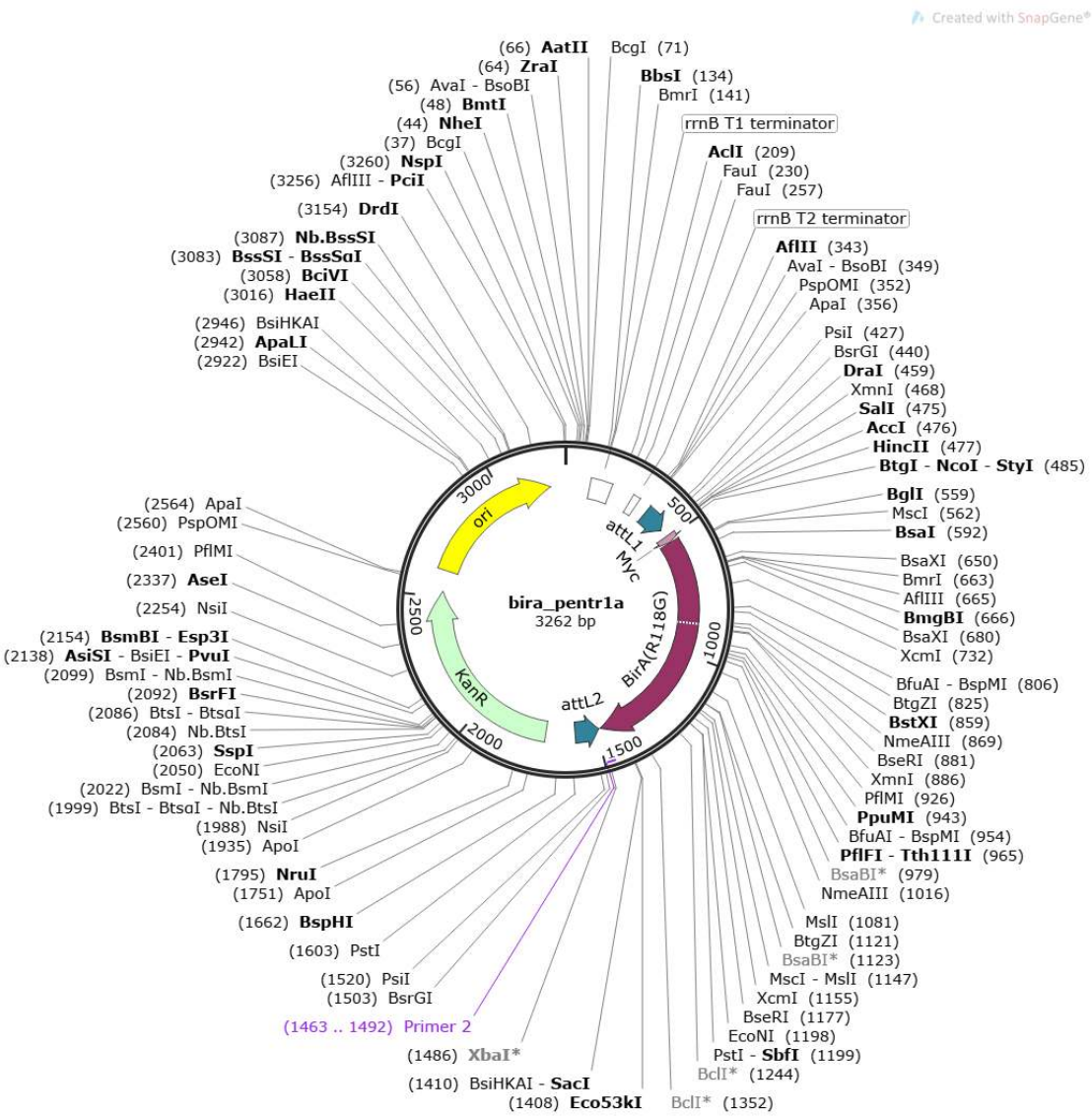

Figure 5.13. Plasmid map for BirA*-NBioID-pENTR1A plasmid after first reaction.

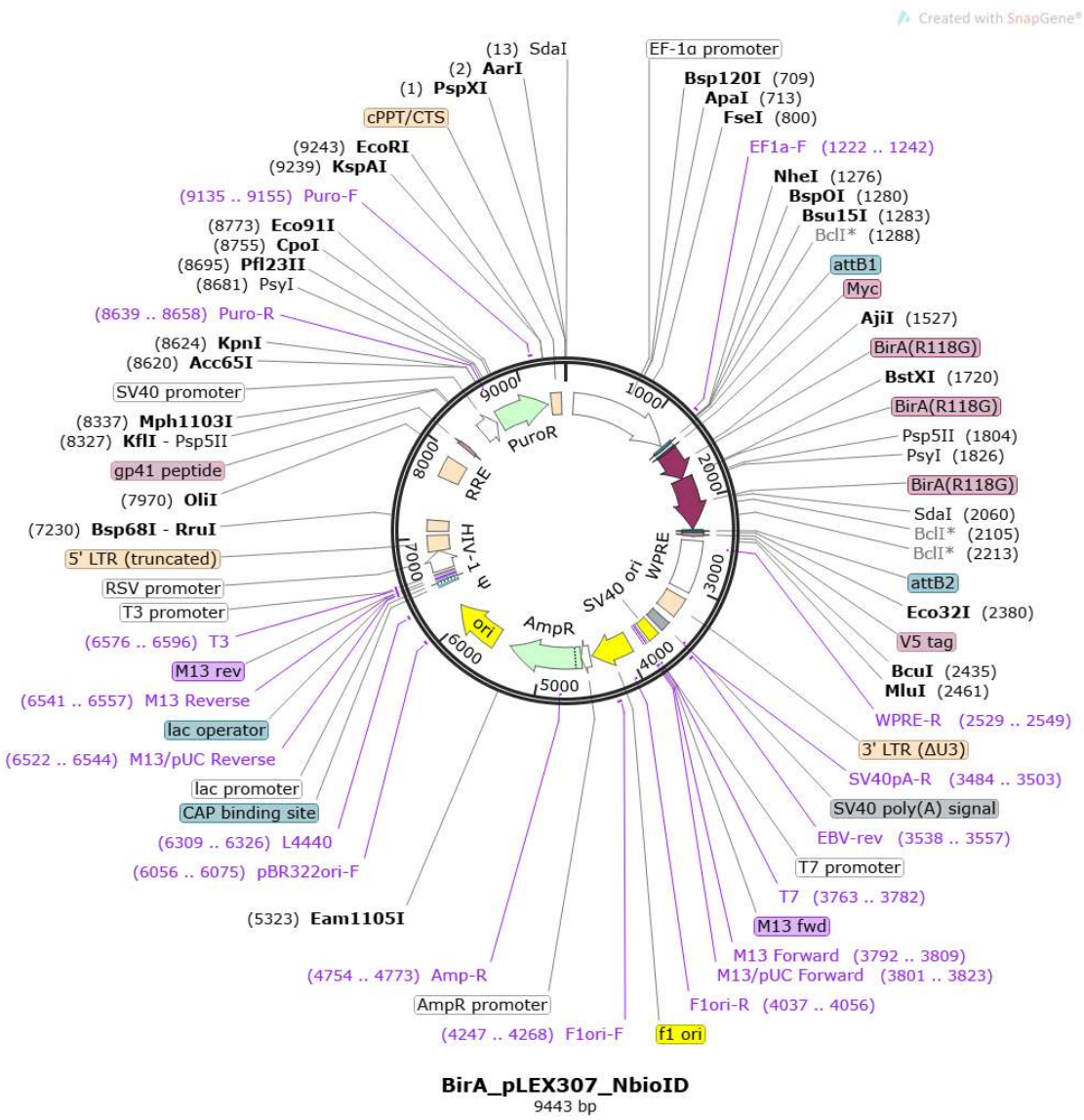

Figure 5.14. Plasmid map for BirA*-NBioID-PLEX307 plasmid after final reaction.

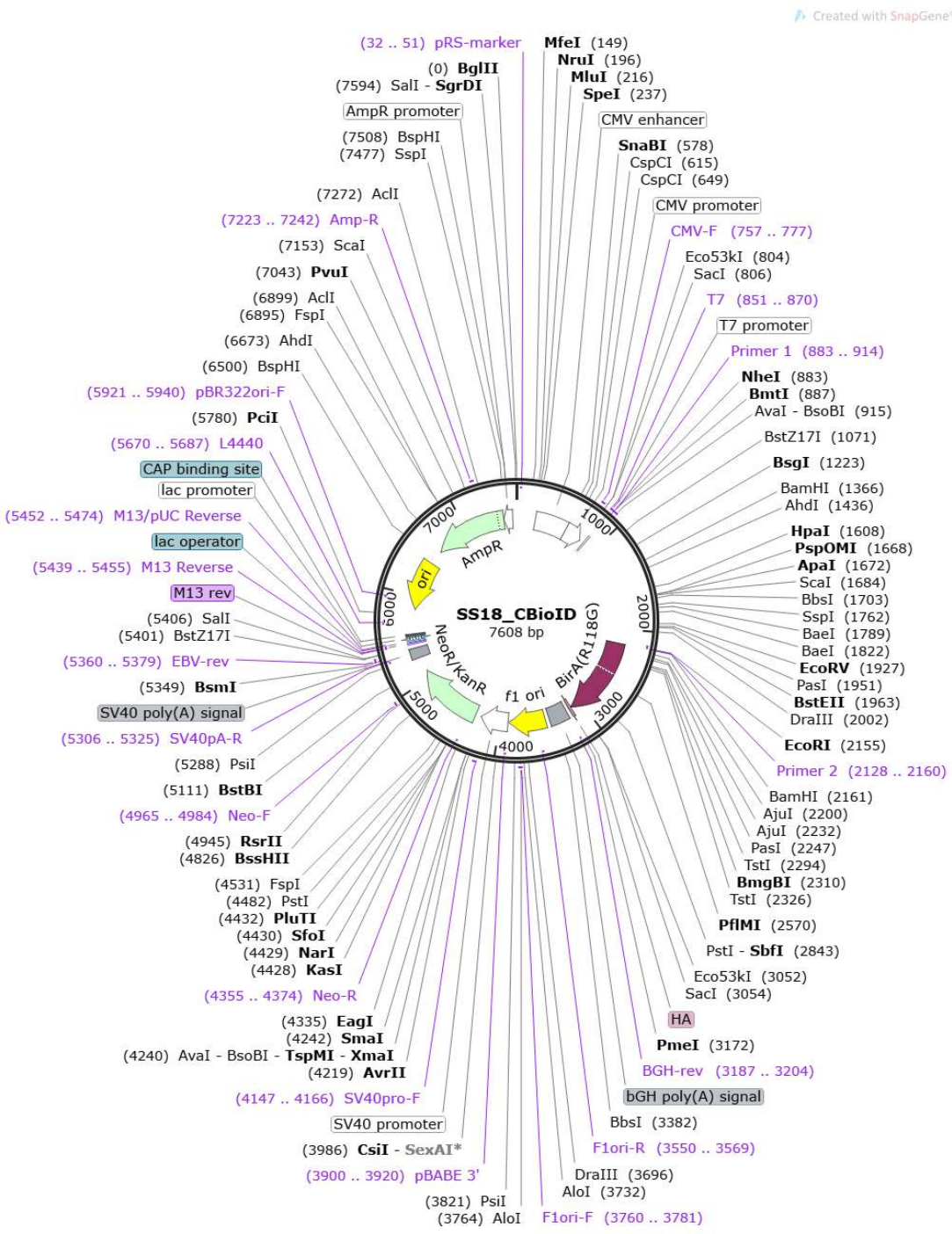

Figure 5.15. Plasmid map for CBioID-SS18 plasmid after first reaction.

Chapter 6:

APPENDIX B: MASS SPECTROMETRY RESULTS RAW DATA

The mass spectrometry analysis results created with MaxQuant (Max Planck Institute of Biochemistry) software and Cassiopeia LFQ software are demonstrated in the tables below. The comparison analysis results are sorted according to the LogFC values with a LogFC>2 cut off.

Table 6.1. Mass spectrometry analysis of SS18L2 interactome raw data. Analyzed with MaxQuant (Max Planck Institute of Biochemistry) software and Cassiopeia LFQ software. LogFC>2 as cut off.

Gene names	LogFC Values	Average Expression	P.Value
SMARCC1	11.48213636	26.07480098	1.95E-06
ARID1B	10.25882666	27.25563934	2.33E-05
SMARCD1	9.74523397	27.06511008	2.00E-06
SMARCD2	9.737819395	27.15288242	0.000493314
HAUS6	8.829588759	24.40453514	5.68E-05
GLTSCR1	8.770081432	24.18443065	8.17E-05
ARID1A	8.705394654	28.59824823	0.000144455
SMARCC2	8.557215093	25.60264494	0.000152386
GIGYF2	8.488533837	25.13016733	5.87E-05
SMARCA2	8.454475438	26.06888012	9.22E-06
DPF2	8.258688518	26.4944805	0.000138108
BCL7C	8.153627193	24.03857132	0.000335279
HAUS8	7.495605967	23.89833694	0.00014684
SMARCA4	7.445137497	29.7366563	1.02E-05
ZBTB33	7.407605986	24.60506602	0.000281588
SS18L2	7.196701901	25.26995532	0.005183013
YEATS2	7.037737148	23.47529467	0.000501025
SMARCE1	7.011189763	25.59204379	0.003420669
SDHB	6.131278953	22.54918852	0.000873726
EEF1D	5.955947412	23.26819081	0.000706697
SIN3B	5.933562248	24.15871915	0.000686416
MRE11A	5.81378994	22.0068686	0.00166448
SMARCD3	5.758967254	25.86089438	0.026298789
DDX46	5.742819333	25.25874007	0.010396629
CSTF2	5.560511589	23.02112242	0.001427791
CPSF6	5.438991947	21.70084259	0.005964613
CD2AP	5.406665126	22.22402028	0.004399534
SMARCB1	5.337457603	26.9767117	0.000145858
JMJD1C	5.3369857	25.40176619	0.000885809
CHD8	5.19621397	22.26962286	0.002793841
ARID2	5.143911862	24.02246146	0.012546401

MKL2	5.082943484	23.17477277	0.001963832
AHDC1	5.080688952	23.58229542	0.002396474
GLTSCR1L	5.073274101	22.83738881	0.002016229
AHNAK2	5.026858118	25.92462536	0.004524394
INTS7	4.966455679	26.35999729	0.001618047
BRD4	4.882129646	26.44979781	0.00061691
DOT1L	4.832054076	21.73659718	0.003303696
CORO1B	4.785254598	23.70958358	0.016462218
RBM27	4.761219527	23.83793305	0.004613873
ZNF687	4.720001891	23.57752459	0.035223739
AHNAK	4.539055824	30.48068922	0.001332903
ZEB2	4.531665772	23.42590429	0.003178125
PHF12	4.375237851	23.08752755	0.005331957
ZNF598	4.318079681	24.51174785	0.019683829
FAM208A	4.236217035	21.7024164	0.019706037
ACTL6A	4.186723965	22.60212398	0.00633743
TUBGCP3	4.160818159	22.4268859	0.011546017
QSER1	4.139291617	21.84577657	0.007861196
SAP130	4.130151786	23.82402441	0.012549799
KIF4A	4.109387815	23.32930262	0.007691696
DIDO1	3.99310038	25.80283316	0.008039221
THRAP3	3.973747636	22.33197298	0.027402073
TAF1	3.960651713	22.88532021	0.007878486
PRPF3	3.916406647	26.60129247	0.001314645
GATAD2B	3.914865824	24.80941492	0.00245245
TPR	3.830446392	25.0864508	0.00785187
BRD9	3.814218233	22.31606704	0.010526693
ZNF609	3.812250289	22.10592514	0.007755276
AIFM1	3.710893919	22.7414237	0.054863902
CHD7	3.692224242	24.74406556	0.037098846
RNF40	3.687432382	21.44374636	0.009651168
FLNA	3.673926004	28.87164766	0.000782011
PUF60	3.664763372	22.77100006	0.021680233
PBRM1	3.61492069	25.69060359	0.001695788
POLR2B	3.532349358	21.8563503	0.048461927
NUP85	3.492193962	24.34134532	0.102091216
EMSY	3.483595605	22.23927004	0.025235004
MED13	3.404727295	21.21339438	0.076851761
SART1	3.39818158	23.9545745	0.0189997
NACC1	3.37560273	21.49005767	0.023311211
SF1	3.302109547	24.70301317	0.031559076
YLPM1	3.29813484	23.49359802	0.01880843
MPG	3.296825147	22.95991474	0.034829653
SCML2	3.292396729	24.41484015	0.047355229
MEF2D	3.267319943	21.63418885	0.024442932
MEF2C	3.267319943	21.63418885	0.024442932
MEF2A	3.267319943	21.63418885	0.024442932

ZZZ3	3.267213369	20.84264769	0.043625499
XRN1	3.266469984	22.660742	0.019283086
NUP50	3.258912438	21.89780398	0.023474502
WDR5	3.246558682	22.20236788	0.019158186
TCF20	3.228524281	21.88208217	0.061941816
CCT8	3.223446185	22.00885872	0.035925227
DDX28	3.1597394	22.76035189	0.085539894
BRD7	3.159508316	22.42567202	0.017893476
SAP30BP	3.144408581	21.26155988	0.06310725
BPTF	3.114388828	22.37054731	0.061176444
TFAP2C	3.097870704	21.22120316	0.137660454
MAFG	3.096018618	22.17081472	0.059065716
MAFK	3.096018618	22.17081472	0.059065716
GTF2F2	3.093080386	21.55943769	0.056026411
RUVBL1	3.088876971	24.84916953	0.008509945
SNRPA1	3.064787153	24.53291574	0.031675225
INTS1	3.053387918	23.37463923	0.117356294
HM13	3.040168819	22.95836188	0.151172776
CHERP	2.932605465	22.81605855	0.024129358
RBBP6	2.920943576	22.54712629	0.069979333
RBM17	2.908821187	22.06411691	0.162580011
CEBPB	2.890272551	21.63619023	0.138323597
ZMYND8	2.882264329	22.89561748	0.058134948
ZNF318	2.873299645	25.64147271	0.007605361
MTA1	2.871006144	24.94830158	0.036342863
BUB3	2.861397603	22.60711356	0.058428715
PAXIP1	2.830479078	22.70267903	0.042540097
MAP4	2.829010752	25.23950507	0.006507224
HCFC1	2.827024388	28.69499606	0.000385298
TAF6	2.81426719	24.04286008	0.031607113
CTTN	2.808453609	22.53628242	0.039579418
BBX	2.802027978	21.55926486	0.049884937
UBAP2	2.790057983	21.46378682	0.029611774
ACIN1	2.781711322	21.13024816	0.097337154
CDK12	2.768560676	21.20617863	0.123793448
PNN	2.735051204	26.59577727	0.116323886
NUDT21	2.722339142	21.12648753	0.087118997
ZHX3	2.707215685	22.02445796	0.129113452
NCOA5	2.690135819	21.61655543	0.097628085
SEPT7	2.687853951	21.48527969	0.072111286
MRPL47	2.68170126	22.24076231	0.249760635
BCLAF1	2.676628388	22.28433452	0.081402056
YTHDF3	2.66986485	22.38622989	0.238218645
TPX2	2.656673289	26.41400644	0.006990048
PHF3	2.639076288	22.10096169	0.042160104
EPC1	2.638451654	20.73965518	0.037454478
DDX24	2.629149464	22.69339509	0.128768176

UBAP2L	2.622780146	23.32677322	0.037701665
SF3B2	2.617339151	26.33196295	0.023195725
KIAA1671	2.610845277	20.79060661	0.052915686
AKAP8	2.596578811	18.77305893	0.086716031
C19orf43	2.59107507	22.63690804	0.036106636
ZNF148	2.572592724	25.70494101	0.044047011
PLRG1	2.563697734	24.20883167	0.167340532
TCOF1	2.523196584	24.59241424	0.069619693
CHAMP1	2.511013752	22.56374436	0.069271616
FOXK2	2.477968591	22.29629277	0.071893136
CISD1	2.466295591	23.32605843	0.206689872
ZW10	2.424698534	23.8365383	0.113031952
RBM10	2.411995643	23.036651	0.203507904
NDUFA4	2.391597392	27.73434707	0.06447628
KHDRBS3	2.357193487	20.00075904	0.059578033
ELMSAN1	2.351561854	22.66399815	0.259528136
ZNF592	2.333189811	22.55101912	0.116049665
NUMA1	2.331978737	26.52510541	0.069671513
PPP1R10	2.294807388	20.67159308	0.156617197
SNW1	2.269079071	22.40483484	0.239791164
TIMM23	2.264988342	22.0942575	0.104665646
FNDC3B	2.262996549	23.82417194	0.245414699
PDLIM5	2.258557015	21.02154954	0.05324766
NDUFB5	2.229336545	24.0416818	0.292152374
MRPS27	2.217311992	20.11631906	0.069256237
MKI67	2.182412796	26.53311296	0.01901523
SNRNP70	2.172470764	23.81112574	0.220728876
EIF3C	2.17202403	20.62879865	0.111507207
EIF3CL	2.17202403	20.62879865	0.111507207
SACM1L	2.164810324	23.69161595	0.063295988
CPSF7	2.161668239	24.24147925	0.075967774
MRPL15	2.129210464	23.35609931	0.067124345
FIP1L1	2.093306597	20.47365274	0.165453658
NOL9	2.090733823	22.59163806	0.247262989
NELFA	2.090455699	22.44446546	0.132038284
SLTM	2.085371361	22.78301813	0.128042314
ACACA	2.08043053	27.88066293	0.003509585
PCF11	2.075972546	23.50435089	0.070571382
APMAP	2.067079571	25.54660138	0.032770846
SF3A2	2.049074347	23.40277595	0.317568993
SUGP2	2.041002957	23.89834866	0.098591511
BCS1L	2.033850768	22.72140533	0.221264193
CHD4	2.02334679	29.48405348	0.010796735
HNRNPDL	2.00735208	21.07138078	0.133691836

Table 6.2. Mass spectrometry analysis of SS18 interactome raw data. Analyzed with MaxQuant (Max Planck Institute of Biochemistry) software and Cassiopeia LFQ software. LogFC>2 as cut off.

Gene names	LogFC Values	Average Expression	P.Value
CPSF6	13.21394032	25.58831678	2.35E-08
SMARCC1	11.76919089	26.21832825	3.17E-08
ARID1B	10.07906716	27.16575959	4.48E-07
NUDT21	10.07073482	24.80068536	1.09E-06
TCF20	9.984539066	25.26008956	4.53E-07
SMARCD1	9.849688508	27.11733735	8.69E-08
SMARCA2	9.765944142	26.72461447	1.31E-07
RBM33	9.670921763	24.84651694	1.21E-06
YLPM1	9.618810645	26.65393592	1.50E-07
FIP1L1	9.513869559	24.18393422	1.27E-06
SMARCD2	9.399524953	26.9837352	2.81E-05
CSTF2	9.346853276	24.91429326	1.16E-06
AKAP8	9.210102265	22.07982066	1.03E-05
ARID1A	9.112782048	28.80194193	2.03E-06
RBM17	8.908031241	25.06372194	7.80E-07
SMARCC2	8.765667338	25.70687106	6.89E-06
GLTSCR1	8.761748159	24.18026401	3.03E-06
NCOA5	8.589019271	24.56599716	3.18E-06
CPSF1	8.571877682	24.52417872	1.27E-05
YEATS2	8.382735479	24.14779383	7.43E-06
GIGYF2	8.368381074	25.07009095	8.10E-06
SF1	8.361390399	27.2326536	3.10E-06
RBM25	8.241612016	24.76274904	2.54E-06
KMT2C	8.150366726	23.45912542	2.66E-05
RBBP6	8.120763834	25.14703642	3.96E-06
BCL7C	8.067537638	23.99552654	3.25E-05
RBM27	8.065863096	25.49025483	6.50E-06
DDX46	8.033107713	26.40388426	0.000106103
SF3A1	8.016637486	26.60723703	2.18E-06
DPF2	7.992977845	26.36162517	1.97E-05
CDK12	7.976003427	23.8099	0.000158667
ZNF609	7.923363564	24.16148178	1.00E-05
SMARCE1	7.905180694	26.03903925	9.06E-05
CDKN2AIP	7.839537554	24.12871181	8.57E-06
RBM10	7.818780287	25.74004332	0.000118193
ARID2	7.773498974	25.33725501	6.57E-05
SMARCA4	7.76490737	29.89654123	3.13E-07
PABPC1	7.608975936	24.78528101	5.90E-06
PABPC3	7.608975936	24.78528101	5.90E-06
RBM12	7.549410272	23.91661368	2.46E-05
PRR12	7.529242614	23.26793796	3.65E-05

PRPF40A	7.471490236	24.34107326	2.23E-05
CHERP	7.448779937	25.07414579	4.72E-06
PCF11	7.430246224	26.18148773	2.08E-06
ZBTB33	7.422215422	24.61237074	1.86E-05
JUNB	7.367335018	24.44001159	1.58E-05
CPSF7	7.362639464	26.84196486	2.75E-06
KHDRBS3	7.344465441	22.49439502	6.18E-05
CD2AP	7.284680919	23.16302818	8.36E-05
RAI1	7.18656979	23.6252162	3.85E-05
PRRC2C	7.12397109	24.00654276	0.001140998
POLR2B	7.098244834	23.63929803	2.71E-05
SCAF8	7.086561065	24.22081889	3.31E-05
FAM208A	7.058486711	23.11355124	0.00016223
SUGP1	7.03801654	24.16836114	2.11E-05
SNRPA1	7.022879434	26.51196188	8.29E-06
ZZZ3	6.986570639	22.70232632	0.000354537
BUB3	6.93961155	24.64622053	3.03E-05
ELMSAN1	6.904121662	24.94027806	0.001403579
PSPC1	6.850648396	23.19512959	0.000118025
JMJD1C	6.845296255	26.15592147	2.34E-05
MED12	6.799523052	24.05586425	5.97E-05
SF3B2	6.792761066	28.41967391	1.16E-06
UBAP2	6.781475262	23.45949546	7.29E-05
TFAP2C	6.780471223	23.06250342	0.000103134
PHF3	6.779247373	24.17104723	3.45E-05
WDR33	6.778322503	27.84926898	8.61E-06
QSER1	6.758054089	23.15515781	0.000206671
SNRNP70	6.73379595	26.09178833	0.000152512
PRRC2A	6.726656609	23.44602083	0.000114447
NKRF	6.725095395	24.60857081	0.000192964
CPSF4	6.672987444	24.44696714	2.57E-05
HNRNPDL	6.63214264	23.38377606	8.13E-05
SF3A2	6.59623998	25.67635876	0.00110925
BCOR	6.593052801	23.01704083	0.000151928
RBM26	6.500012459	24.20640588	4.92E-05
BCLAF1	6.475644869	24.18384276	0.000190545
WDR5	6.347102667	23.75263988	0.000402693
DIDO1	6.343264862	26.9779154	0.000103441
SNRPA	6.336388113	23.54437769	7.28E-05
FUBP1	6.304048779	27.6974652	1.14E-06
SRRM2	6.302861237	22.39633643	0.000240393
SNW1	6.288272397	24.41443151	0.00019826
GLTSCR1L	6.287676819	23.44459017	0.000123818
CCAR1	6.286751733	25.48621009	1.16E-05
ZFR	6.252137412	28.02280986	9.83E-07
CHD8	6.250396857	22.7967143	0.000216329
SMARCD3	6.243775395	26.10329845	0.003747363

ACIN1	6.225726765	22.85225588	0.000185915
RBM6	6.209561565	23.4399257	0.00167897
MRE11A	6.205964621	22.20295594	0.0002954
PUF60	6.180223892	24.02873032	6.47E-05
THRAP3	6.177285229	23.43374177	0.000364441
MED14	6.106182678	23.15959881	0.000261407
NCOA6	6.104216644	24.3172062	8.12E-05
LENG8	6.102818573	22.81077057	0.000171658
SS18	6.09312189	25.12716339	2.88E-05
RPRD2	6.056210434	27.1370863	4.49E-05
APOBEC3B	6.015638465	23.10213624	0.000347654
APOBEC3A	6.015638465	23.10213624	0.000347654
AHDC1	6.004870054	24.04438597	0.000107826
SAP130	6.004076534	24.76098678	0.000235505
BRD4	5.946883284	26.98217463	2.64E-05
FOSL1	5.931890233	23.0111786	0.000392006
ZC3H14	5.834883043	25.06482582	0.000503615
AHNAK2	5.82055108	26.32147184	0.000118427
ZNF326	5.787935826	23.19505207	0.000288802
AGO2	5.756139674	23.05492411	0.000336282
ZNF687	5.747085731	24.09106651	0.002745503
PPP1CB	5.687004122	23.84311765	0.000451935
DDX42	5.602058266	25.07231087	0.000848616
HNRNPF	5.592232143	23.67366815	0.00018255
SIN3B	5.587501951	23.98568901	0.000172464
NCOR2	5.586465999	23.11738232	0.000609985
DOT1L	5.58603527	22.11358778	0.000666685
RBM22	5.56247661	23.70978281	0.000374506
HMG20A	5.542405043	22.3945746	0.000589197
CCNK	5.533035093	23.87648961	0.000161575
KHSRP	5.525762105	29.96237943	4.93E-07
SNRPB2	5.504177615	23.44954691	0.000240468
TPR	5.444372723	25.89341396	4.24E-05
PAXIP1	5.439777409	24.0073282	0.000768775
SAP30BP	5.380272152	22.37949167	0.000817635
CREBBP	5.365132678	22.50337922	0.004107394
MEF2D	5.36448038	22.68276907	0.000879405
MEF2C	5.36448038	22.68276907	0.000879405
MEF2A	5.36448038	22.68276907	0.000879405
MKL2	5.362114019	23.31435804	0.000550401
ZNF638	5.359225639	28.11342565	4.49E-05
SDHB	5.331998825	22.14954846	0.000755222
CORO1B	5.330105622	23.98200909	0.001520722
SAFB	5.32312484	23.27665504	0.000681057
SMARCB1	5.320224884	26.96809535	4.45E-05
ATXN2	5.271555303	23.59686344	0.001115218
POLR2A	5.252173976	24.25380176	0.001611585

PBRM1	5.251227055	26.50875678	0.000360077
MED13	5.238552071	22.13030676	0.001640442
CBLL1	5.210382321	22.93119279	0.000950662
SLTM	5.196098469	24.33838168	0.000515211
WBP11	5.180206071	23.30059793	0.00055781
RNF40	5.14801737	22.17403885	0.002117119
TNRC6B	5.140308263	21.9983663	0.001878547
NACC1	5.135160083	22.36983635	0.001023423
GATAD2B	5.130412952	25.41718849	9.68E-05
TUBGCP3	5.126563629	22.90975864	0.000791349
HAUS6	5.11815772	22.54881962	0.001116044
UBAP2L	5.111342054	24.57105418	0.00038452
PLRG1	5.089355785	25.47166069	0.00379255
PLK1	5.088178998	23.57259543	0.005537874
FAM98A	5.076818106	23.30553531	0.000514011
AHNAK	5.058477877	30.74040025	4.22E-05
SAFB2	5.052673493	25.55277038	0.002380421
GTF2F2	5.029769651	22.52778232	0.000879687
ZHX3	5.02712958	23.18441491	0.002247781
C19orf43	4.996470919	23.83960597	0.000378092
RBM14	4.988908749	28.90191089	3.95E-06
MITF;TFEB	4.984107472	21.72888943	0.001852427
PPP1R10	4.979250873	22.01381482	0.002863722
KMT2D	4.935539437	25.6594255	0.000468238
MED17	4.918815186	23.45458929	0.001010214
KIAA1671	4.915165627	21.94276679	0.001921277
ILF3	4.898247606	28.80194589	0.000272494
TAF1	4.897593142	23.35379092	0.000908461
SFPQ	4.890401052	30.59701188	6.64E-06
YTHDF3	4.855749845	23.47917238	0.017810109
XRN1	4.855521782	23.4552679	0.000669014
SART1	4.836581955	24.67377469	0.000248489
LUC7L3	4.831578876	23.32654044	0.000730164
EWSR1	4.807409813	24.40578748	0.000323668
LARP1	4.799914204	22.43858108	0.001401515
DROSHA	4.790539155	21.82744899	0.001887753
CHD7	4.775707327	25.2858071	0.003570421
CRNKL1	4.767881697	22.83502681	0.002300451
BPTF	4.762083714	23.19439475	0.001187232
WAC	4.745956389	22.78895218	0.001664044
DHX15	4.741859427	26.80360113	0.000389258
UPF1	4.741815536	22.22013286	0.001952258
STAU1	4.741423735	22.8894607	0.001343313
MED1	4.740372296	22.73678993	0.002492235
CTTN	4.665154711	23.46463297	0.001019979
SON	4.650145048	25.97961536	0.000465644
CDK13	4.634472491	22.68076265	0.002149699

BRD9	4.630787254	22.72435155	0.003048047
TCERG1	4.620998772	27.73786791	3.75E-05
BRD7	4.613844166	23.15283995	0.005553171
EEF1D	4.609461671	22.59494794	0.001478192
BBX	4.577761194	22.44713147	0.002105935
PHF14	4.570685766	22.44794966	0.001827617
INTS7	4.570609672	26.16207429	0.000167907
TFIP11	4.565418041	23.11147259	0.002308813
FLNA	4.541283565	29.30532644	9.34E-06
U2AF2	4.533574106	23.98715522	0.000984602
ZNF592	4.527555291	23.64820186	0.00316699
RBM15	4.507142382	23.61153239	0.000949723
SF3B6	4.466204405	22.94358533	0.001678581
ZNF318	4.428689608	26.41916769	0.000228552
CCNT1	4.405302593	24.94852077	0.000431047
FXR2	4.384379112	23.25883249	0.00213055
CIZ1	4.383830304	28.41555992	1.75E-05
XAB2	4.382603211	25.97897705	0.000643304
NUP50	4.368277412	22.45248646	0.003084607
SCAF11	4.35661182	23.17827367	0.00224826
SUGP2	4.348104618	25.05189949	0.000611938
ZNF598	4.320120698	24.51276836	0.006737107
PHRF1	4.296434551	23.98033709	0.003953221
CLK3	4.252551414	22.26431349	0.004302392
HELZ	4.20432118	22.68933585	0.00498124
RUNX1	4.203865239	22.96843339	0.008729721
U2SURP	4.173173204	29.92901845	1.58E-05
EMSY	4.156598171	22.57577132	0.011276997
ZEB2	4.155896831	23.23801982	0.002547278
CBX8	4.144760084	22.71112288	0.002958078
HNRNPUL1	4.141326199	27.03340025	0.00010866
NUFIP2	4.123938923	24.81543591	0.001350314
ASH2L	4.118971782	21.19650088	0.028732017
CDC40	4.100179246	22.49172922	0.022895824
FUBP3	4.097918788	28.65079161	1.62E-05
ACTL6A	4.092850551	22.55518727	0.003831617
ZMYND8	4.08481618	23.49689341	0.003030598
DDX23	4.081171376	21.99647053	0.005944177
SLU7	4.070973136	21.82197195	0.006224479
CAPRIN1	4.022475097	23.2467566	0.002454149
CPSF3	4.000801878	22.71351626	0.013404361
KIF4A	3.994871889	23.27204466	0.004060057
PDLIM5	3.990957218	21.88774964	0.006724264
HNRNPR	3.987405698	27.06398334	0.000126455
RAVER1	3.983820097	22.4777155	0.011584106
GOSR2	3.983802357	22.98895954	0.003067379
HNRNPK	3.954554367	30.30079565	1.30E-05

SATB2	3.940008177	22.13462887	0.006484748
SATB1	3.940008177	22.13462887	0.006484748
FASN	3.929538014	22.77064614	0.004285649
SNRPC	3.918417349	23.05650176	0.006582893
DAZAP1	3.916565493	25.35339244	0.001510739
NONO	3.912914478	30.93120383	3.16E-06
DDX24	3.911902018	23.33477137	0.015065518
SEPT7	3.893832521	22.08826897	0.008166839
KIAA0907	3.891496971	23.43561201	0.009312246
HNRNPH3	3.87169569	23.49273094	0.005030826
SF3A3	3.864522665	21.59907382	0.008479375
ZMYM3	3.818155464	22.34140238	0.005543231
PABPC4	3.808288343	22.70563942	0.004759016
MLLT10	3.804712395	21.94274907	0.014368332
ZNF217	3.79067251	22.15845692	0.010465027
MED15	3.781263698	21.95025799	0.010365323
PNN	3.769325727	27.11291453	0.011025577
KDM6A	3.752510086	24.86502616	0.007580715
HAUS8	3.74195647	22.02151219	0.012330925
SCAF4	3.724892053	25.1578366	0.002059686
HNRNPD	3.712267105	25.45859109	0.002817951
ZC3H11A	3.709145188	25.45667593	0.000906954
ILF2	3.692810828	26.70289526	0.00165102
CEBPB	3.691883614	22.03699576	0.014064982
CPSF2	3.676415227	23.00893403	0.011125724
CSTF3	3.66584918	26.54241949	0.000812776
MTA1	3.655630117	25.34061356	0.004017987
ZFC3H1	3.652750849	21.62569416	0.017500127
EP300	3.64085961	23.1465971	0.013246356
TAF6	3.635569485	24.45351123	0.00706838
RUVBL1	3.625863186	25.11766263	0.001288871
EPC1	3.607147334	21.22400302	0.014450824
MATR3	3.578631654	29.58118897	0.000556336
RANGAP1	3.57670287	22.36712258	0.025896885
HNRNPA2B1	3.575137767	28.90290443	9.76E-05
SNRPD3	3.550292744	26.89161078	0.00231705
ADAR	3.549352035	25.85321792	0.0053273
SYNCRIP	3.520349072	23.41855383	0.007917235
NCOR1	3.506951881	22.02779496	0.01579705
CNOT1	3.505282391	25.54874221	0.001220802
SIN3A	3.49505321	26.46549247	0.003664808
FOXK2	3.471693849	22.7931554	0.013318075
FXR1	3.408967651	22.45693365	0.024107499
NUMA1	3.358726943	27.03847951	0.005729564
IGF2BP2	3.357537381	22.75848901	0.016442515
HNRNPC	3.350309051	30.07504619	0.000350055
STAT3	3.341639589	22.95526691	0.026839593

SPEN	3.33638262	25.06628455	0.003048862
ZNF148	3.331607616	26.08444845	0.005560773
SNRPN	3.321109011	27.08237691	0.000758045
SNRPB	3.321109011	27.08237691	0.000758045
FAM120A	3.288724516	26.03196746	0.001840057
RANBP2	3.280536311	29.34052712	5.81E-05
RGPD4	3.245275674	21.66406401	0.020529344
DNTTIP1	3.244941587	24.99799507	0.007378994
MED19	3.242558865	21.60337572	0.055075176
PHF12	3.208248084	22.50403267	0.022343818
HNRNPUL2	3.189451653	27.006003	0.005049348
RBPJ	3.186968286	22.82356751	0.012406784
ASXL2	3.178540587	20.70513963	0.056275027
AQR	3.175750687	22.77466746	0.023286068
BAP1	3.174504705	21.74637409	0.029678657
NOL9	3.160735799	23.12663905	0.049474428
CDK8	3.158180582	21.77640015	0.038856811
CDK19	3.158180582	21.77640015	0.038856811
DKC1	3.134415437	22.40327327	0.046685006
XRN2	3.121461581	27.74567078	0.003541968
SF3B3	3.117769947	23.62210057	0.058365642
MRPS27	3.117404629	20.56636538	0.039897328
DYNLL1	3.098939353	23.14446617	0.039039045
DYNLL2	3.098939353	23.14446617	0.039039045
PQBP1	3.096527853	22.12475071	0.019304553
DMAP1	3.062310154	23.00183518	0.02899079
SUPT16H	3.047384867	21.80129737	0.072228001
SKIV2L2	3.014922713	23.96874161	0.043806684
TPX2	3.01031009	26.59082484	0.002430276
FUS	2.997309839	28.59106834	0.000437202
PAXBP1	2.983863322	21.94004359	0.032520211
SRCAP	2.958095289	24.7104175	0.008548898
OGT	2.956330413	21.12572016	0.051072122
HCFC1	2.905570506	28.73426912	0.000605718
PRPF3	2.880148361	26.08316332	0.003670539
RBM12B	2.854148755	22.23543386	0.048934185
SRSF3	2.848418356	22.10266908	0.034213071
RALY	2.823582573	29.84703619	0.001991521
CHD4	2.81646617	29.88061317	0.000568586
TCOF1	2.811449279	24.73654059	0.024795839
JUN	2.8105608	22.32239232	0.030532984
JUND	2.8105608	22.32239232	0.030532984
DHX8	2.785545654	22.2522764	0.041930833
YBX1	2.737537992	27.39824247	0.004366461
PML	2.69255743	23.93285887	0.063290798
DDX6	2.666145895	25.91334434	0.005622654
LUC7L2	2.612972694	24.66832074	0.015289395

TAF2	2.612189355	21.72782568	0.073315868
PCBP1	2.58856272	28.31297375	0.000932665
BOP1	2.542896749	22.01524603	0.085724211
GATAD2A	2.533353514	27.19713383	0.019032006
IQCG	2.524466703	23.18770611	0.085307564
RNMT	2.48966993	23.46644505	0.119363458
HNRNPA1	2.47589974	29.24753388	0.000820996
HNRNPA1L2	2.47589974	29.24753388	0.000820996
CCT8	2.470483712	21.63237748	0.076305331
ADNP	2.448827987	28.89353144	0.0006059
FNDC3B	2.448223702	23.91678552	0.131788504
TRRAP	2.428581822	25.77634714	0.02360289
ZNF207	2.421477396	26.91528189	0.00431349
HDAC2	2.369918948	24.7918871	0.025884174
IK	2.362115313	23.69911514	0.049255206
KAT2B	2.353196578	21.65449984	0.106930924
KAT2A	2.353196578	21.65449984	0.106930924
KPNA3	2.351271049	22.46068611	0.057428442
CHAMP1	2.334978988	22.47572698	0.058228211
MRPL47	2.326638488	22.06323092	0.22179095
NUP85	2.310325138	23.75041091	0.179232623
INTS1	2.283489653	22.9896901	0.173176397
TOP2A	2.28061688	24.14032139	0.120260668
SNRNP200	2.271356198	27.74168419	0.012185201
HLTF	2.257905877	21.90778806	0.089211238
DHX30	2.249929123	22.1135619	0.122712301
TAF9B	2.249911326	21.29278775	0.11357728
TAF9	2.249911326	21.29278775	0.11357728
GAR1	2.243776534	21.79800171	0.101873331
CELF1	2.241225868	25.43926717	0.020270597
EIF2AK2	2.219874371	22.09552247	0.080185489
BCS1L	2.213378087	22.81116899	0.144907187
SYNM	2.188887419	22.2671399	0.097144553
MDC1	2.182502652	24.45018585	0.037574407
SYMPK	2.162349729	28.0046097	0.005822689
CSNK2A2	2.160937897	22.50994516	0.075262052
RBMX	2.150198335	27.14103427	0.025229005
FANCI	2.149576158	25.35384496	0.036034432
SCML2	2.149018021	23.8431508	0.127794235
DDX5	2.129360057	29.13636282	0.003077663
EIF4A3	2.119236883	27.43682218	0.011992524
KIF23	2.109199744	24.67842515	0.033594491
NOL10	2.109096133	20.22615823	0.133208044
KHDRBS1	2.102016748	27.41605205	0.013248385
HDAC1	2.089314284	26.86289741	0.019193931
CENPB	2.036724168	21.61722804	0.123815259
MED31	2.036072989	21.81592987	0.160597074

MKI67	2.020942543	26.45237783	0.02426002
CWC22	2.019661252	21.56893702	0.114569305
SMCHD1	2.014658153	22.16484357	0.137129022
MRPS34	2.010555822	22.12688553	0.129403702

Table 6.3. Mass spectrometry analysis of SS18L1 interactome raw data. Analyzed with MaxQuant (Max Planck Institute of Biochemistry) software and Cassiopeia LFQ software. LogFC>2 as cut off.

Gene names	LogFC Values	Average Expression	P.Value
SMARCC1	12.00168803	26.33457682	7.66E-07
CPSF6	10.88168012	24.42218668	1.11E-05
ARID1B	10.43628661	27.34436931	2.79E-05
SMARCA2	10.36134017	27.02231249	6.85E-07
SMARCD1	10.0357976	27.2103919	7.05E-07
SMARCC2	9.581117802	26.11459629	8.57E-05
SMARCD2	9.509373383	27.03865942	0.000865878
ARID1A	9.39436865	28.94273523	0.000139185
GLTSCR1	9.38126207	24.49002097	1.06E-05
TCF20	8.992061455	24.76385076	3.25E-05
NUDT21	8.487997437	24.00931667	0.000104464
GIGYF2	8.477481986	25.1246414	5.17E-05
BCL7C	8.42490294	24.1742092	0.000193943
SS18L1	8.302607512	24.38937036	2.53E-05
YLPM1	8.179892747	25.93447697	2.89E-05
SMARCA4	8.059467488	30.04382129	9.75E-06
ARID2	7.959636635	25.43032384	0.000754738
ZBTB33	7.950049452	24.87628775	6.74E-05
FIP1L1	7.881785466	23.36789217	6.00E-05
SMARCE1	7.879725368	26.02631159	0.001886973
RBM17	7.682373238	24.45089294	3.09E-05
DPF2	7.577779282	26.15402589	0.000307975
CSTF2	7.325779779	23.90375651	0.000113043
HAUS6	6.96070672	23.47009412	0.000168837
YEATS2	6.921639166	23.41724568	0.00018804
JUNB	6.881939025	24.19731359	0.000192302
AHDC1	6.725274064	24.40458797	6.66E-05
GLTSCR1L	6.69644239	23.64897296	0.000114388
RBM25	6.679217367	23.98155172	9.10E-05
RBM10	6.673442847	25.1673746	0.00500122
RBM33	6.67197404	23.34704308	0.000344554
AKAP8	6.647860746	20.7986999	0.000337997
RBM27	6.637727646	24.77618711	0.000457027
DDX46	6.520852312	25.64775656	0.00644022
JMJD1C	6.433769296	25.95015799	0.000254948
KMT2C	6.426449654	22.59716689	0.000383489
SF3A1	6.36622132	25.78202895	0.000308298

EIF3C	6.257615379	22.67159432	0.000412807
EIF3CL	6.257615379	22.67159432	0.000412807
SF1	6.235433033	26.16967491	0.000768949
MRE11A	6.142678178	22.17131272	0.000623835
DOT1L	6.111038254	22.37608927	0.000551546
PCF11	6.07513619	25.50393271	5.98E-05
NCOA5	5.964459539	23.25371729	0.000316246
TFAP2C	5.922306497	22.63342105	0.000962035
SMARCD3	5.8891498	25.92598565	0.027933573
SIN3B	5.88126666	24.13257136	0.000158635
CPSF7	5.858180787	26.08973552	0.000282585
FAM208A	5.824584858	22.49660031	0.002653008
SNRPA1	5.820370753	25.91070754	0.000577916
BRD4	5.802650269	26.91005812	0.00015992
LENG8	5.700167571	22.60944507	0.000506253
CPSF1	5.563699279	23.02008952	0.002054071
RBBP6	5.535945628	23.85462732	0.000734284
POLR2B	5.496588993	22.83847011	0.000516745
RBM12	5.462475976	22.87314653	0.000784978
CEBPB	5.382082802	22.88209536	0.003084344
SF3B2	5.324857221	27.68572198	0.000218817
CHERP	5.321379733	24.01044569	0.000391895
HAUS8	5.291873602	22.79647076	0.000635817
SMARCB1	5.200529516	26.90824766	0.000211592
ACTL6A	5.193895713	23.10570985	0.000911642
GTF2F2	5.183056391	22.60442569	0.001565515
PABPC1	5.124021059	23.54280357	0.000701004
PABPC3	5.124021059	23.54280357	0.000701004
HNRNPF	5.123965414	23.43953479	0.000642824
RAI1	5.07384571	22.56885416	0.002228986
PRPF40A	5.031714986	23.12118563	0.002251616
ZNF687	4.970583939	23.70281562	0.023366414
INTS7	4.965413769	26.35947634	0.000386262
PUF60	4.935695094	23.40646592	0.000724626
KHDRBS3	4.892179528	21.26825206	0.002935177
NCOR2	4.872909178	22.76060391	0.00128624
MEF2D	4.853671771	22.42736476	0.002244825
MEF2C	4.853671771	22.42736476	0.002244825
MEF2A	4.853671771	22.42736476	0.002244825
UBAP2	4.823485798	22.48050073	0.002031678
SNRNP70	4.803983555	25.12688213	0.013424241
CDK12	4.735904985	22.18985078	0.015223912
ZNF326	4.716343828	22.65925607	0.001467567
BUB3	4.70622307	23.52952629	0.003113186
CHD8	4.698011469	22.02052161	0.001516266
PRRC2A	4.681808415	22.42359674	0.002479883
THRAP3	4.666145137	22.67817173	0.007048645

GATAD2B	4.6661426	25.18505331	0.000327408
QSER1	4.633722218	22.09299187	0.003623307
MKL2	4.627382666	22.94699236	0.001502191
HNRNPDL	4.586646902	22.36102819	0.001512382
SUGP1	4.554789684	22.92674771	0.001490355
SNW1	4.553427752	23.54700918	0.008925951
PAXIP1	4.540221747	23.55755037	0.002849743
BRD9	4.538526975	22.67822141	0.002067343
SAP130	4.528755215	24.02332612	0.005466099
WDR33	4.494513257	26.70736435	0.002217208
ZNF598	4.488082435	24.59674922	0.0166923
ZC3H14	4.463927277	24.37934794	0.013949029
BCLAF1	4.426443869	23.15924226	0.010038937
PSPC1	4.42416743	21.9818891	0.006517985
CPSF4	4.384836797	23.30289181	0.001495163
DDX42	4.365653034	24.45410825	0.017799525
ZNF609	4.31443103	22.35701551	0.001935663
SF3A2	4.299952187	24.52821487	0.052433521
ZFR	4.28871935	27.04110083	0.00018202
FAM98A	4.274173889	22.9042132	0.003295752
ZEB2	4.264316761	23.29222978	0.00195506
ZNF318	4.259047755	26.33434676	0.000522681
SRRM2	4.251462386	21.370637	0.004479725
RBM26	4.190110485	23.05145489	0.002201166
NACC1	4.171767436	21.88814002	0.007882633
BRD7	4.149860054	22.92084789	0.001939739
TPR	4.148681044	25.24556812	0.000877528
C19orf43	4.135928644	23.40933483	0.001646861
NCOA6	4.134324353	23.33226006	0.001536057
PRRC2C	4.116491201	22.50280281	0.068329927
RPRD2	4.115164161	26.16656316	0.005123529
KHSRP	4.114477073	29.25673691	3.83E-05
PPP1CB	4.109654009	23.05444259	0.009182299
ACIN1	4.061091034	21.76993801	0.011007554
AHNAK2	4.057465224	25.43992891	0.004738618
ZHX3	4.035369044	22.68853464	0.00430187
SAP30BP	4.011868494	21.69528984	0.005189868
NKRF	4.006599146	23.24932268	0.026482727
APOBEC3B	3.934896078	22.06176505	0.006100774
APOBEC3A	3.934896078	22.06176505	0.006100774
ASH2L	3.931100432	21.10256521	0.007595455
RNF40	3.925990523	21.56302543	0.005464751
SCAF8	3.917303997	22.63619036	0.003802732
ILF3	3.9071702	28.30640718	0.009167871
NOL9	3.887544274	23.49004329	0.039718165
CHD7	3.885920778	24.84091382	0.030231931
UBAP2L	3.849024176	23.93989524	0.003052864

CCAR1	3.840579134	24.26312379	0.002018639
CD2AP	3.832089353	21.4367324	0.00679011
SUGP2	3.799260459	24.77747741	0.002008361
CDKN2AIP	3.781183347	22.09953471	0.004455761
PHF12	3.771796453	22.78580686	0.004101697
BBX	3.699552303	22.00802703	0.005932609
XAB2	3.66925927	25.62230508	0.004520956
PBRM1	3.66715387	25.71672018	0.00092626
PLRG1	3.628868582	24.74141709	0.063003361
TAF1	3.619407453	22.71469808	0.004184046
CBLL1	3.558749147	22.1053762	0.018047538
TOP2A	3.544316321	24.77217111	0.026241208
AHNAK	3.537402565	29.97986259	0.003104664
SON	3.533932384	25.42150903	0.007043497
ZNF148	3.521339343	26.17931432	0.011422971
POLR2A	3.505738695	23.38058412	0.039944443
FLNA	3.504285847	28.78682758	7.69E-05
DIDO1	3.503805456	25.55818569	0.016714561
MED13	3.466440543	21.244251	0.032643779
MTA1	3.464030226	25.24481362	0.012530372
PLK1	3.444858383	22.75093512	0.092035615
MITF	3.420058611	20.946865	0.062466429
TFEB	3.420058611	20.946865	0.062466429
RUVBL1	3.40910118	25.00928163	0.001735068
NUP50	3.370716998	21.95370626	0.011076306
DHX15	3.34762305	26.10648294	0.015177715
EP300	3.321900344	22.98711747	0.025715422
PHF3	3.320901703	22.4418744	0.012411321
ZMYND8	3.302730367	23.1058505	0.011842223
FXR1	3.302280539	22.4035901	0.041038177
SART1	3.283419274	23.89719335	0.004655074
SFPQ	3.268620732	29.78612172	0.001750898
BPTF	3.268420592	22.44756319	0.011863541
CCNK	3.268076706	22.74401042	0.006449273
CORO1B	3.266260164	22.95008637	0.05208868
KIAA1671	3.26162511	21.11599653	0.010959447
TAF6	3.258364726	24.26490885	0.008048078
FOXK2	3.258156224	22.68638658	0.013193919
CNOT1	3.256441765	25.42432189	0.001195186
MATR3	3.243351989	29.41354913	0.002812774
PNN	3.225642607	26.84107297	0.069093065
PML	3.225125847	24.19914308	0.02924457
TOP2B	3.224382684	21.33611576	0.028148364
SF3B3	3.174517465	23.65047433	0.101810642
FUBP1	3.140796353	26.11583899	0.001425991
KIF4A	3.140149658	22.84468354	0.015829169
PHRF1	3.111648734	23.38794419	0.037006119

CAPRIN1	3.107713782	22.78937594	0.008765152
ZNF638	3.107519957	26.98757281	0.011803587
SF3B6	3.083928741	22.2524475	0.011050765
HNRNPK	3.05523498	29.85113595	0.00070087
U2SURP	2.984149521	29.3345066	0.001317298
LUC7L3	2.980122263	22.40081213	0.010114389
RBM6	2.92914054	21.79971518	0.123543473
FUBP3	2.918146739	28.06090559	0.00051489
SLTM	2.916841137	23.19875302	0.032640492
CIZ1	2.914430453	27.68085999	0.000722918
CHD4	2.901873457	29.92331681	0.00175485
ELMSAN1	2.894730579	22.93558252	0.174786806
IQCG	2.882718699	23.36683211	0.081347858
KMT2D	2.881118969	24.63221527	0.012624187
CPSF2	2.87234312	22.60689798	0.018967083
PHF14	2.87014057	21.59767706	0.019047536
HNRNPUL2	2.845057835	26.83380609	0.010309225
CRNKL1	2.823222313	21.86269712	0.020659645
CDC40	2.789264343	21.83627177	0.114733333
SCML2	2.777135149	24.15720936	0.27891045
SKIV2L2	2.774651107	23.84860581	0.099244506
HNRNPD	2.762387279	24.98365118	0.011373225
PRR12	2.758672088	20.88265269	0.018636867
FOSL1	2.73115123	21.4108091	0.128864953
RANBP2	2.728422338	29.06447013	0.000327809
NUP85	2.723966114	23.9572314	0.189331732
DAZAP1	2.708072081	24.74914574	0.021564246
SYNCRIP	2.687232577	23.00199558	0.034428247
SNRNP200	2.671463748	27.94173796	0.004688206
INTS1	2.670359397	23.18312497	0.158778288
ILF2	2.646365729	26.17967271	0.027780518
SF3A3	2.638400846	20.98601291	0.029043337
NONO	2.609172792	30.27933299	0.000170141
CISD1	2.589143053	23.38748216	0.167975506
HNRNPA2B1	2.587093025	28.40888206	0.002545325
TRIM25	2.536077608	23.53420235	0.080571025
CBX8	2.514023964	21.89575482	0.022277332
SUPT16H	2.503718196	21.52946403	0.117578267
BCS1L	2.496541319	22.9527506	0.136820662
SCAF4	2.494361027	24.54257108	0.022856182
KPNA3	2.493247215	22.5316742	0.026433865
SEPT7	2.491029302	21.38686736	0.132640312
CTTN	2.48544127	22.37477625	0.023411784
BAP1	2.474977665	21.39661057	0.041673002
EMSY	2.456740636	21.72584255	0.043897354
MED14	2.423735605	21.31837527	0.050055552
ADNP	2.413095843	28.87566537	0.000431266

HNRNPC	2.4112126	29.60549797	0.006341345
GATAD2A	2.406158848	27.1335365	0.042953801
MED17	2.377608605	22.183986	0.036487692
MED1	2.377562531	21.55538505	0.031644149
HCFC1	2.37374571	28.46835672	0.000694268
CCNT1	2.363310719	23.92752484	0.015431301
YBX1	2.36014957	27.20954826	0.019330288
AURKB	2.328982194	20.31629568	0.087878934
RBM14	2.319260792	27.56708691	0.002227152
ITPR3	2.313335834	24.64282717	0.079950231
CPSF3	2.310512542	21.86837159	0.077066909
CLK3	2.306267682	21.29117162	0.046586839
SNRPD3	2.27082823	26.25187852	0.051797072
LAS1L	2.249233186	24.01740188	0.098553868
ADAR	2.235699352	25.19639158	0.060788966
MYH9	2.225848916	23.24566387	0.18147625
RALY	2.223346354	29.54691808	0.010372349
ZNF217	2.218771064	21.3725062	0.075401295
SMU1	2.215118074	22.60367349	0.117267763
SDHB	2.212916932	20.59000751	0.245142404
SMCHD1	2.158901349	22.23696517	0.099081729
CC2D1A	2.155056416	23.4016376	0.158796723
SPEN	2.150535452	24.47336097	0.020781847
HNRNPR	2.145369245	26.14296511	0.021101277
BCOR	2.13946362	20.79024624	0.052910203
FAM120A	2.138919542	25.45706498	0.009119467
SENP3	2.13887247	25.27393713	0.04210945
IK	2.132747208	23.58443109	0.05931365
STAU1	2.128633746	21.5830657	0.058326067
ZC3H11A	2.126452828	24.66532975	0.036383595
TPX2	2.111905032	26.14162231	0.011953282
KMT2A	2.095949772	24.3269082	0.13794589
FANCI	2.093644701	25.32587923	0.038899031
TNRC6B	2.07053566	20.46348	0.097416412
SRCAP	2.052019238	24.25737947	0.032549028
SNRPN	2.041038439	26.44234162	0.017592137
SNRPB	2.041038439	26.44234162	0.017592137
U2AF2	2.028190798	22.73446357	0.061600774
EIF2AK2	2.02635163	21.9987611	0.045790345
NUP133	2.012778277	25.97763802	0.088529354
YTHDF3	2.002620126	22.05260752	0.362103116
TTK	2.000827334	22.31535925	0.173514522

Table 6.4. Mass spectrometry analysis of SS18L2 with SS18 as control raw data. Analyzed with MaxQuant (Max Planck Institute of Biochemistry) software and Cassiopeia LFQ software. LogFC>2 as cut off.

Gene names	LogFC Values	Average Expression	P.Value
SS18L2	8.346672511	24.69497001	3.10E-05
IBA57	4.750701063	22.04300088	0.004099677
HAUS8	3.753649497	25.76931518	0.005177547
HAUS6	3.711431039	26.963614	0.001144636
PNO1	3.677424385	22.67894289	0.00783193
ATP5B	3.528830736	22.7636457	0.019868295
MPG	3.357164668	22.92974498	0.019328206
EIF2S2	3.280579248	25.25769669	0.006876318
SH3BP4	3.122138507	22.80139879	0.017656411
PHGDH	3.106199729	20.94046969	0.100137923
APOO	3.050435489	24.80089831	0.080048194
PARL	3.020057443	24.98049119	0.055901598
UTP18	3.007742722	21.13209909	0.032238462
STK10	2.988045512	23.75906163	0.010899007
NDUFB5	2.871824293	23.72043792	0.018831212
PPIA	2.858883895	26.45351352	0.001769251
NPM1	2.856450213	28.36902854	0.001187488
CSRP2	2.766078002	25.41634078	0.008123998
MAGED2	2.707259544	22.41381628	0.161136796
SLC27A1	2.699671125	23.83942229	0.018097739
CCDC137	2.646681235	24.99348755	0.011468514
RAB38	2.64437447	23.31574433	0.029390381
HELLS	2.617400413	25.57720451	0.005635742
ZNF593	2.584545842	24.51438962	0.014582278
TMEM2	2.574292382	21.74953306	0.043269063
SRFBP1	2.534528226	26.31714658	0.004067925
ARL6IP5	2.49030668	25.34699056	0.077475604
TOMM22	2.48029979	22.19907281	0.042392128
RAB5A	2.4713033	22.28113089	0.052844151
CCT6A	2.467662692	22.49484107	0.059857839
TIMM23	2.454775397	21.99936398	0.059783397
SEC22B	2.446568164	24.6826985	0.04745365
FAR1	2.435991081	24.56952357	0.016998128
MAFF	2.421620523	26.68700389	0.004149393
SLC25A30	2.407808239	23.1133051	0.065999846
SLC25A14	2.407808239	23.1133051	0.065999846
DCAF13	2.386830968	23.64657025	0.067080012
NSUN5	2.366380021	26.92013932	0.005061597
NR2F6	2.357539949	22.16308882	0.099005213
HM13	2.337450235	23.30972117	0.116408421
TFB2M	2.319028853	27.96805628	0.007340458
C19orf52	2.317287138	26.09330518	0.006774996

SURF4	2.311355747	21.32151358	0.076748456
PTRH1	2.295362747	24.55816771	0.025625866
LGALS1	2.27512717	23.83237161	0.030482092
IDH3B	2.271026192	24.45329873	0.026283387
AAAS	2.264388905	21.97893876	0.160889988
SLC25A4	2.256991205	27.78234718	0.002968293
MSH2	2.254048157	27.90665627	0.012190975
SLC25A22	2.222781593	29.74514763	0.001029708
SLC25A19	2.193024419	28.1519007	0.009521825
BTAF1	2.138851751	22.70272184	0.178635525
DIMT1	2.1331682	27.416703	0.0041079
KLF16	2.132528599	24.80932691	0.035451453
BAZ2B	2.132098855	28.03242285	0.00924293
TAP2	2.074401976	24.87851731	0.050100397
ABCB10	2.074401976	24.87851731	0.050100397
ZNF787	2.068681604	24.21806058	0.05768277
PSIP1	2.066125395	21.89604759	0.093133043
NHP2L1	2.062467515	26.39531644	0.019040469
MAFG	2.060681059	22.68848349	0.125349137
MAFK	2.060681059	22.68848349	0.125349137
RPL17	2.058975342	33.31124442	0.000655774
THAP11	2.040656931	21.36526153	0.261814725
GPNMB	2.03815173	28.88158298	0.004890507
HADHB	2.032945213	27.73737201	0.009111562
NDUFA12	2.026356786	24.92650956	0.027808967
RRP8	2.024199048	26.27202062	0.011708266
KDM2A	2.003331075	22.08942667	0.147568952
OXA1L	2.003067268	27.34452275	0.010562641

Table 6.5. Mass spectrometry analysis of SS18L2 with SS18L1 as control raw data. Analyzed with MaxQuant (Max Planck Institute of Biochemistry) software and Cassiopeia LFQ software. LogFC>2 as cut off.

Gene names	LogFC Values	Average Expression	P.Value
SS18L2	7.916708725	24.90995191	7.49E-06
NDUFB5	5.349309138	22.4816955	0.000850765
RAB27A	5.278273917	22.62468268	0.003096692
SH3BP4	4.69598069	22.01447769	0.001788273
LGALS1	4.684134918	22.62786773	0.005329737
TAP2	4.426431912	23.70250235	0.006099257
ABCB10	4.426431912	23.70250235	0.006099257
EEF1D	4.29344621	24.09944142	0.056224592
MRPL15	4.265777654	22.28781572	0.005906832
APOO	4.228935316	24.21164839	0.067764429
DDX28	4.036193772	22.3221247	0.00305371
PARL	3.956293213	24.51237331	0.011237294
MPG	3.953785203	22.63143471	0.009992166
STT3A	3.938837886	24.46860327	0.002610136
SDHB	3.918362021	23.65564699	0.037283127
TUBGCP3	3.74553748	22.63452624	0.008776151
SLC27A1	3.703808949	23.33735338	0.004325201
MAGT1	3.636141093	21.67462127	0.083343154
RAB38	3.370298576	22.95278228	0.009091391
ARL6IP6	3.350599078	21.95323179	0.017217039
TRMT10C	3.34557988	23.70429172	0.030680079
WDR5	3.266270614	22.19251192	0.009514495
NDUFB6	3.264624397	22.7286402	0.014404416
EIF5A2	3.209791595	25.05404367	0.002615586
EIF5AL1	3.209791595	25.05404367	0.002615586
EIF5A	3.209791595	25.05404367	0.002615586
AIFM1	3.194602693	22.99956931	0.018428849
TMCO1	2.985044438	23.66743827	0.015943209
HELLS	2.975014913	25.39839726	0.002663498
MAFG	2.942771738	22.24743816	0.04684545
MAFK	2.942771738	22.24743816	0.04684545
NDUFB4	2.935681498	24.17901625	0.023050641
IBA57	2.870583496	22.98305966	0.032934784
KLF16	2.827916848	24.46163278	0.007704443
CSRP2	2.817317859	25.39072086	0.004086046
SLC27A4	2.749917635	22.89264731	0.069558791
TMEM2	2.639332142	21.71701318	0.030356906
TUBGCP2	2.503440004	22.39117474	0.069971254
MAP4	2.482400623	25.41281013	0.028960226
PPIA	2.416979184	26.67446588	0.009757591
TMEM209	2.399575563	23.38791963	0.057646501
NELFA	2.399086535	22.29015004	0.036082937

ZNF592	2.378573191	22.52832743	0.174257327
CKAP4	2.351043641	27.37635077	0.020417344
HM13	2.281775931	23.33755832	0.041656172
DAD1	2.276219809	26.25949537	0.018677625
SEC63	2.264575443	23.77907719	0.051763475
ACACA	2.256011899	27.79287225	0.000738658
TIMM23	2.244460602	22.10452137	0.065064866
HAUS8	2.203732365	26.54427374	0.002767761
TECR	2.202024418	28.29348562	0.001688423
TMEM201	2.197747421	23.45585837	0.033556744
PRPF3	2.192375892	27.46330785	0.002553333
MED31	2.165265747	21.23597754	0.112350471
SS18	2.110026587	20.56631146	0.086414613
CCDC137	2.109284396	25.26218597	0.012079647
MTG2	2.093563542	25.50796582	0.013997644
PHGDH	2.027901936	21.47961859	0.217519946
MRPL47	2.024824015	22.56920093	0.129655841
CCDC47	2.014681117	24.06231626	0.039398323
MPDU1	2.014477402	24.4627432	0.060180445
THAP11	2.009491588	21.3808442	0.117877444
SLC25A30	2.002982288	23.31571808	0.161708336
SLC25A14	2.002982288	23.31571808	0.161708336
FAM210A	2.00292784	27.10534321	0.002750422
GALNT2	2.001453434	24.42318404	0.062947274

BIBLIOGRAPHY

Al-Amrani, S., Al-Jabri, Z., Al-Zaabi, A., Alshekaili, J., & Al-Khabori, M. (2021). Proteomics: Concepts and applications in human medicine. *World J Biol Chem*, *12*(5), 57-69. https://doi.org/10.4331/wjbc.v12.i5.57

Alerasool, N., Leng, H., Lin, Z. Y., Gingras, A. C., & Taipale, M. (2022). Identification and functional characterization of transcriptional activators in human cells. *Mol Cell*, *82*(3), 677-695 e677. https://doi.org/10.1016/j.molcel.2021.12.008

Alfert, A., Moreno, N., & Kerl, K. (2019). The BAF complex in development and disease. *Epigenetics Chromatin*, *12*(1), 19. https://doi.org/10.1186/s13072-019-0264-y

Almansour, N. M. (2022). Triple-Negative Breast Cancer: A Brief Review About Epidemiology, Risk Factors, Signaling Pathways, Treatment and Role of Artificial Intelligence. *Front Mol Biosci*, *9*, 836417. https://doi.org/10.3389/fmolb.2022.836417

AlphaFold Protein Structure Database. (2021).

Audia, J. E., & Campbell, R. M. (2016). Histone Modifications and Cancer. *Cold Spring Harb Perspect Biol*, *8*(4), a019521. https://doi.org/10.1101/cshperspect.a019521

Bajrami, E., & Spiroski, M. (2016). Genomic Imprinting. *Open Access Maced J Med Sci*, *4*(1), 181-184. https://doi.org/10.3889/oamjms.2016.028

Bannister, A. J., & Kouzarides, T. (2011). Regulation of chromatin by histone modifications. *Cell Res*, *21*(3), 381-395. https://doi.org/10.1038/cr.2011.22

Bayani, J., & Squire, J. A. (2004). Fluorescence in situ Hybridization (FISH). *Curr Protoc Cell Biol*, *Chapter 22*, Unit 22 24. https://doi.org/10.1002/0471143030.cb2204s23

Berger, S. L. (2007). The complex language of chromatin regulation during transcription. *Nature*, *447*(7143), 407-412. https://doi.org/10.1038/nature05915

Biswas, S., & Rao, C. M. (2018). Epigenetic tools (The Writers, The Readers and The Erasers) and their implications in cancer therapy. *Eur J Pharmacol*, *837*, 8-24. https://doi.org/10.1016/j.ejphar.2018.08.021

Bock, C., Datlinger, P., Chardon, F., Coelho, M. A., Dong, M. B., Lawson, K. A., Lu, T., Maroc, L., Norman, T. M., Song, B., Stanley, G., Chen, S., Garnett, M., Li, W., Moffat, J., Qi, L. S., Shapiro, R. S., Shendure, J., Weissman, J. S., & Zhuang, X. (2022). High-content CRISPR screening. *Nat Rev Methods Primers*, *2*(1). https://doi.org/10.1038/s43586-022-00098-7

Butler, M. G. (2009). Genomic imprinting disorders in humans: a mini-review. *J Assist Reprod Genet*, *26*(9-10), 477-486. https://doi.org/10.1007/s10815-009-9353-3

Cancer Dependency Map. (2021). Broad Institute. https://depmap.org/portal/

Cheng, Y., Shen, Z., Gao, Y., Chen, F., Xu, H., Mo, Q., Chu, X., Peng, C. L., McKenzie, T. T., Palacios, B. E., Hu, J., Zhou, H., & Long, J. (2022). Phase transition and remodeling complex assembly are important for SS18-SSX oncogenic activity in synovial sarcomas. *Nat Commun*, *13*(1), 2724. https://doi.org/10.1038/s41467-022-30447-9

Cooper, G. W., & Hong, A. L. (2022). SMARCB1-Deficient Cancers: Novel Molecular Insights and Therapeutic Vulnerabilities. *Cancers (Basel)*, *14*(15). https://doi.org/10.3390/cancers14153645

Cunha-Ferreira, I., Chazeau, A., Buijs, R. R., Stucchi, R., Will, L., Pan, X., Adolfs, Y., van der Meer, C., Wolthuis, J. C., Kahn, O. I., Schatzle, P., Altelaar, M., Pasterkamp, R. J., Kapitein, L. C., & Hoogenraad, C. C. (2018). The HAUS Complex Is a Key Regulator of Non-centrosomal Microtubule Organization during Neuronal Development. *Cell Rep*, *24*(4), 791-800. https://doi.org/10.1016/j.celrep.2018.06.093

Damiani, E., Duran, M. N., Mohan, N., Rajendran, P., & Dashwood, R. H. (2020). Targeting Epigenetic 'Readers' with Natural Compounds for Cancer Interception. *J Cancer Prev*, *25*(4), 189-203. https://doi.org/10.15430/JCP.2020.25.4.189

Dawson, M. A., & Kouzarides, T. (2012). Cancer epigenetics: from mechanism to therapy. *Cell*, *150*(1), 12-27. https://doi.org/10.1016/j.cell.2012.06.013

de Bruijn, D. R., & Geurts van Kessel, A. (2006). Common origin of the human synovial sarcoma associated SS18 and SS18L1 gene loci. *Cytogenet Genome Res*, *112*(3-4), 222-226. https://doi.org/10.1159/000089874

de Bruijn, D. R., Kater-Baats, E., Eleveld, M., Merkx, G., & Geurts Van Kessel, A. (2001). Mapping and characterization of the mouse and human SS18 genes, two human SS18-like genes and a mouse Ss18 pseudogene. *Cytogenet Cell Genet*, *92*(3-4), 310-319. https://doi.org/10.1159/000056920

Dhar, G. A., Saha, S., Mitra, P., & Nag Chaudhuri, R. (2021). DNA methylation and regulation of gene expression: Guardian of our health. *Nucleus (Calcutta)*, *64*(3), 259-270. https://doi.org/10.1007/s13237-021-00367-y

dos Santos, N. R., de Bruijn, D. R., Balemans, M., Janssen, B., Gartner, F., Lopes, J. M., de Leeuw, B., & Geurts van Kessel, A. (1997). Nuclear localization of SYT, SSX and the synovial sarcoma-associated SYT-SSX fusion proteins. *Hum Mol Genet*, *6*(9), 1549-1558. https://doi.org/10.1093/hmg/6.9.1549

Ehrlich, M. (2009). DNA hypomethylation in cancer cells. *Epigenomics*, *1*(2), 239-259. https://doi.org/10.2217/epi.09.33

Ehrlich, M. (2019). DNA hypermethylation in disease: mechanisms and clinical relevance. *Epigenetics*, *14*(12), 1141-1163. https://doi.org/10.1080/15592294.2019.1638701

Foulkes, W. D., Smith, I. E., & Reis-Filho, J. S. (2010). Triple-negative breast cancer. *N Engl J Med*, *363*(20), 1938-1948. https://doi.org/10.1056/NEJMra1001389

Gabel, C. A., Li, Z., DeMarco, A. G., Zhang, Z., Yang, J., Hall, M. C., Barford, D., & Chang, L. (2022). Molecular architecture of the augmin complex. *Nat Commun*, *13*(1), 5449. https://doi.org/10.1038/s41467-022-33227-7

Gade, P., & Kalvakolanu, D. V. (2012). Chromatin immunoprecipitation assay as a tool for analyzing transcription factor activity. *Methods Mol Biol*, *809*, 85-104. https://doi.org/10.1007/978-1-61779-376-9_6

Gazendam, A. M., Popovic, S., Munir, S., Parasu, N., Wilson, D., & Ghert, M. (2021). Synovial Sarcoma: A Clinical Review. *Curr Oncol*, *28*(3), 1909-1920. https://doi.org/10.3390/curroncol28030177

Guo, C., Ma, X., Gao, F., & Guo, Y. (2023). Off-target effects in CRISPR/Cas9 gene editing. *Front Bioeng Biotechnol*, *11*, 1143157. https://doi.org/10.3389/fbioe.2023.1143157

Gure, A. O., Wei, I. J., Old, L. J., & Chen, Y. T. (2002). The SSX gene family: characterization of 9 complete genes. *Int J Cancer*, *101*(5), 448-453. https://doi.org/10.1002/ijc.10634

Hale, R., Sandakly, S., Shipley, J., & Walters, Z. (2019). Epigenetic Targets in Synovial Sarcoma: A Mini-Review. *Front Oncol*, *9*, 1078. https://doi.org/10.3389/fonc.2019.01078

Hamilton, J. P. (2011). Epigenetics: principles and practice. *Dig Dis*, *29*(2), 130-135. https://doi.org/10.1159/000323874

Handy, D. E., Castro, R., & Loscalzo, J. (2011). Epigenetic modifications: basic mechanisms and role in cardiovascular disease. *Circulation*, *123*(19), 2145-2156. https://doi.org/10.1161/CIRCULATIONAHA.110.956839

Honke, K. (2018). Biological functions of sulfoglycolipids and the EMARS method for identification of co-clustered molecules in the membrane microdomains. *J Biochem*, *163*(4), 253-263. https://doi.org/10.1093/jb/mvx078

Iorio, E., Caramujo, M. J., Cecchetti, S., Spadaro, F., Carpinelli, G., Canese, R., & Podo, F. (2016). Key Players in Choline Metabolic Reprograming in Triple-Negative Breast Cancer. *Front Oncol*, *6*, 205. https://doi.org/10.3389/fonc.2016.00205

Jin, B., Li, Y., & Robertson, K. D. (2011). DNA methylation: superior or subordinate in the epigenetic hierarchy? *Genes Cancer*, *2*(6), 607-617. https://doi.org/10.1177/1947601910393957

Kadoch, C., & Crabtree, G. R. (2013). Reversible disruption of mSWI/SNF (BAF) complexes by the SS18-SSX oncogenic fusion in synovial sarcoma. *Cell*, *153*(1), 71-85. https://doi.org/10.1016/j.cell.2013.02.036

Kawazu, M., Kojima, S., Ueno, T., Totoki, Y., Nakamura, H., Kunita, A., Qu, W., Yoshimura, J., Soda, M., Yasuda, T., Hama, N., Saito-Adachi, M., Sato, K., Kohsaka, S., Sai, E., Ikemura, M., Yamamoto, S., Ogawa, T., Fukayama, M., . . . Mano, H. (2017). Integrative analysis of genomic alterations in triple-negative breast cancer in association with homologous recombination deficiency. *PLoS Genet*, *13*(6), e1006853. https://doi.org/10.1371/journal.pgen.1006853

Kouzarides, T. (2007). Chromatin modifications and their function. *Cell*, *128*(4), 693-705. https://doi.org/10.1016/j.cell.2007.02.005

Kraus, J., Travis, S. M., King, M. R., & Petry, S. (2023). Augmin is a Ran-regulated spindle assembly factor. *J Biol Chem*, *299*(6), 104736. https://doi.org/10.1016/j.jbc.2023.104736

Kuang, J., Zhai, Z., Li, P., Shi, R., Guo, W., Yao, Y., Guo, J., Zhao, G., He, J., Xu, S., Wu, C., Yu, S., Zhou, C., Wu, L., Qin, Y., Cai, B., Li, W., Wu, Z., Li, X., . . . Pei, D. (2021). SS18 regulates pluripotent-somatic transition through phase separation. *Nat Commun*, *12*(1), 4090. https://doi.org/10.1038/s41467-021-24373-5

Kulka, J., Tokes, A. M., Kaposi-Novak, P., Udvarhelyi, N., Keller, A., & Schaff, Z. (2006). Detection of HER-2/neu gene amplification in breast carcinomas using quantitative real-time PCR - a comparison with immunohistochemical and FISH results. *Pathol Oncol Res*, *12*(4), 197-204. https://doi.org/10.1007/BF02893412

Lawo, S., Bashkurov, M., Mullin, M., Ferreria, M. G., Kittler, R., Habermann, B., Tagliaferro, A., Poser, I., Hutchins, J. R., Hegemann, B., Pinchev, D., Buchholz, F., Peters, J. M., Hyman, A. A., Gingras, A. C., & Pelletier, L. (2009). HAUS, the 8-subunit human Augmin complex, regulates centrosome and spindle integrity. *Curr Biol*, *19*(10), 816-826. https://doi.org/10.1016/j.cub.2009.04.033

Lehmann, B. D., Bauer, J. A., Chen, X., Sanders, M. E., Chakravarthy, A. B., Shyr, Y., & Pietenpol, J. A. (2011). Identification of human triple-negative breast cancer

subtypes and preclinical models for selection of targeted therapies. *J Clin Invest*, *121*(7), 2750-2767. https://doi.org/10.1172/JCI45014

Lehmann, B. D., Jovanovic, B., Chen, X., Estrada, M. V., Johnson, K. N., Shyr, Y., Moses, H. L., Sanders, M. E., & Pietenpol, J. A. (2016). Refinement of Triple-Negative Breast Cancer Molecular Subtypes: Implications for Neoadjuvant Chemotherapy Selection. *PLoS One*, *11*(6), e0157368. https://doi.org/10.1371/journal.pone.0157368

Lehmann, B. D., & Pietenpol, J. A. (2014). Identification and use of biomarkers in treatment strategies for triple-negative breast cancer subtypes. *J Pathol*, *232*(2), 142-150. https://doi.org/10.1002/path.4280

Li, Y., & Sousa, R. (2012). Expression and purification of E. coli BirA biotin ligase for in vitro biotinylation. *Protein Expr Purif*, *82*(1), 162-167. https://doi.org/10.1016/j.pep.2011.12.008

Martire, S., Nguyen, J., Sundaresan, A., & Banaszynski, L. A. (2020). Differential contribution of p300 and CBP to regulatory element acetylation in mESCs. *BMC Mol Cell Biol*, *21*(1), 55. https://doi.org/10.1186/s12860-020-00296-9

Mashtalir, N., D'Avino, A. R., Michel, B. C., Luo, J., Pan, J., Otto, J. E., Zullow, H. J., McKenzie, Z. M., Kubiak, R. L., St Pierre, R., Valencia, A. M., Poynter, S. J., Cassel, S. H., Ranish, J. A., & Kadoch, C. (2018). Modular Organization and Assembly of SWI/SNF Family Chromatin Remodeling Complexes. *Cell*, *175*(5), 1272-1288 e1220. https://doi.org/10.1016/j.cell.2018.09.032

May, D. G., Scott, K. L., Campos, A. R., & Roux, K. J. (2020). Comparative Application of BioID and TurboID for Protein-Proximity Biotinylation. *Cells*, *9*(5). https://doi.org/10.3390/cells9051070

McBride, M. J., Pulice, J. L., Beird, H. C., Ingram, D. R., D'Avino, A. R., Shern, J. F., Charville, G. W., Hornick, J. L., Nakayama, R. T., Garcia-Rivera, E. M., Araujo, D. M., Wang, W. L., Tsai, J. W., Yeagley, M., Wagner, A. J., Futreal, P. A., Khan, J., Lazar, A. J., & Kadoch, C. (2018). The SS18-SSX Fusion Oncoprotein Hijacks BAF Complex Targeting and Function to Drive Synovial Sarcoma. *Cancer Cell*, *33*(6), 1128-1141 e1127. https://doi.org/10.1016/j.ccell.2018.05.002

Mehus, A. A., Anderson, R. H., & Roux, K. J. (2016). BioID Identification of Lamin-Associated Proteins. *Methods Enzymol*, *569*, 3-22. https://doi.org/10.1016/bs.mie.2015.08.008

Middeljans, E., Wan, X., Jansen, P. W., Sharma, V., Stunnenberg, H. G., & Logie, C. (2012). SS18 together with animal-specific factors defines human BAF-type SWI/SNF complexes. *PLoS One*, *7*(3), e33834. https://doi.org/10.1371/journal.pone.0033834

Mittal, P., & Roberts, C. W. M. (2020). The SWI/SNF complex in cancer - biology, biomarkers and therapy. *Nat Rev Clin Oncol*, *17*(7), 435-448. https://doi.org/10.1038/s41571-020-0357-3

Moore, L. D., Le, T., & Fan, G. (2013). DNA methylation and its basic function. *Neuropsychopharmacology*, *38*(1), 23-38. https://doi.org/10.1038/npp.2012.112

Morey, L., & Helin, K. (2010). Polycomb group protein-mediated repression of transcription. *Trends Biochem Sci*, *35*(6), 323-332. https://doi.org/10.1016/j.tibs.2010.02.009

Murtaza, N., Uy, J., & Singh, K. K. (2020). Emerging proteomic approaches to identify the underlying pathophysiology of neurodevelopmental and neurodegenerative disorders. *Mol Autism*, *11*(1), 27. https://doi.org/10.1186/s13229-020-00334-5

Nguyen, T. M. T., Kim, J., Doan, T. T., Lee, M. W., & Lee, M. (2020). APEX Proximity Labeling as a Versatile Tool for Biological Research. *Biochemistry*, *59*(3), 260-269. https://doi.org/10.1021/acs.biochem.9b00791

Ohe, S., Kubota, Y., Yamaguchi, K., Takagi, Y., Nashimoto, J., Kozuka-Hata, H., Oyama, M., Furukawa, Y., & Takekawa, M. (2022). ERK-mediated NELF-A phosphorylation promotes transcription elongation of immediate-early genes by releasing promoter-proximal pausing of RNA polymerase II. *Nat Commun*, *13*(1), 7476. https://doi.org/10.1038/s41467-022-35230-4

Oostdyk, L. T., Shank, L., Jividen, K., Dworak, N., Sherman, N. E., & Paschal, B. M. (2019). Towards improving proximity labeling by the biotin ligase BirA. *Methods*, *157*, 66-79. https://doi.org/10.1016/j.ymeth.2018.11.003

Pagliaroli, L., & Trizzino, M. (2021). The Evolutionary Conserved SWI/SNF Subunits ARID1A and ARID1B Are Key Modulators of Pluripotency and Cell-Fate Determination. *Front Cell Dev Biol*, *9*, 643361. https://doi.org/10.3389/fcell.2021.643361

Perani, M., Antonson, P., Hamoudi, R., Ingram, C. J., Cooper, C. S., Garrett, M. D., & Goodwin, G. H. (2005). The proto-oncoprotein SYT interacts with SYT-interacting protein/co-activator activator (SIP/CoAA), a human nuclear receptor co-activator with similarity to EWS and TLS/FUS family of proteins. *J Biol Chem*, *280*(52), 42863-42876. https://doi.org/10.1074/jbc.M502963200

Perani, M., Ingram, C. J., Cooper, C. S., Garrett, M. D., & Goodwin, G. H. (2003). Conserved SNH domain of the proto-oncoprotein SYT interacts with components of the human chromatin remodelling complexes, while the QPGY repeat domain forms homo-oligomers. *Oncogene*, *22*(50), 8156-8167. https://doi.org/10.1038/sj.onc.1207031

Pradhan, A., & Liu, Y. (2005). A multifunctional domain of the calcium-responsive transactivator (CREST) that inhibits dendritic growth in cultured neurons. *J Biol Chem*, *280*(26), 24738-24743. https://doi.org/10.1074/jbc.M504018200

Roux, K. J., Kim, D. I., Burke, B., & May, D. G. (2018). BioID: A Screen for Protein-Protein Interactions. *Curr Protoc Protein Sci*, *91*, 19 23 11-19 23 15. https://doi.org/10.1002/cpps.51

Sadasivam, S., & DeCaprio, J. A. (2013). The DREAM complex: master coordinator of cell cycle-dependent gene expression. *Nat Rev Cancer*, *13*(8), 585-595. https://doi.org/10.1038/nrc3556

Sears, R. M., May, D. G., & Roux, K. J. (2019). BioID as a Tool for Protein-Proximity Labeling in Living Cells. *Methods Mol Biol*, *2012*, 299-313. https://doi.org/10.1007/978-1-4939-9546-2_15

Sheikh, B. N., Guhathakurta, S., & Akhtar, A. (2019). The non-specific lethal (NSL) complex at the crossroads of transcriptional control and cellular homeostasis. *EMBO Rep*, *20*(7), e47630. https://doi.org/10.15252/embr.201847630

Shen, A., Liu, L., Huang, Y., Shen, Z., Wu, M., Chen, X., Wu, X., Lin, X., Chen, Y., Li, L., Cheng, Y., Chu, J., Sferra, T. J., Wei, L., Zhuang, Q., & Peng, J. (2021). Down-Regulating HAUS6 Suppresses Cell Proliferation by Activating the p53/p21 Pathway in Colorectal Cancer. *Front Cell Dev Biol*, *9*, 772077. https://doi.org/10.3389/fcell.2021.772077

Sokpor, G., Xie, Y., Rosenbusch, J., & Tuoc, T. (2017). Chromatin Remodeling BAF (SWI/SNF) Complexes in Neural Development and Disorders. *Front Mol Neurosci*, *10*, 243. https://doi.org/10.3389/fnmol.2017.00243

Sugeedha, J., Gautam, J., & Tyagi, S. (2021). SET1/MLL family of proteins: functions beyond histone methylation. *Epigenetics*, *16*(5), 469-487. https://doi.org/10.1080/15592294.2020.1809873

Tamaki, S., Fukuta, M., Sekiguchi, K., Jin, Y., Nagata, S., Hayakawa, K., Hineno, S., Okamoto, T., Watanabe, M., Woltjen, K., Ikeya, M., Kato, T., Jr., & Toguchida, J. (2015). SS18-SSX, the Oncogenic Fusion Protein in Synovial Sarcoma, Is a Cellular Context-Dependent Epigenetic Modifier. *PLoS One*, *10*(11), e0142991. https://doi.org/10.1371/journal.pone.0142991

Tong, X., Tang, R., Xu, J., Wang, W., Zhao, Y., Yu, X., & Shi, S. (2022). Liquid-liquid phase separation in tumor biology. *Signal Transduct Target Ther*, *7*(1), 221. https://doi.org/10.1038/s41392-022-01076-x

Uscanga-Perales, G. I., S.K., S.-F., & Ortiz-López, R. (2016). Triple negative breast cancer: Deciphering the biology and heterogeneity. *Medicina Universitaria*.

Wang, T., Wei, J. J., Sabatini, D. M., & Lander, E. S. (2014). Genetic screens in human cells using the CRISPR-Cas9 system. *Science*, *343*(6166), 80-84. https://doi.org/10.1126/science.1246981

Wang, X., Lee, R. S., Alver, B. H., Haswell, J. R., Wang, S., Mieczkowski, J., Drier, Y., Gillespie, S. M., Archer, T. C., Wu, J. N., Tzvetkov, E. P., Troisi, E. C., Pomeroy, S. L., Biegel, J. A., Tolstorukov, M. Y., Bernstein, B. E., Park, P. J., & Roberts, C. W. M. (2017). SMARCB1-mediated SWI/SNF complex function is essential for enhancer regulation. *Nature Genetics*, *49*(2), 289-295. https://doi.org/10.1038/ng.3746

Wanior, M., Kramer, A., Knapp, S., & Joerger, A. C. (2021). Exploiting vulnerabilities of SWI/SNF chromatin remodelling complexes for cancer therapy. *Oncogene*, *40*(21), 3637-3654. https://doi.org/10.1038/s41388-021-01781-x

Wei, J., Huang, K., Yang, C., & Kang, C. (2017). Non-coding RNAs as regulators in epigenetics (Review). *Oncology reports*, *37*. https://doi.org/10.3892/or.2016.5236

Wiggs, A., Molina, S., Sumner, S. J., & Rushing, B. R. (2022). A Review of Metabolic Targets of Anticancer Nutrients and Nutraceuticals in Pre-Clinical Models of Triple-Negative Breast Cancer. *Nutrients*, *14*(10). https://doi.org/10.3390/nu14101990

Won, K. A., & Spruck, C. (2020). Triple-negative breast cancer therapy: Current and future perspectives (Review). *Int J Oncol*, *57*(6), 1245-1261. https://doi.org/10.3892/ijo.2020.5135

Yang, A. Y., Kim, H., Li, W., & Kong, A. N. (2016). Natural compound-derived epigenetic regulators targeting epigenetic readers, writers and erasers. *Curr Top Med Chem*, *16*(7), 697-713. https://doi.org/10.2174/1568026615666150826114359

Yang, X., Liu, M., Li, M., Zhang, S., Hiju, H., Sun, J., Mao, Z., Zheng, M., & Feng, B. (2020). Epigenetic modulations of noncoding RNA: a novel dimension of Cancer biology. *Mol Cancer*, *19*(1), 64. https://doi.org/10.1186/s12943-020-01159-9

Yedier-Bayram, O., Gokbayrak, B., Kayabolen, A., Aksu, A. C., Cavga, A. D., Cingoz, A., Kala, E. Y., Karabiyik, G., Gunsay, R., Esin, B., Morova, T., Uyulur, F., Syed, H., Philpott, M., Cribbs, A. P., Kung, S. H. Y., Lack, N. A., Onder, T. T., & Bagci-Onder, T. (2022). EPIKOL, a chromatin-focused CRISPR/Cas9-based screening platform, to identify cancer-specific epigenetic vulnerabilities. *Cell Death Dis*, *13*(8), 710. https://doi.org/10.1038/s41419-022-05146-4

Yin, L., Duan, J. J., Bian, X. W., & Yu, S. C. (2020). Triple-negative breast cancer molecular subtyping and treatment progress. *Breast Cancer Res*, *22*(1), 61. https://doi.org/10.1186/s13058-020-01296-5