Mechanisms controlling SMAD-dependent transcription and chromatin modification

Roy Baas

ISBN: 978-94-6295-641-4
Printed by: ProefschriftMaken || www.proefschriftmaken.nl

About the cover: In this thesis, confocal microscopy was used to study the localization of GFP-fusion proteins (green). Additional staining identifies important cellular structures like the DNA in the nucleus (blue) and Actin fibers that are found in the cytoplasm (red). Image by Roy Baas.

Copyright © 2017 by Roy Baas

Mechanisms controlling SMAD-dependent transcription and chromatin modification

Mechanismen voor regulatie van SMAD-afhankelijke transcriptie en chromatine modificaties

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht
op gezag van de rector magnificus, prof.dr. G.J. van der Zwaan,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen
op donderdag 18 mei 2017 des middags te 2.30 uur

door

Roy Baas

geboren op 8 augustus 1988
te Vlaardingen

Promotor: Prof.dr. H.Th.M. Timmers

Table of contents

Chapter 1
General introduction 7
Thesis outline 21

Chapter 2
A novel microscopy-based high-throughput screening method to identify proteins that regulate global histone modification levels. 23

Chapter 3
Quantitative proteomics of the SMAD (SMA and MAD related) transcription factor family identifies Importin 5 as a Bone Morphogenetic Protein receptor-SMAD specific Importin. 63

Chapter 4
The Mixed Lineage Leukemia 4 (MLL4) methyltransferase complex is a mediator of Transforming Growth Factor beta (TGF-β)-activated gene transcription. 85

Addendum
Interactome of the KDM6A/UTX subunit of the MLL3 and MLL4 histone methyltransferase complexes. 95

Chapter 5
General discussion 107

Appendices
References 116
List of abbreviations 134
Nederlandse samenvatting voor niet-ingewijden 138
Dankwoord 142
Curriculum vitae 148
List of publications 149

Chapter 1

General introduction

Regulation of molecular processes in the cell is critical for proper cellular function and identity. To maintain this function and identity, proteins derived from approximately 20,000 protein coding genes encoded on the human genome need to be present at the right place and at the right time. Eukaryotic cells possess various strategies to control the activity and spatiotemporal distribution of proteins on transcriptional and post-transcriptional levels.

Transcription

The process of RNA polymerase II (RNA pol II) mediated protein gene transcription starts with transcription initiation. Various genomic features surrounding the promoter, like the B recognition element (1), CpG-islands (2), downstream promoter element (3), initiator (4), motif 10 element (5) and TATA box (6-8) contribute to promoter recognition and the initiation of transcription. Transcription starts by basal transcription factors (GTFs) binding to the core promoter, resulting in the assembly of the pre-initiation complex (PIC). Formation of the PIC starts with the association of the megadalton-sized multi-protein complex transcription factor II D (TFIID) that consists of TATA-binding protein (TBP) and 13 TBP associated factors (TAFs) that are present in one or two copies (9). TFIID binds to DNA through TBP and subsequently TFIIA, TFIIB, TFIIF bind and recruit RNA pol II (6,10,11). Following assembly, TFIIE and TFIIH are recruited which allows for ATP-dependent opening of the promoter via the helicase domain located on the TFIIH subunits XPD/ERCC3 (12-14). The opening of the DNA allows single stranded DNA to enter the active site of the RNA pol II (6,10,15). TFIIH subunits CDK7 and cyclin H phosphorylate RNA pol II on its C-terminal domain stimulating polymerase escape and mRNA transcription to start.

Chromatin and transcriptional control

The human reference genome has an approximated length of 3.2 giga base pairs (Gbp). With an average length of 3.4 Å per base pair (16) this roughly translates to 2 m of DNA per diploid genome that needs to be fit inside its 6 μm diameter cell nucleus (17). Nevertheless, these cells need to able to access the proper genetic information stored in DNA for their cellular function. To efficiently store and access the genetic information in the nucleus the DNA is wrapped around an octamer of core histone proteins (Fig. 1). Four pairs of histones H2A, H2B, H3 and H4 form a positively charged protein core around which ~146-bp of negatively charged DNA is wrapped forming the nucleosome. The DNA is held in place by histone H1 that sits on top of the wound DNA. The packing of DNA into nucleosomes allows for packing into a higher order chromatin structure. Densely packed heterochromatin allows for efficient storage of DNA, however transcription from these regions is restricted. Alternatively, DNA can be packed into euchromatin, which is less densely packed in comparison to heterochromatin. This reduced compaction allows the DNA to remain accessible for transcription by the basal transcription machinery of the cell.

Histone modifications

The histones in the nucleosome consists of a globular core from which unstructured N- and C-terminal tails protrude (18). These histone tails contain several residues that can be chemically modified. These covalent bound post-translational modifications (PTMs) can be placed ("written") and removed ("erased") by catalytic enzymes (19,20). These histone PTMs include small molecular modifications like acetylation (21) on lysine residues, methylation

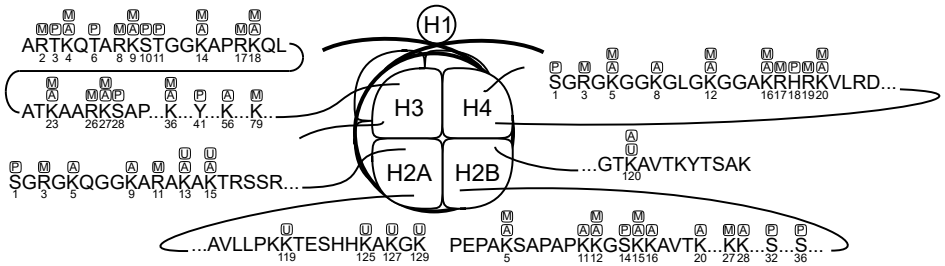


Figure 1. Schematic representation of the nucleosome. DNA is wrapped around four pairs of histone proteins (H2A, H2B, H3 and H4) to form the nucleosomal particle. Linker histone H1 has been proposed to bind at the dyad axis. C- and N-terminal tails protrude from the nucleosome that can be modified with various modifications. Positions of modified amino acids and modifications are indicated. A: acetylation, M: methylation, P: phosphorylation, U: ubiquitination. Figure based on (275,276).

(22) on lysine and arginine residues and phosphorylation (23) of serine, threonine and tyrosine residues (24,25), but also conjugation of macromolecules via ubiquitination (26) or less studied sumoylation (27) or neddylation (28) of lysine residues. The presence of these PTMs in combination with their location on the histone tails dictates whether the chromatin environment is transcriptionally repressive or permissive. This is a consequence of the proteins that can recognize specific chromatin modifications at distinct places on the histones, the so-called "readers" that each have their specific functions.

Histone acetylation and deacetylation

Histone acetylation is a PTM that modifies lysine residues which can generally be found near transcriptionally active enhancers and promoters (29). Lysine acetylation causes the respective lysine residue to lose its positive charge to become neutral. DNA wrapped around the histone proteins is negatively charged and charge reduction of the positive core weakens the interaction between the histone proteins and the DNA. The loosening of the DNA results in increased accessibility for proteins that interact with DNA like the basal transcription machinery and TFIID in particular. Additionally, the TFIID largest subunit TAF1 specifically detects and binds to acetylated H3 and H4 histones via its double bromodomain (30), facilitating recruitment of TFIID and PIC assembly required for transcription (30). Deposition of acetylation on histones is catalyzed by histone acetyltransferases (HATs). The capacity to both bind and write these acetylation marks allows TAF1 to propagate a transcriptional favorable environment. Additional coactivator complexes possess HAT activity by which transcription is promoted. For example, binding of gene-specific transcription factors (GSTFs) to DNA recruit CREB-binding protein (CBP) or its homologue p300. CBP/p300 acetylates Histone 3 lysine 27 (H3K27ac), which antagonizes polycomb-mediated silencing (see below) (31,32). An additional coactivator complex that promotes gene transcription is the Spt-Ada-Gcn5 acetyltransferase (SAGA) complex. SAGA consists of three modules that contain the HAT activity towards histone H3, deubiquitinase (DUB, see below) activity towards H2Bub and the capability to bind to TBP (33-35). Gcn5 is the catalytic subunit for histone HAT activity and is regulated by Ada2 and Ada3 (36,37). Furthermore, SAGA shares TAF5/6/9/10/12 subunits with TFIID and was shown to display some degree of redundancy with TFIID mediated gene transcription (38). Binding of TBP to the TATA box is regulated in SAGA by Spt3 and Spt8 (39). Approximately 10% of the yeast genome is dominated by SAGA and most of these genes are stress induced (40), however SAGA transcription is less specific in humans (41). Removal of histone acetylation

is mediated by histone deacetylases (HDACs) in various multi-protein complexes like the well-studied Nucleosome Remodeling Deacetylase (NuRD) complex (42), Sin3 complex (43) or CoREST complex (44). A less well studied HDAC-containing protein complex is the mitotic deacetylase (MiDAC) complex (45). This protein complex was identified to consist of the HDAC1/2, ELMSAN1 and DNTTIP1 proteins and is stabilized during mitosis. However the precise function of MiDAC remains an open question (46-48).

The variable response of histone methylation

In contrast to lysine acetylation, methylation can occur on both lysine and arginine residues and this modification can occur multiple times on the same residue. Methylation of a target residue requires the S-adenosyl-L-methionine (SAM) molecule as a donor for the methyl group (49). Lysine residues can be modified with up to three methyl groups, while arginine residues can be modified twice, but di-methylation of arginine occurs in stereochemically different manners. Furthermore, methylation of lysine and arginine residues on histone tails does not remove the positive charge of the respective residues (26,50).

Arginine methylation and demethylation

Arginine residue methylation is catalyzed by the protein arginine methyltransferases (PRMTs) that are classified into three different subgroups based on the formation of methylation state of the arginine. Arginine di-methylation can be symmetrical (Rme2s) or asymmetrical (Rme2a) and the placement of modification can have different biological outcomes (51). Once such example is H3R2 methylation. H3R2me1 and H3R2me2s is a PTM associated with gene activation (52-54) while H3R2me2a is associated with gene repression (55-57). The type 1 class contains PRMT1/2/3/4/6/8 and generates Rme2a, while PRMT5/7 belong to the type 2 class that can generate Rme2s. Type 3 only consists of PRMT9, which can only form Rme1. Removal of methylation from arginine residues is catalyzed by deamination enzymes that convert methylated arginine into citrulline (58,59).

Euchromatic lysine methylation

Up to three methyl groups can be placed onto the ε-amino group of lysine residues. Like the methylation of arginine residues, the cellular biological response depends on the methylation state and respective location of the modified residue on the histone proteins. Most lysine methylations are catalyzed by the Su (var), Enhancer of Zeste, and Trithorax (SET) domain-containing proteins (60). The exception is the disruptor of telomeric silencing 1 like (DOT1L) methyltransferase (KMT) that catalyzes mono-, di- and tri-methylation of H3K79, located in the globular part of the histone (61,62). Methylated lysines associate to both euchromatic and heterochromatic regions. Euchromatic methylation marks include H3K4 on enhancers and promoters and H3K36 at elongating gene bodies. H3K36 methylation is furthermore also implicated in alternative splicing, DNA replication, recombination and repair (63). H3K79 methylation is associated with active genes transcription but is also implicated in DNA double strand repair (64-66).

Active enhancers are generally mono-methylated at H3K4 (H3K4me1) while tri-methylated H3K4 (H3K4me3) is indicative of transcriptionally active promoter sites. Many different proteins can recognize and associate with H3K4me3 (67-69). Examples are Jumonji (JmjC) domain-containing protein 2A (Jmjd2A) (70) and the SAGA subunit Sgf29 that binds H3K4me3 via its double Tudor domain (67). TFIID also interacts with H3K4me3

via the plant homeodomain (PHD) finger in TAF3 (68). This association is antagonized by phosphorylation of the neighboring residue H3T3 by the kinase Haspin, forming a molecular switch during mitosis (71). In yeast, Set1 is the sole KMT responsible for mono-, di- and tri-methylation of H3K4 and is part of the multiprotein complex of proteins associated with Set1 (COMPASS). In *Drosophila*, Set1 contains two additional homologues called Trithorax (Trx) and Trithorax-related (Trr), which each reside in a COMPASS-like complex. In humans, writers of the H3K4 methylation are six distinct multi-protein complexes that are descendants of the *Drosophila* methyltransferase proteins. These complexes either contain SET domain-containing protein 1A/B (SET1A/B, also known as KMT2F/G) or Mixed Lineage Leukemia 1-4 (MLL1-4, also known as KMT2A-D) methyltransferases as catalytic subunit. SET1A/B and are descendants of the fly SET1, MLL1/2 descend from Trx and MLL3/4 originate from Trr. The SET1A and SET1B complexes deposit bulk methylation of H3K4me3 (72), whereas MLL1 and MLL2 complexes tri-methylate H3K4 on a select subset of genes like the *Hox* genes (73). MLL3 and MLL4 complexes provide genome-wide mono-methylation of H3K4 on the regulatory enhancer sites (74,75). All six human SET/MLL complexes consist of a WRD5, RBBP5, ASH2L and DPY30 (WRAD) core (76) and a specific subset of proteins that only interact with a specific subset of these complexes. The SET1A, SET1B, MLL1 and MLL2 complexes interact with HCFC1 and OGT. These complexes are differentiated by interaction of MENIN and PSIP with the MLL1/2 complexes and CFP1 and WDR82 with SET1A/B. Furthermore, BOD1 and BOD1L are found exclusively in the SET1B complex. MLL3/4 complexes interact with NCOA6, PA1, PTIP and UTX (also known as lysine demethylase 6A [KDM6A]) (77). Originally, *MLL1* was identified as *ALL1* after a t(4;11) chromosome translocation event forming a fusion protein with AF4 resulting in acute lymphoblastic leukemia (ALL) (78). *ALL1* was renamed to *MLL1*, when more types of leukemia like acute myelogenous leukemia (AML) were found to be associated with *MLL1* translocations. Up to 10% of detectable translocation events involve MLL1 for certain types of leukemia (79) and *MLL1* is in 93% the main translocation partner in newborn infants (<1 month) with congenital ALL (80). Furthermore, patients with *MLL1* translocations run a high risk of relapse after treatment (79) and event free survival ranges from 22 to 54% (81-84). Multiple *MLL1* fusion partners recruit MLL to transcription favorable complexes like the DOT1L H3K79 methyltransferase complex (DotCom) and the super elongation complex (SEC) (85-89). AF4 is part of the SEC that stimulates RNA pol II to form full-length mRNA transcripts. SEC subunit ELL was shown to interact to RNA pol II subunit Rbp2 with the same properties as TFIIF, reducing RNA polymerase pausing and stimulating mRNA elongation (90).

Constitutive heterochromatin lysine methylation

The mammalian genome consists of various genomic elements required for proper gene function such as protein coding genes and elements controlling gene transcription like enhancers or CTCF insulator sites. However, other sequences like endogenous retroviruses, retrotransposons and repeat-rich sequences near centromeres and telomeres are also embedded into the genome. These sequences are a liability for genome integrity because of their ability to transcribe, self-duplicate and to recombine in an error prone manner. Therefore, it is important for the cell to silence these regions by packing these sequences into transcriptionally repressing heterochromatin. Heterochromatin methylation marks can typically be found on H3K9 or H3K27. Genomic unstable elements are silenced by packing into constitutive heterochromatin in all cell types and is considered to be very static (91). This

constitutive heterochromatin keeps DNA regions transcriptionally silent and are often gene poor. The constitutive heterochromatin is formed by deposition of H3K9me1/2 by G9a (92-95) and GLP (94) followed by H3K9me2/3 by SETDB1 (96), SUV39H1 and SUV39H2 (97,98). H3K9me2/3 can be bound by heterochromatin protein 1 (HP1) paralogues chromobox homolog 1/3/5 (CBX1/3/5) that multimerize and recruit additional repressive histone modifying enzymes (99-102). This results in reduction of lysine acetylation and gain of methylation of DNA at CpG islands (103,104), spreading of the constitutive heterochromatin state to nearby chromatin that can become over one Mbp in length (105).

Facultative heterochromatin lysine methylation

Cell types only need a subset of genes to perform properly, *e.g.* a neuron requires a different gene repertoire to properly function compared to a hepatocyte and the transcriptional profile from a proliferating stem cell deviates from a differentiated cell. Genes that are not required in a specific cell type are often silenced by deposition of H3K27me3 and formation of facultative heterochromatin (106). The presence of H3K27me3 on a certain chromatin region therefore differs between cell types and is dependent on the requirement of these cell type specific genes. Specific events like differentiation induce demethylation of these genomic regions and gain euchromatin marks, losing its repressive state and becoming transcriptionally active. H3K27me2/3 deposition is catalyzed by the polycomb repressive complex 2 (PRC2) that consists of the enhancer of zeste homologue 2 (EZH2) or EZH1 catalytic subunit (107), SUZ12 and EED and RBBP4/7 (108). The polycomb group (PcG) proteins were initially described in *Drosophila* where gene mutations were found to influence the fly body plan. Male flies carrying these mutations developed additional sex combs on the second and third legs, instead of only on the first leg (109). Later studies identified more protein complexes to be part of the PcG family like PRC1 and the polycomb repressive deubiquitinase complex (PR-DUB, see below). The PRC1 complex initially was identified in fractionation experiments in *Drosophila melanogaster* which core components were Polycomb (Pc), Polyhomeotic (Ph), Posterior Sex combs (Psc) and Ring1/Sce (110). In humans, various homologues of these proteins exist. CBX2/4/6/7/8, PHC1-3 and RING1A/B are homologous to Pc, Ph and Sce respectively (111). Additionally, six different homologs of the *Drosophila* Psc protein exist named Polycomb group RING finger protein 1-6 (PCGF1-6). These six proteins define the different PRC1 complexes that can be identified that assemble with different additional interacting proteins (112-114). Deposition of H3K27me2/3 is recognized by the PRC1 complex through the chromo-domain of the CBX proteins (115). Subsequently, RING1 mono-ubiquitinates histone H2A at lysine 119 (H2AK119ub1), leading to heterochromatin formation and transcriptional repression. In contrast to H3K27me3, H3K27 monomethylation is found to be associated with active gene transcription and enhancer fidelity (116). H3K27me1 correlates with active enhancers and promoters and can be found in combination with H3K4 mono-methylation in primary cells (117,118). These bivalently marked genes are thought to maintain the activation potential of genes without silencing them completely in case these genes are required during cellular differentiation (118).

Lysine demethylation

Lysine methylation removal is catalyzed by the LSD and Jumonji (JmjC) domain containing lysine demethylating (KDM) enzymes (89,119,120). LSD1 (also known as KDM1A) requires flavin adenine dinucleotide (FAD) as cofactor to demethylate H3K4 and

H3K9 mono- and dimethyl groups via its FAD-dependent amine oxidase domain (119). LSD1 is found to be enriched at gene promoters and the association of LSD1 in protein complexes determines its activity as a transcriptional repressor or activator. Interaction with the Co-REST complex demethylates H3K4 while interaction with the androgen receptor results in demethylation of H3K9 (119,121). In contrast, homologue LSD2 (KDM1B) is often found at gene bodies (122) and has high expression in oocytes where it facilitates *de novo* methylation of imprinted genes (123). JmjC domain containing proteins require α-ketoglutarate and iron as cofactor to demethylate mono-, di- and tri-methyl lysine (120). The JmjC domain containing proteins can be divided into seven subgroups, which can act on different histone residues (89,124). Jarid1A-D (also known as KDM5A-D) can remove H3K4 di- and trimethylation. Interestingly, KDM5C can dampen transcription but stimulate enhancers by reducing H3K4 methylation levels (125). JHDM2 and JMJD2A-D proteins (KDM3A and KDM4A-D) modulate K9 mono- and dimethylation or di- and tri-methylation respectively on both H3K9 and H3K36. Finally, UTX, JMJD3 and UTY (KDM6A-C) demethylate H3K27. UTX is located on the X chromosome and its homologue UTY (KDM6C) is male specific and located on the Y chromosome (126). The active site of UTY is conserved with UTX and JMJD3, however its activity is reduced.

Lysine ubiquitination

Acetylation and methylation of histones results in a relative small chemical modification of the target residues (addition of 42 and 14 Da, respectively). Ubiquitin is a small 8.5 kDa protein that can be conjugated to lysine residues. Because ubiquitin in itself is not reactive, ubiquitination is catalyzed in a three-step process. First, Ubiquitin gets bound by one of two ubiquitin E1 activating enzymes, which are very general ubiquitin binding proteins and do not show any target specificity. In an ATP-dependent step ubiquitin is transferred from the E1 to a more target specific E2 ubiquitin conjugating enzyme. Finally, the E3 enzymes catalyze the transfer of ubiquitin to specific target lysine. Ubiquitin itself contains seven different lysines (K6, K11, K27, K29, K33, K48 and K63) that again can be modified by ubiquitination, resulting in the possibility to generate poly-ubiquitin chains leading to various cellular responses (127). Examples include cell fate decisions like apoptosis (128), cell cycle progression (129), differentiation (130) and development (131) and misregulation can lead to tumorigenesis (132). The best-known function of ubiquitin conjugation is to target modified proteins for degradation by the proteasome, but ubiquitination can also alter protein activity. Like methylation, histone ubiquitination effects are dependent on the amount of ubiquitin moieties and the location of the modified residue. H2AK119ub1 by PRC1 results in restraining of RNA pol II (133), while mono-ubiquitination of H2AK13 or H2AK15 by RNF168 is involved in the DNA damage response signaling (134,135). Conversely, H2BK120 is ubiquitinated by RNF20/40 E3 ligase complex using UBCH6 as E2 (136,137). H2BK120ub1 is a hallmark of RNA pol II transcribed genes and prerequisite for the euchromatic H3K4 and H3K79 methylation marks (138-140).

Removal of ubiquitin from modified residues is catalyzed by deubiquitinating enzymes (DUBs). The 95 DUBs are classified in five different groups, based on the structure of their catalytic domain. Ubiquitin-specific proteases (USPs), otubain domain-containing proteins (OTUs), Machado-Joseph domain (Josephin domain)-containing proteins (MJDs), Ubiquitin C-terminal hydrolases (UCHs) are cysteine proteases and the jab1/mov34/mpn (JAMM) domain DUBs are Zn2+ metalloproteases (141,142). The smallest class are the UCH

DUBs consisting of Ubiquitin carboxyl-terminal hydrolase isozyme L1 (UCHL1), UCHL3, UCHL5 and the BRCA1-associated protein 1 (BAP1) (142). BAP1 is the catalytic enzyme of the PR-DUB complex that deubiquitinates H2AK119ub1. The intrinsic catalytic activity of BAP1 for its ubiquitinated substrate is low. However BAP1 activity is greatly enhanced when BAP1 is interacting with the additional sex comb like 1-3 (ASXL1-3) proteins. In PR-DUB, the ASXL proteins are mutual exclusive, generating specific PR-DUB.1-3 subcomplexes. These subcomplexes then interact with YY1 for localization of the complex and submodules either containing OGT and HCFC1 or FOXK1/2, MBD5/6, KDM1B and KDM6A (114,143,144). Yeast H2BK123ub1 is catalyzed by the DUB module of SAGA (35,145). This module consists of Ubp8, Sgf11, Sus11 and Sgf73 and mainly contacts H2A and H2B (146). Several DUBs of the human USP family have been identified that can deubiquitinate H2Bub1. USP3/7/12/22/44/46 and 49 show dub activity towards H2Bub1 (130,147-151), of which USP22 is part of the human SAGA complex (150).

Switching from transcriptional control to protein modulation

The histone modifications and interplay by histone modifying enzyme complexes allow for extensive cellular control of transcription and are important tools for controlling cellular molecular processes. Once transcription followed by cytoplasmic protein synthesis is established, the cell can modulate protein activity, binding partners and location to maintain molecular control. Physical separation of enzyme and substrate by sequestering the enzyme in a complex or controlled microenvironment allows control over catalysis. Furthermore, cells are able to quickly respond when these molecular processes are needed by simply relocating the separated proteins to their substrates instead of relying on *de novo* synthesis of the required proteins.

Protein nuclear transport

In eukaryotic cells, the nuclear envelope enclosing the cell nucleus provides physical separation of chromatin and other nuclear factors from the cytoplasm, providing a controlled microenvironment. Nuclear pore complexes located in the nuclear envelope surrounding the nucleus allow for passive diffusion of small molecules and globular proteins up to approximately 50 kDa that can freely move between cytoplasm and nucleus. Furthermore, the nuclear pore controls active transport of larger macromolecules. Many proteins functioning in the nucleus require active transport to reach their targets, either as a control mechanism or after cytoplasmic protein synthesis. Conversely, nuclear acting proteins can be exported to physically separate them from their substrate. This active protein nuclear transport is driven by the small GTPase Ran, which is part of the RAS superfamily of GTPases (152). Small GTPases can be either bound to GDP or GTP and function as a molecular switch (153,154). When GTPases are bound to GTP the GTPase is in an active state. GTPase-activating proteins (GAPs) promote the hydrolysis of GTPase-bound GTP into GDP + Pi which causes the GTPase to become inactive. Reactivation of small GTPases is promoted by guanine exchange factors (GEFs) that stimulate GDP to GTP exchange.

The mammalian karyopherin family of proteins consists of 20 proteins that interact and mediate the transport of cargo proteins between the cytoplasm and the nucleus (155,156). Karyopherins that recognize a nuclear localization signal (NLS) transport their cargo into the nucleus are classified as importins. Conversely, exportins recognize nuclear export signals (NES) and transport their cargo out of the nucleus. The process of karyopherin cycling between the cytoplasm and the nucleus is driven by the energy that is provided

during the Ran cycling (157,158) (Fig. 2).

In the cytoplasm, cargo proteins are bound by a heterodimeric importin-α/β complex or by single importins. Importin-α proteins bind the NLS of the cargo and bind to importin-β that in turn can interact with the nuclear pore (159). The cargo-importin complex then migrates together with GDP bound Ran (RanGDP) into the nucleus. Nuclear transport factor 2 (NTF2) has a high affinity for the cytoplasmic RanGDP and causes Ran to be concentrated in the nucleus (160,161). Once present in the nucleus, regulator of chromosome condensation (RCC1) act as a GEF for Ran promoting exchange of GDP for GTP (162,163). This results in a gradient of the GTP bound form of Ran (RanGTP) in the nucleus that is approximately 100-fold greater than the cytoplasm (164). The interaction with RanGTP causes a conformational change in the importin proteins, resulting in release of their cargo in the nucleus and for importin-β to form a new complex with RanGTP. The released importin-α then interacts with exportin-2 and enters the exportin cycle (see below) (165). The importin-RanGTP complex then migrates back through the nuclear pore to the cytoplasm. Interaction of the importin-RanGTP complex with Ran binding protein 1 (RanBP1), RanBP2 and RanGAP1 stimulates the hydrolysis of GTP, followed by dissociation of importin that resets the transport cycle (163,166).

Nuclear export of proteins follows a route that is very similar to protein import. Cargo residing in the nucleus that contains a NES becomes bound by exportin proteins and RanGTP. The heterotrimeric complex then transfers through the nuclear pore into the cytoplasm where RanGAP promotes GTP hydrolysis and the exportin is released. The exportin protein then independently migrates back to the nucleus to bind a new substrate.

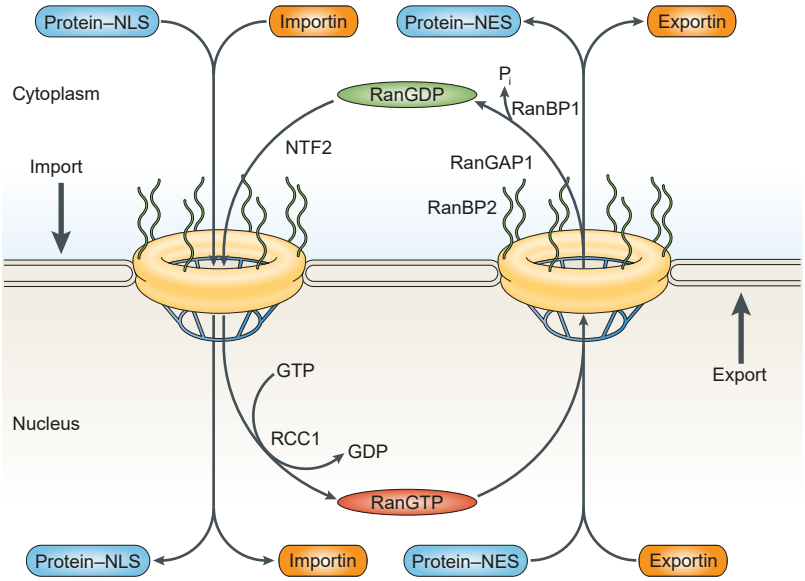


Figure 2. Schematic representation of Ran cycle dependent nuclear protein import and export. Ran cycles between GDP and GTP bound state between cytoplasm and nucleus, facilitated by RCC1, RanGAP1, RanBP1, RanBP2 and NTF2. Importin (like IPO5 and IPO7) and exportin (like CRM1) proteins can bind proteins containing an NLS and NES, respectively and transport their cargo through the nuclear pore in a Ran-dependent manner. Image adapted from (277).

Cells need to be able to respond quickly to cues from its environment to alter target gene transcription. Signals in the form of ligands present in the extracellular space can bind cell surface receptors that initiate protein-mediated signaling cascades. A molecular response can be initiated by signaling proteins facilitated by nuclear transport that enter the nucleus and promote transcription of target genes. One such classical example is the SMA and MAD related (SMAD)-mediated Bone Morphogenetic Protein (BMP) and Transforming Growth Factor-β (TGF-β) protein signaling cascade.

TGF-β and SMADs

Cellular signaling via cytokines is vital for proper functioning of cells throughout the entire organismal development and its lifespan. TGF-β signaling affects stem cell maintenance, cell proliferation, differentiation, migration and cell death and is evolutionarily conserved (167). The TGF-β protein superfamily members are secreted factors that allow these signaling events to occur. The first described TGF-β protein is TGF-β1 (168) and since then, multiple proteins in different subclasses are described. These subclasses include TGF-β, Activin, Inhibin, Anti-Müllerian Hormone (AMH), BMP, Nodal and Growth Differentiation Factors (GDFs) (169) and is based on structural similarities of these proteins. TGF-β signaling is mediated through a signaling cascade mediated by the SMAD protein family. SMAD proteins are the human homologues of the *Caenorhabditis elegans* protein SMA (Small) (170) and the *Drosophila melanogaster* protein mothers against decapentaplegic (MAD) (171). SMAD proteins are gene-specific transcription factors (GSTFs) consisting of eight members, which are divided into three separate groups: receptor activated SMADs (R-SMADs, SMAD1/2/3/5/9), the co-activator SMAD (Co-SMAD, SMAD4) and inhibitory SMADs (I-SMADs, SMAD6/7). R-SMADs are subdivided based on their activating ligands into SMAD1/5/9 for BMP and GDF activation or SMAD2/3 for TGF-β, activin and NODAL activation (172). All eight SMADs consist of an N-terminally located MAD homology (MH1) domain that can recognize sequence-specific DNA elements. The MH1 is coupled by a proline-rich linker to a C-terminal MH2 domain that is required for SMAD-SMAD protein binding. Furthermore, all R-SMADs contain C-terminal SSXS residues, which are critical for R-SMAD phosphorylation resulting in activation of these R-SMADs during canonical SMAD-mediated TGF-β signaling. The co-activator SMAD4 is the only SMAD that contains a SMAD activating domain (SAD), which is required for target gene activation.

Canonical SMAD-mediated signaling and transcription regulation

SMAD signaling starts with the extracellular binding of a ligand to one of five different Activin receptor-like kinase (ALK) type II serine/threonine receptor kinase pairs (Fig. 3). Activation of the type II receptors results in intracellular phosphorylation of one of seven different ALK type I receptor pairs. The ligand-bound activated receptors then internalize into endocytic vesicles and are transported to early endosomes (173,174). The endosome-associated FYVE domain-containing proteins Endofin (also known as ZFYVE16) and SMAD anchor for receptor activation (SARA) are protein scaffolds for BMP and TGF-β R-SMADs respectively (174-176). These proteins require their FYVE domain to bind to phosphatidylinositol 3-phosphate (PtdIns(3)P) for localization to early endosomes and assist to position the R-SMADs in close proximity and directionality to the activated type I receptors (174). This allows the ALK type I receptors to phosphorylate the R-SMADs on the SSXS residues that results in a conformational change in the phosphorylated R-SMAD and dissociation from the protein scaffold (177). The activated R-SMADs then form a hetero-

trimeric complex consisting of two R-SMADs and one copy of SMAD4 (178). SMADs require active protein transport between cytoplasm and nucleus for proper functioning that is mediated by specific nucleoporin proteins (159). An NLS sequence (KKLKK) that is recognized by importin proteins IPO7, IPO8 and Importin-β is present in the R-SMADs MH1 domain, but is absent in SMAD4 and I-SMADs (179-182). SMAD4 contains a different bipartite NLS in its MH1 domain that mediates association with Importin-α (183). A NES sequence is also present in the R-SMADs and SMAD4 that is required for recognition and nuclear export by CRM1. The presence of both the NLS and NES sequences in the R-SMADs and SMAD4 results in constant shuttling between cytoplasm and nucleus. R-SMADs show a steady-state pattern under basal conditions, whereby the R-SMADs are mostly cytoplasmic located. SMAD4 requires binding to an activated R-SMAD to find its DNA target the nucleus. This binding event results in sequestering of the SMAD4 NLS and nuclear accumulation of the R-SMADs/Co-SMAD complex (184).

Once present in the nucleus, the R-SMADs/Co-SMAD complex can bind to specific DNA sequences via an eleven-residue β-hairpin fold that is embedded in the major groove of the DNA (185-187). SMAD3 and SMAD4 can bind the palindromic GTCTAGAC sequence (188). This sequence is composed of two GTCT SMAD binding elements (SBEs), whereby one SMAD molecule binds to one SBE (185). In SMAD2/3 mediated TGF-β signaling SMAD3 and SMAD4 bind to these SBEs while SMAD2 is not able to bind DNA because of an encoded insert that blocks DNA binding via the β-hairpin (185,189). SMAD1/5/4 complexes recognize GRCGNC-N5-GTCT sequences where N is any nucleotide, R is A or G and V is A, C or G (190-192). The best-defined GRCGNC DNA binding site for SMAD1 and SMAD5 is GGCGCC and the GTCT site is the SMAD4 SBE binding site and is often found

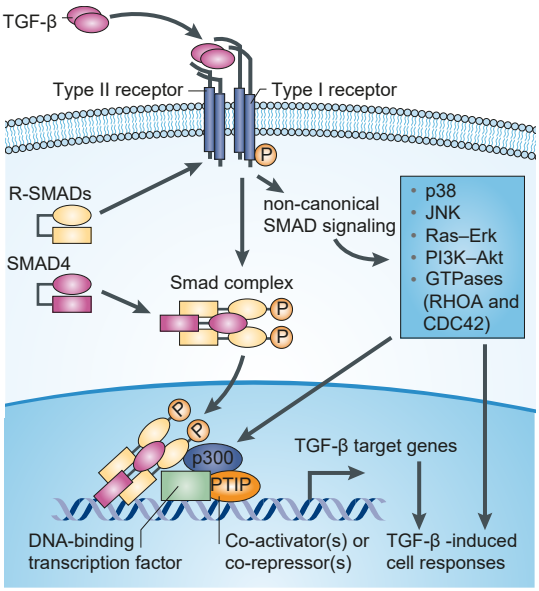


Figure 3. Schematic representation of SMAD-mediated TGF-β signaling. TGF-β binds and activates type II and type I receptors resulting in canonical and non-canonical SMAD signaling. Canonical SMAD signaling results in phosphorylation and hetero-trimerization of R-SMADs and SMAD4. This complex accumulates in the nucleus and activates target gene transcription with assistance of other transcription factors. Image adapted from (204).

in BMP target genes (193-196). In general, SMADs possess a low DNA-binding affinity and binding require interaction with various other cell type-specific transcription factors like AP-1 (197), ETS (198), C/EBP (199), FoxH1 (200,201), FoxO (202), MYOD1 (203), OCT4 (203) and PU1 (203). Interaction with these transcription factors enables activated SMAD complexes to target and regulate specific target genes like plasminogen activator inhibitor-1 (PAI-I, also known as Serpine1), SMAD6, SMAD7 and inhibitor of differentiation-1 (ID1) (204-208). Loss of these cell type-specific transcription factors results in reduced TGF-β responsiveness on target genes and loss of SMAD binding (203). Once bound to the target genes, SMAD complexes recruit coactivators like p300/CBP or corepressors such as ATF-3 to initiate or repress target gene transcription (209,210). The requirement of a cell type- and lineage-specific transcription factor is well established and these cofactors can affect SMAD signaling in a variety of ways. Association of SMAD2/4 with *Xenopus* Foxh1 stimulates the transcription of the *Mixer* gene (211). Mixer also associates with SMAD2/4 and enables transcription of the *Goosecoid* genes via enhancer binding. Conversely, SMAD3/4 binding to ATF3 results in SMAD3/4/ATF3 mediated *ID1* repression (210). Another type of activation is via recruitment of histone modifying enzymes. SMAD2/3 was shown in human embryonic stem cells (hESCs) to bind H3K27 demethylase KDM6B/JMJD3, which is recruited to the SMAD2/3 binding sites upon Activin signaling (212). This reduces H3K27 methylation levels resulting in a transcriptionally permissive chromatin state and bivalent marked gene activation. Transcriptional repression is achieved by recruitment of HDACs and subsequent deacetylation of histone residues, like the recruitment of HDAC4/5 by SMAD3 (213). Other mechanisms of SMAD repression include TEAD/SMAD/OCT4 (TSO) complex mediated recruitment of the Nucleosome Remodeling and Deacetylase (NuRD) complex resulting in repression of pluripotency genes (214).

SMAD regulation

The SMAD-mediated signaling cascade responds to extracellular stimuli by modulating gene expression. Amongst the upregulated target genes of SMAD signaling are the I-SMADs that negatively regulate the SMAD signaling cascade (215-218). BMP R-SMADs induce expression of SMAD6 while TGF-β R-SMADs induce expression of SMAD7. I-SMADs compete with the R-SMADs for binding with the type I receptors (219,220). SMAD6 also inhibits SMAD signaling by competing for binding to SMAD4 (221). Additionally, I-SMADs interact with the E3 ubiquitin ligases SMURF1 and SMURF2 which ubiquitinate the activated type I receptors (222-224). Ubiquitination of the receptors then results in proteasomal degradation of the receptor. Alternative proteins that modulate SMAD signaling include the SKI (also known as SNO), SKI-like (SKIL, also known as SNON) and MAN1 (also known as LEMD3) proteins (225-228). SKI and SKIL genes are, like I-SMADs, upregulated upon SMAD signaling. SKI and SKIL can bind to active R-SMAD complexes and recruit the repressive NCOR complex to inhibit SMAD2/3/4 mediated transcription (229-231). However, TGF-β signaling also results in degradation of SKI and SKIL (232,233). The ratio between SMAD2/3/4 and SKI/SKIL determines whether SKI and SKIL are degraded or will repress SMAD signaling (225,232). MAN1 is located at the inner nuclear membrane and, like SKI and SKIL, binds with R-SMADs resulting in suppression of TGF-β signaling (227,228).

SMAD signaling crosstalk

Apart from canonical TGF-β signaling, SMAD signaling has been shown to crosstalk

with other pathways. Alternative signaling pathways can induce SMAD target genes like I-SMADs that will repress canonical SMAD signaling like the interferon-γ (IFN-γ)/STAT and the nuclear factor κB (NF-κB)/RelA (234,235). Additionally, TGF-β stimulation of target receptors can activate different signaling pathways. Examples of this crosstalk include ERK activation by TGF-β, which is mediated through the GTP loading of RAS (236,237). Activated ERK is important for epithelial to mesenchymal transition (EMT) during embryonic development (238). A second described pathway is the activation of mitogen-activated protein kinases (MAPK) leading to activation of JNK and p38 (239-241). Induction of MAPK by TGF-β is shown to act together with canonical TGF-β signaling on EMT and apoptosis (242,243). A pathway acting independently of canonical TGF-β signaling is activation or RhoA, which can rapidly be activated by TGF-β activated receptors. RhoA signaling affects cytoskeleton organization and cell motility (244) and induction of TGF-β induces cell mobility and stress fiber formation (245,246). Finally, TGF-β signaling can induce PI(3)K/Akt and results in EMT and fibroblast proliferation (247-249).

SMAD signaling in cancer

TGF-β and the associated SMAD signaling proteins play a central role by acting on various important molecular pathways from controlled EMT during embryonic development to control of homeostasis during adult life (250). Loss of heterozygosity analysis of genes has shown various TGF-β signaling associated genes to be functioning as tumor suppressors (251). In cancer genome-wide association studies (GWAS), chromatin-associated proteins and GSTFs were shown to be often mutated in various solid tumor types (252-254). Canonical non-canonical TGF-β signaling like MAPK and PI(3)K pathways are often mutated in colo-rectal carcinomas and other tumor types. Colon and rectal carcinomas contain mutations in the type I and II receptors like TGF-β receptor type 2 (TGFBR2, 2.6%), Activin receptor type 1B (ACVR1B, 3.6%) and ACVR2A (2.6%). Additionally, colon and rectal carcinomas contain *SMAD2* and *SMAD4* mutations in 5.7 and 9.8% of the cases, respectively (252). Generally, these TGF-β signaling components are often down-regulated and loss of these proteins correlates with tumor grade, metastasis and poor prognosis (255,256). Conversely, negative regulators of TGF-β signaling like SMAD7, SMURF2, SKI and SKIL are often upregulated. In contrast to signaling pathway downregulation, the ligand TGF-β1 is often found up-regulated in colon, gastric and prostate cancer and correlates with metastasis and disease progression indicating that TGF-β can also act as a tumor promoter (257). Tumors with high TGF-β expression correlate with angiogenesis and metastasized tumors show greater TGF-β expression (258-260) suggesting TGF-β might be required for induction of EMT (245,256,261). The complex interplay of TGF-β signaling and its importance in cancer underscores the need to understand the regulation of this signaling pathway.

Studying (SMAD) regulation

In molecular biology, many different tools are available to study SMAD mediated TGF-β signaling. Over the last decades extensive molecular biology techniques have been developed to study protein expression, localization and function. One of the tools that was critical for studying protein localization was the utilization of a green fluorescent protein (GFP) identified in the jellyfish *Aequorea victoria* (262). GFP is a 238 amino acid protein that forms a beta sheet barrel carrying an internal chromophore absorbing light in the ultraviolet and blue spectrum and emitting in the lower green spectrum of visible light (263). Current versions of GFP have an excitation maximum at 488 nm and emission maximum at 509 nm

and are optimized in fluorescence and photo-stability, extinction coefficient and quantum yield (264-266). Furthermore, many different fluorescent molecules have been identified and engineered to cover the entirety of the visible light spectrum. Because of the capability to emit light, GFP and its optimized versions have been used massively in fluorescence wide-field and confocal microscopy to study protein expression and localization.

Protein interaction studies are often performed by affinity purification of a tagged recombinant protein followed by analysis by immunoblotting or by mass spectrometry. For proper protein complex purification, this approach requires very specific high affinity binding by antibodies. Conventional antibodies consist of a heavy and light chain and often bind nonspecific. Generation of camelid single-chain nanobodies against GFP allowed for efficient affinity purification of GFP-tagged proteins and identification of their interacting partners by mass spectrometry (267-271).

Once information about interaction partners is obtained, molecular tools can be used to functionally study transcription factors like the SMADs. SMADs require phosphorylation of their C-terminal tail to enter an active state that is catalyzed by their activating receptors. Molecules can be used as tools to artificially modulate protein catalytic activity. For example, SMAD signaling can be induced by addition of recombinant TGF-β1 to tissue culture medium. Conversely, small molecule inhibitors like LDN-193189 or SB-431542 inhibit the BMP and TGF-β receptors, thereby preventing activation of SMAD signaling (272-274). Once a small molecule inhibitor and its effects are established, clinical trials are performed to test the compounds in the clinic as a potential new medicine.

Thesis outline

Transcriptional and post-transcriptional protein control mechanisms are vital for proper cellular functioning and the misregulation of these processes can lead to improper cell function and disease. This thesis aims to investigate various control mechanisms.

Concerning gene expression, cells use PTMs on histone proteins to influence the chromatin environment and thereby exert control at a transcriptional level. **Chapter 2** describes Chromatin Regulation Ontology SiRNA Screening (CROSS), a microscopy-based high-throughput screening method that allows for identification of proteins influencing histone PTM levels. By combining siRNA knockdown with automated antibody-based staining and automated microscopy we show CROSS can be used to identify writers and erasers of chromatin marks.

The spatiotemporal distribution of proteins and their respective targets determines if these proteins can interact with these targets. GSTFs act on DNA located in the cell nucleus. By constraining these GSTFs in the cytoplasm the activity of these proteins can be controlled by modulation of their nuclear import and export. SMAD proteins are GSTFs that in inactive state are localized in the cytoplasm. Upon activation these proteins migrate to the nucleus via active transport. **Chapter 3** describes affinity purification coupled with mass spectrometry to identify binding partners of SMAD transcription factors. IPO5 is identified as a *de novo* binding partner of BMP-activated SMAD1. Using confocal microscopy we show that IPO5 is a BMP-activated SMAD specific transporter and that the lysine stretch found in the NLS of the R-SMADs is a determinant for IPO specificity.

Upon activation, GSTFs localize to specific DNA elements found at enhancers or promoters near target genes. Upon binding, these GSTFs can recruit multi-meric protein complexes to modulate transcription of the target genes. **Chapter 4** describes the involvement of the MLL4 methyltransferase complex in TGF-β induced SMAD signaling. Direct interaction of MLL3/4 subunits PA1 and PTIP is shown with SMAD2/3/4. Expression analysis of TGF-β target genes shows that target gene activation is dependent on the MLL4 complex.

In **addendum 1** we sidetrack by exploring the interactome composition of the MLL3/MLL4 complex subunit UTX. By mass spectrometric analysis we confirm the interaction with MLL3/4 and PR-DUB. Additionally, we identify UTX to be part of the MiDAC complex.

Overall, cellular control over molecular processes regulates proper functioning of the cell. Here, we investigate this complex network of control pathways at various stages in the cell. Proper understanding of these mechanisms helps to understand how cells maintain a healthy environment and how diseases influence these mechanisms.

Chapter 2

A novel microscopy-based high-throughput screening method to identify proteins that regulate global histone modification levels.

Roy Baas[1], Daphne Lelieveld[2], Hetty A.A.M. van Teeffelen[1], Philip Lijnzaad[1], Bas Castelijns1, Frederik M.A. van Schaik[1], Michiel Vermeulen[1,3], David A. Egan[2], H.Th. Marc Timmers[1] and Petra de Graaf[1]

Adapted from:
J Biomol Screen. 2014 19, 287-296
doi: 10.1177/1087057113515024

[1]Department of Molecular Cancer Research, University Medical Centre, Utrecht, The Netherlands
[2]Cell Screening Center, Department of Cell Biology, University Medical Centre, Utrecht, The Netherlands
[3]Department of Medical Oncology, University Medical Centre, Utrecht, The Netherlands

Abstract

Posttranslational modifications of histones play an important role in the regulation of gene expression and chromatin structure in eukaryotes. The balance between chromatin factors depositing (writers) and removing (erasers) histone marks regulates the steady-state levels of chromatin modifications. Here we describe a novel microscopy-based screening method to identify proteins that regulate histone modification levels in a high-throughput fashion. We named our method CROSS, for Chromatin Regulation Ontology SiRNA Screening. CROSS is based on an siRNA library targeting the expression of 529 proteins involved in chromatin regulation. As a proof of principle, we used CROSS to identify chromatin factors involved in histone H3 methylation on either lysine-4 or lysine-27. Furthermore, we show that CROSS can be used to identify chromatin factors that affect growth in cancer cell lines. Taken together, CROSS is a powerful method to identify the writers and erasers of novel and known chromatin marks and facilitates the identification of drugs targeting epigenetic modifications.

Introduction

The accessibility of the eukaryotic genome is compromised by the dense compaction of DNA in the form of chromatin. The nucleosome represents the repeating unit of chromatin and consists of two copies of each of the canonical histones (H3, H4, H2A, and H2B) around which ~146 bp of genomic DNA is wrapped. The molecular structure of the nucleosome revealed the protrusion of the N- and C-terminal tails of histones from the globular particle (18). A large body of work during the past decade has revealed that these accessible parts of the core histones can be modified by specific enzymes. These modifications include the acetylation and methylation of lysine residues, which are mediated by lysine acetyltransferases (KATs) and lysine methyltransferases (KMTs), respectively. Enzymes that deposit histone modifications are commonly referred to as chromatin "writers" (19). Although often referred to as epigenetic marks, all types of histone modifications are biochemically reversible (278). Histone acetylation is removed by histone deacetylases (HDACs and sirtuins), whereas histone methylation is erased by lysine demethylases (KDMs). These enzymes are classified as chromatin "erasers." It has become increasingly clear that mutations in chromatin-modifying enzymes are important drivers of tumorigenesis (279,280). Together, these findings also indicate that the modification state of chromatin is highly dynamic and results from the interplay between enzymes depositing and removing chemical groups. Furthermore, disturbance of the epigenetic landscape can severely affect cell fate and homeostasis.

A third class of chromatin proteins is the so-called readers. A chromatin reader is defined as a protein bearing a domain that can specifically recognize a particular histone modification. Examples of such reader domains include bromodomains, which bind to acetylated lysines, and PHD fingers, which "read" methylated lysines. Other described readers include Chromo, Tudor, MBT, and PWWP domains, each of which displays affinity for a particular histone modification (278,281-283). Chromatin-binding protein complexes have often evolved to contain multiple reader domains for different modifications. Histone modifications co-occurring on active or repressive chromatin can be recognized by readers in a combinatorial manner, thus giving rise to a high affinity interaction driven by multiple relatively low-affinity single interactions. This also allows for a sophisticated fine-tuning of interactions rather than simple binary kinetics. Together, this illustrates the combinatorial nature or "crosstalk" of histone modifications (284). An example is the binding of TFIID to chromatin. The TFIID subunit TAF3 reads trimethylated histone H3 lysine-4 (H3K4me3) via a PHD finger in its C-terminus (68). Binding of TFIID to chromatin is further enhanced by acetylation of histone H3. These marks are recognized by the double bromodomain of TAF1 (68). Deregulation of chromatin readers can also result in cancer as exemplified by the oncogenic fusion of the nuclear pore complex subunit NUP98 to the C-terminal PHD domain of the KDM5A/JARID1A demethylase for H3K4 or the PHD domain of PHF23 protein. This fusion of NUP98 to a PHD domain is one of the most prevalent gene rearrangements found in various forms of hematopoietic malignancies (279). The oncogenic potential of the fusion protein relies on its ability to bind to H3K4me2/3-modified nucleosomes, preventing demethylation and thereby inactivation of developmentally crucial genes (279).

Efforts to determine which histone modifications co-occur on chromatin and therefore function in chromatin signaling pathways have revealed crosstalk between distinct histone modifications (278). Relevant in this respect are seminal contributions made

by GPS (global proteomic analysis in *Saccharomyces cerevisiae*) developed by Schneider and colleagues (285). GPS relies on screening protein extracts of the yeast deletome library (~4800 viable single-gene knockout strains) by immunoblotting with histone modification-specific antibodies. This unbiased approach revealed that the PAF1 elongation complex and H2B mono-ubiquitylation by Rad6/Bre1 are upstream of H3K4 methylation by the Set1/COMPASS complex (286). The GPS screening method involved a significant amount of experimental work (~4800 lysates on ~300 separate blots per antibody tested). Coverage of mammalian genomes would increase the effort by a factor of four.

Here we describe a novel approach based on high-throughput microscopy to investigate histone modification pathways in human cells. We generated a small interfering RNA (siRNA) library targeting the expression of 529 genes, including annotated writers, erasers, readers, and cofactors of chromatin-modifying complexes. Following siRNA-mediated knockdown in a 96-well format, quantitative immunofluorescence microscopy with histone modification-specific antibodies is used to determine changes in global histone modification levels. We name this method CROSS for Chromatin Regulation Ontology SiRNA Screening. As a proof of principle, we used CROSS to screen H3K4 and H3K27 methylation in U2OS and MCF7 cells. In both cell lines, we identified the writers and erasers for these marks. We also used CROSS to identify chromatin factors that affect cell growth, which in the future could provide novel candidates for the development of drugs targeting epigenetic modifications in relation to cancer.

Materials and Methods

Antibodies

Antibodies used are listed here: H3 (ab1791; Abcam, Cambridge, UK), H3K4me1 (pAb037-050; Diagenode, Liège, Belgium), H3K4me2 (pAb035-050; Diagenode), H3K4me3 (pAb003-050 [Diagenode] or 04-745 [Millipore, Billerica, MA]), H3K27me1 (pAb045-050; Diagenode), H3K27me3 (pAb069-050; Diagenode), H3K27ac (pAb174-050; Diagenode), H3K9me1 (ab8896; Abcam), H3K9me3 (ab8898; Abcam), ASH2L (kind gift from Winship Herr), RBBP5 (BL766; Bethyl Laboratories, Montgomery, TX), GAPDH (clone 6C5, mAb374; Millipore), and EZH2 (clone AC22, #3147; Cell Signaling Technology, Danvers, MA).

Cell Lines

Osteosarcoma cells (U2OS) and breast carcinoma cells (MCF7) were obtained from the ATCC (American Type Culture Collection, Manassas, VA) and cultured in Dulbecco's modified Eagle's medium (DMEM; Lonza, Verviers, Belgium) supplemented with 10% fetal bovine serum (Lonza) and antibiotics (100 U/mL penicillin and 100 µg/mL streptomycin, Lonza) at 37 °C, 5% CO_2 in a humidified atmosphere.

siRNA Library

Based on the chromatin library described in Mulder et al. (287), we included complex members of chromatin modifiers and binders. The chromatin library, targeting expression of 529 genes, was purchased from Thermo Fisher Scientific Dharmacon (Lafayette, CO, custom-made library, order numbers 245120-245146, 245170, and 246203). A full gene list is given in the supplemental data.

siRNA Transfection

Cells were transfected in a reverse transfection protocol using Hiperfect (Qiagen, Valencia, CA) according to the manufacturer's instructions. In short: 5 µL of a SMARTpool of four siRNAs (25 nM each) from the chromatin library was dispensed in 96-wells plates (Costar, 3599; Corning, New York, NY) (triplicates for each screen). Transfection mix (0.4 µL Hiperfect in total volume of 15 µL Opti-MEM [Invitrogen, Carlsbad, CA] per well) was added using the multidrop dispenser (Multidrop combi Type 836; Thermo Scientific, Waltham, MA), followed by a 30-min incubation at room temperature (RT). Meanwhile, cells were counted and 3000 cells/well were added (in 80 µL). Cells were fixed 72 h after transfection.

Fluorescent Microscopy

Depending on the antibody, cells were fixed using formaldehyde or cold methanol. In case of formaldehyde, cells were fixed for 20 min at RT with 4% formaldehyde in phosphate-buffered saline (PBS), followed by a 30-min permeabilization with 0.5% Triton X-100 in PBS and a 10-min incubation with 50 mM glycine in PBS. In case of methanol, cells were fixed and permeabilized by a 5-min incubation with cold methanol, after removal cells were rehydrated by washing twice with PBS, followed by washing once with PBS/0.1% Triton X-100. Cells were blocked using 5% NGS in PBS/0.1% Triton X-100, followed by incubation with appropriate antibodies diluted in blocking solution. After antibody staining, cells were incubated with DAPI solution (2 mg/L 4',6-diamidino-2-phenylindole [Sigma-Aldrich, St. Louis, MO] in 50 mM Tris-HCl [pH 7.4], 100 mM NaCl).

Antibody Quality Control

Antigen specificity was tested using in-house synthesized biotin-labeled peptides, both in immunofluorescence microscopy competition assays as well as on immunoblots with the appropriate peptides. For competition assays, peptides were added in a 2-µM concentration to the primary antibody solution. To prepare immunoblots, peptides were bound to streptavidin, run on a 10% SDS-PAGE, and transferred onto PVDF membranes. The membrane was developed with the appropriate antibodies and ECL (Pierce, Thermo Fisher Scientific Inc., Rockford, IL).

CROSS Data Analysis

Average nucleus fluorescence intensity values and cell count per field were extracted from the microscopic images using Cellomics vHCS: View 1.6.1.613 (Thermo Fisher Scientific Inc., Rockford, IL). Robust *z* scores were calculated using R version 2.15.1 (2012-06-22) for all variables per plate for all chromatin library siRNAs using equation (1):

$$z_{i,robust} = \frac{x_i - median\,(x_n)}{\sigma_{MAD}\,(x_n)},$$

whereby $\sigma_{MAD}\,(x_n) = c \cdot median \mid x_i - median\,(x_n) \mid$ and $c \approx 1.4826$.

CROSS robust *z* score data sets of the same residues were merged. Genes that were significantly up- or downregulating global histone modification levels ($z < -1.96$ or $z > 1.96$ unless otherwise noted) in two of three replicates in at least one screen were selected. To exclude effects caused by cell death or growth effects, we excluded genes showing a significant

reduction in objects per field. When necessary, known associated genes were added to the list (WDR5, RBBP5, ASH2L, and DPY30 for H3K4 and EZH2, EED, SUZ12, and KDM6A for H3K27).

Clustering of Growth Affecting Genes

CROSS H3K4 robust *z* score data sets were merged, and genes that were significantly downregulating valid objects per microscopic field ($z < -1.96$) in two of three replicates in at least one screen were selected. Clustering was performed on Euclidean distance using Ward's method.

MTT Assay

Wild-type U2OS and MCF7 cells were seeded at 1000 cells per well in 100 µL normal DMEM medium in a 96-well plate. After 24 h, medium was replaced by medium containing either I-BET inhibitor (I-BET151 at concentrations of 0.1, 0.3, 1, 3, or 10 µM), 0.1% DMSO, 2 mg/L puromycin, or medium without addition. Seventy-two hours after adding I-BET, MTT (3-(4,5-dimethylthiazolyl-2)-2, 5-diphenyltetrazoliumbromide; Sigma-Aldrich) was added to a final concentration of 0.5 mg/mL (288). Cells were incubated for 4 h under normal culturing conditions, followed by careful removal of medium, and 200 µL DMSO was added to dissolve the formazan crystals. Spectramax M5E (Molecular Devices, Sunnyvale, CA) was used to measure absorbance at 570-nm and 630-nm wavelengths. Relative growth was calculated using equation (2):

$$\frac{A_{570,\text{sample}} - A_{630,\text{sample}}}{A_{570,\text{untr}} - A_{630,\text{untr}}}.$$

Crystal Violet Staining

For long-term colony formation assays, 2000 cells/well were seeded in 12-well plates. I-BET151 was added at 0.5 and 1 µM; DMSO was used as vehicle control. Cells were incubated at 37 °C for 6 days. After the incubation period, the medium was aspirated and cells were fixed with ice cold methanol (100%) for 10 min at room temperature, with a subsequent staining with crystal violet (0.25% w/v crystal violet in 25% methanol) for 10 min. Cells were washed with water and air-dried overnight.

Results

CROSS for H3K4 Methylation Identifies Known Major Methyltransferase and Demethylase Complexes

Epigenetic modifications are affected in a variety of tumors, and these aberrant modification patterns can significantly alter gene expression patterns. We therefore decided to profile a subset of histone modifications in two distinct human cancer cell lines, MCF7 (breast carcinoma derived) and U2OS (osteosarcoma derived), using immunoblotting and immunofluorescence (Fig. 1A, B). Immunoblotting revealed comparable levels of H3K4 methylation (mono, di, and tri), as well as H3K27me1 and H3K9 methylation (mono and tri) between U2OS and MCF7 (Fig. 1A). Interestingly, H3K27me3 levels are greatly reduced in U2OS compared with MCF7, while H3K27ac is higher in U2OS. The difference in H3K27me3 levels was confirmed by immunofluorescence (Fig. 1B). EZH2 expression, however, is similar in both cell lines (Fig. 1A). This indicates that factors other than EZH2 may be responsible for the reduced H3K27me3 levels in U2OS cells compared with MCF7.

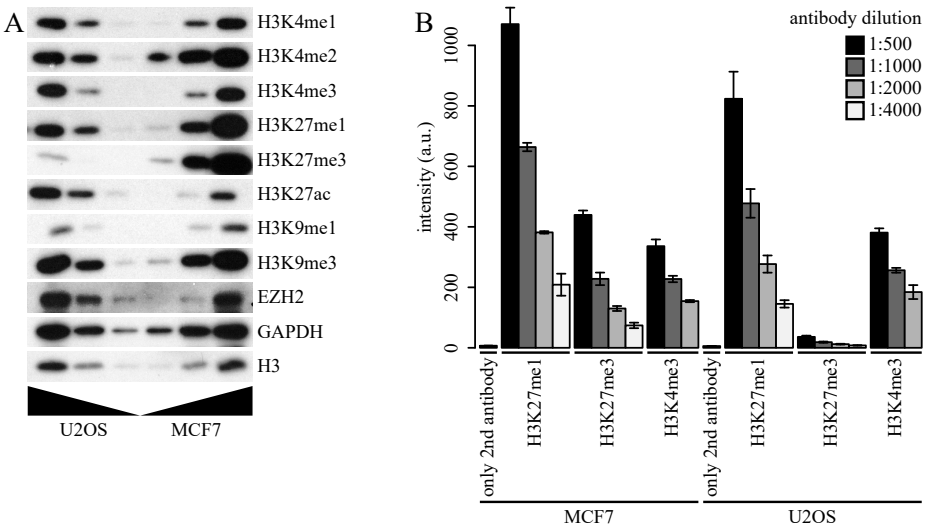
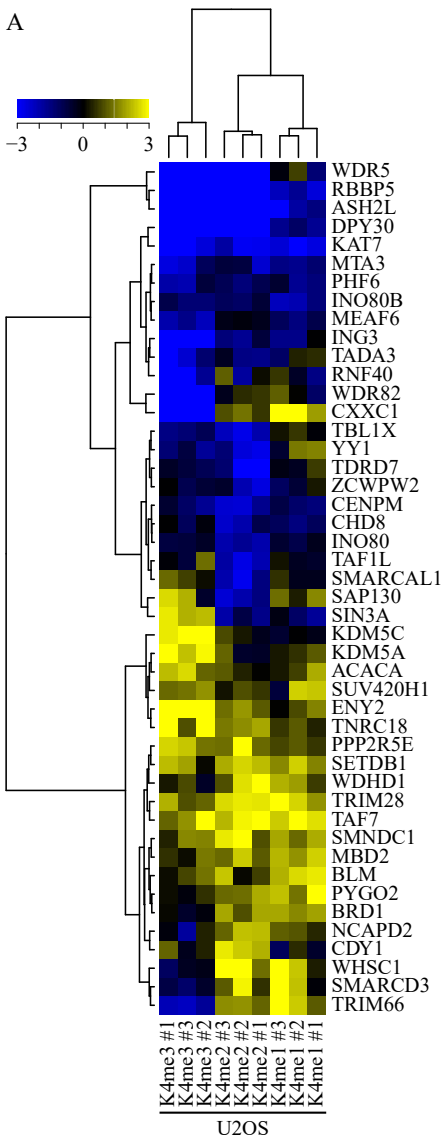


Figure 1. Global histone mark levels in U2OS and MCF7. (A) Immunoblot analysis of U2OS and MCF7 cell lysates. Immunoblots were stained by the indicated antibodies. GAPDH and H3 antibodies are used as loading controls. (B) Immunofluorescence analysis of global histone levels in U2OS and MCF7 cells. Shown are average intensity levels per nucleus of at least 3000 nuclei per condition.

To identify complexes or cofactors that influence the global levels of chromatin modifications, we combined siRNA-mediated knockdown with high-throughput immunocytochemistry and automated microscopy. We named this method CROSS for Chromatin Regulation Ontology SiRNA Screening. To this end, a custom siRNA library was designed targeting human chromatin-associated proteins, including writers, erasers, readers, and cofactors of known chromatin-modifying complexes, resulting in a gene list of 529 genes (full list in Suppl. Table S1). Each gene in our library was targeted by a SMARTpool of four individual siRNAs. Because our screening results are directly dependent on the specificity of antibodies, we performed quality control experiments for all used antibodies. These control experiments consisted of peptide competition experiments combined with immunofluorescence, as well as immunoblotting (Suppl. Figs. S1–S3).

We applied CROSS first to H3K4 methylation, a well characterized histone modification for which the enzymes, their complexes, and even regulatory pathways are known (289). In short, U2OS and MCF7 cells were reverse transfected in triplicate using the chromatin library consisting of 529 siRNAs divided over a series of 96-well plates (21 plates per antibody staining). Cells were fixed 72 h after transfection and stained using three antibodies that recognize respectively histone H3K4 mono-, di-, or trimethylation (Suppl. Fig. S2). Staining was performed using a Caliper Life Sciences (Hopkinton, MA) automated robotic liquid handling platform, which allows fully automated staining of 96-well plates. An ArrayScan VTi automated wide-field microscope (Thermo Fisher Scientific Inc., Rockford IL) was used to screen for global changes in H3K4 methylation levels by measuring average nuclear fluorescent intensity values for 1000 cells per well. The CROSS pipeline is depicted in Supplemental Figure S4. Algorithms in the ArrayScan software were used to exclude dying cells, mitotic cells, or cells with an otherwise deformed nucleus. From these averaged data, robust *z* scores were calculated per plate using an automated R script (290) to



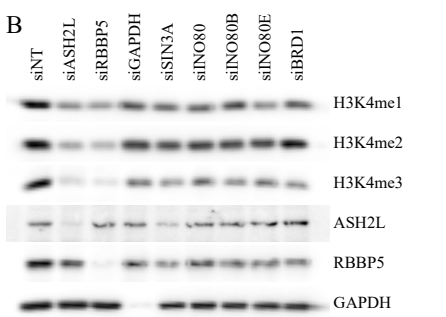


Figure 2. CROSS for H3K4 methylation in U2OS cells. (A) Hierarchical clustering of mono-, di-, and trimethylation of H3K4 in U2OS. Robust *z* scores were calculated using an R script (see Materials and Methods section). Replicates are indicated on the bottom and identified hits are on the right. Coloring (ascending from blue to yellow via black) is based on the robust *z* score for the indicated antibody signal. (B) Immunoblot analysis of global K4 methylation levels for siRNA-treated cells. Immunoblots were stained by the indicated antibodies. GAPDH is used as a loading control.

negate any variation introduced as a result of the automated staining procedure. Plotting the robust *z* values per plate shows bell-shaped curves for each replicate and a limited number of significant hits (Suppl. Figs. S5 and S6, $\alpha = 0.05$, significance borders depicted by the gray lines).

Each individual plate also contained nontargeting, ASH2L, and PLK1 siRNA controls. These controls are expected to respectively show no effect on H3K4 methylation signal (negative control, nontargeting siRNA), a reduction in H3K4 methylation signal (positive control, siRNA targeting ASH2L), and cell mitotic defects resulting in reduction of cell count and an increase of H3K4 methylation signal caused by nuclear condensation (siRNA targeting PLK1). Robust *z* scores for these controls are plotted in Supplemental Figures S5 and S6.

Occasionally, knockdown of a gene causes growth effects such as a cell cycle arrest or even apoptosis. In these cases, chromatin structure and function are severely altered, thereby influencing the readout of CROSS. Therefore, we also calculated robust *z* scores on the amount of selected cell nuclei per field. This allowed us to filter out genes that induce cell cycle arrest or apoptosis. Remaining genes that influenced global H3K4 methylation levels with a *z* score of 1.96 (corresponding to the 95% confidence interval) in either positive or negative direction in at least two of three technical replicates in at least one of the three modification screens (H3K4 mono-, di-, or tri-methyl) were selected and subjected to hierarchical clustering (Fig. 2A and Suppl. Fig. S7). Clustering analysis revealed a clear separation between MCF7 and U2OS cell lines (data not shown). The three different methylation states clustered together in U2OS, indicating reproducibility as well as distinct biology associated with each of these different modifications. As expected, negative z values were found for the known canonical methyltransferase complexes for H3K4 (SET1A/B and MLL1-4). A reduction in methylation levels was observed only for the noncatalytic core WRAD module (WDR5, RBBP5, ASH2L, and DPY30), which is shared among all SET/MLL-complexes (76). Interestingly, although efficient knockdown is achieved (Suppl. Fig. S8), the loss of methyltransferase enzymes themselves does not result in a global loss of methylation (Fig. 2A), which could be explained by (partial) redundancy or by the fact that each of these enzymes is responsible only for a small percentage of H3K4 methylation. Typical images as obtained by the automated microscopes show clear reduction of nuclear signal in case of ASH2L and DPY-30 knockdown (Suppl. Fig. S9A), whereas no difference is observed after knockdown of the MLL methyltransferases (Suppl. Fig. S9B). Four homologous proteins, KDM5A–D, mediate demethylation of H3K4. These KDM5 proteins exert their demethylase activity via the JmjC domain that preferentially removes di- and trimethylation (291). Knockdown of the H3K4 demethylases KDM5A and KDM5C (JARID1A/ RBP2 and JARID1C/SMCX, respectively) results in an increase of global H3K4me3 levels in U2OS cells (Fig. 2A and Suppl. Fig. S9C).

In addition to these expected methyltransferase and demethylase complexes, a number of other proteins that affect H3K4 methylation levels either positively or negatively were identified. Examples include subunits of the INO80 chromatin remodeling complex as well as subunits of the SIN3 histone deacetylase complex and the BRD1 bromo-domain protein. Currently, it is not clear how the known function of these proteins could affect H3K4 methylation levels, but further experiments could help to shed light on this.

As a confirmation of our CROSS results, we performed immunoblot with a selection of hits from CROSS. Knockdown of ASH2L and RBBP5 reduced all three methylation states of H3K4 (Fig. 2B). However, no effect on methylation levels detected by immunoblotting was obvious after knockdown of INO80 complex, SIN3A, or BRD1 (Fig. 2B and Suppl. Fig. S10). This strengthens the CROSS procedure: first, the dynamic range of measuring methylation levels is larger compared with immunoblotting (Suppl. Fig. S10C), and second, averaging signals of 1000 cell nuclei using automated microscopy results in accurate and quantitative data in a fast and reliable manner.

Cell-Type Specific K27 Methylation Patterns and Regulation in U2OS and MCF7 Cells

Using CROSS, genes affecting global levels of histone modifications can be identified. To measure these changes, a histone modification of interest should be of reasonable abundance under normal culturing conditions. As H3K27me3 is of low abundance in U2OS cells (Fig. 1A,B), we applied CROSS for H3K27me3 only in MCF7 cells, whereas CROSS was applied for H3K27me1 in both cell lines (Fig. 3). As expected, knockdown of PRC2 complex subunits EZH2, SUZ12, and EED results in a reduction of global H3K27me3 levels in MCF7 cells, confirming the ability of CROSS to reveal the enzyme(s) and associated factors that regulate particular histone modification(s). Interestingly, knockdown of PRC2 subunits has different effects on H3K27me1 in MCF7 and U2OS cells. While depletion of PRC2 results in a downregulation of H3K27me1 in U2OS cells, this effect is not observed in MCF7 cells. This may imply that in MCF7 cells, H3K27me1 may be (partially) catalyzed by a PRC2-unrelated methyltransferase. Knockdown of the known H3K27 demethylase KDM6A (aka UTX) showed an overall increase in H3K27 methylation levels in MCF7 but not U2OS, again showing the potential of CROSS to identify novel epigenetic regulators.

CROSS Can Reveal Cell Line–Specific Growth Effects on Gene Knockdown

Due to their general importance inside the nucleus, knocking down chromatin-associated proteins can in some cases induce a growth defect. To measure these effects in U2OS and MCF7 cells, we calculated robust *z* scores for the amount of cells that were present per optical field. Clustering analysis of these data revealed a clear separation of the two cell lines (Fig. 4A). Three clusters of genes can be identified: genes that specifically reduce cell numbers in U2OS (majority of the genes), genes leading to a reduction in cell number in MCF7, and genes that affect growth in both cell lines (PLK1, SGOL1, and ESPL1). In the cluster of genes affecting growth in MCF7, the BET proteins BRD2 and BRD4 are found. BRD4 also reduces growth in U2OS, although to a lesser extent than in MCF7. In contrast, BRD2 does not affect cell growth in U2OS (Fig. 4A), while knockdown is efficient (Suppl. Fig. S10B). The BET family of bromodomain-containing proteins binds acetylated histones to influence transcription (292). Recently, BRD4 was identified as a therapeutic target in acute myeloid leukemia (293), and BET-specific inhibitors such as JQ1 and I-BET151 were developed (294,295). We tested the I-BET151 inhibitor in MCF7 and U2OS cells to investigate a potential differential dependence on BET signaling in both cell lines. We therefore treated U2OS and MCF7 cells for 3 days with increasing concentration of I-BET inhibitor (Fig. 4B). Interestingly, a reduction of 50% in cell number was achieved at a concentration of 1 μM in MCF7, whereas this reduction in U2OS was achieved at least at a three times higher concentration, indicating that MCF7 cells are more dependent on BET function than U2OS for their growth. This effect was confirmed by a crystal violet staining (Fig. 4C). The different effects of I-BET151 on cell number in the two cell lines may be explained by the fact that in

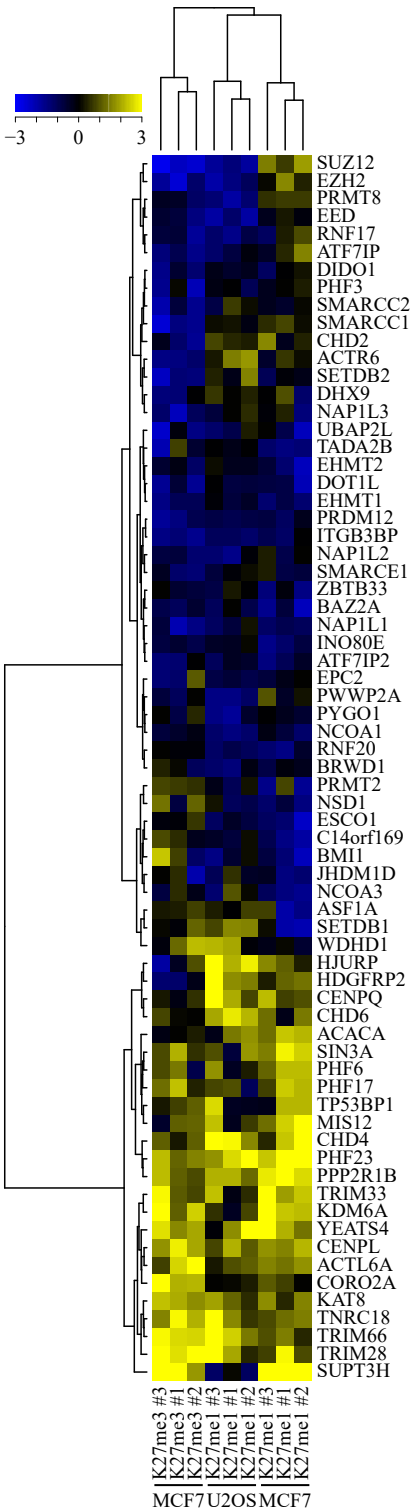


Figure 3. CROSS for H3K27 methylation in U2OS and MCF7. Hierarchical clustering of mono- and trimethylation of H3K27 in U2OS and MCF7. Robust *z* scores were calculated using an R script (see Materials and Methods section, with negative $z < -1.20$). Replicates are indicated on the bottom and identified hits are on the right. Coloring (ascending from blue to yellow via black) is based on the robust *z* score for the indicated antibody.

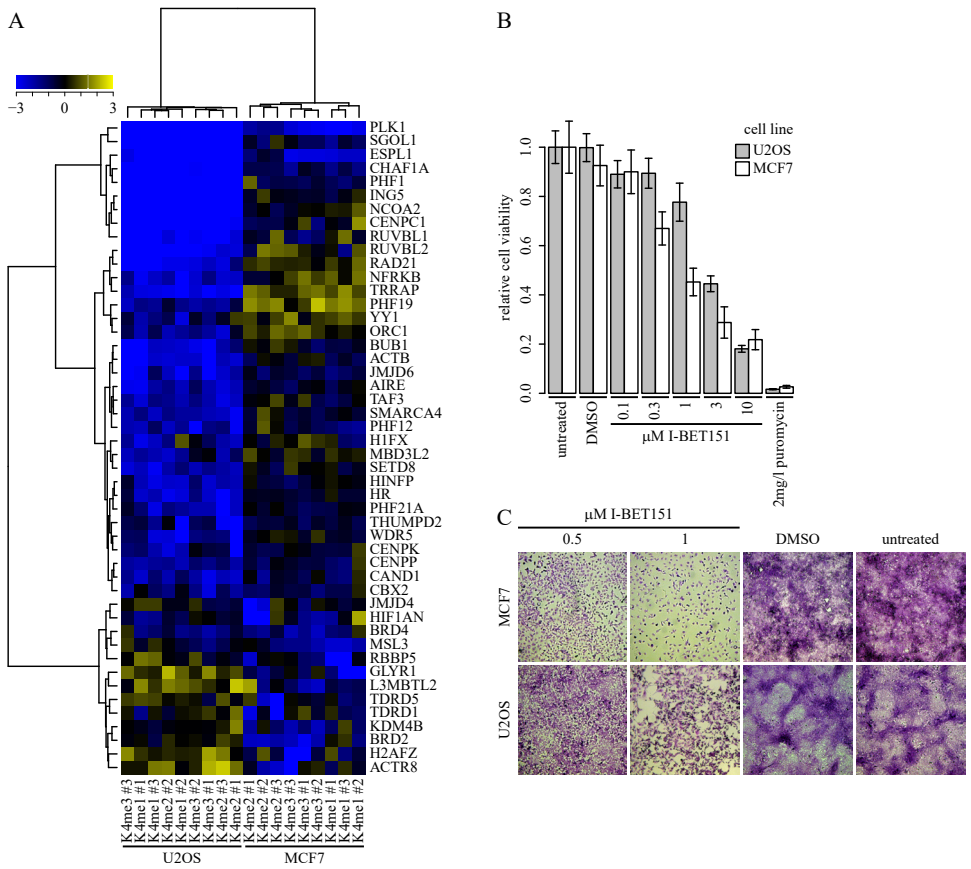


Figure 4. Effect on cell growth on U2OS and MCF7 by siRNA transfection and I-BET151 treatment. (A) Hierarchical clustering of valid object count per field for U2OS and MCF7 from the CROSS results for H3K4 methylation. Robust *z* scores were calculated using an R script (see Materials and Methods section). Replicates are indicated on the bottom and identified hits affecting growth are on the right. Coloring (ascending from blue to yellow via black) is based on the robust *z* score. (B) Cell viability was identified using an MTT assay as described in the Materials and Methods section. Relative cell viability was plotted against increasing I-BET151 concentration, and untreated cells were set at 1. (C) Crystal violet staining of I-BET151– treated cells. Cells were grown and stained as described in the Materials and Methods. Images were taken using a Zeiss light microscope (Jena, Germany) and Canon digital camera at (Tokyo, Japan) 40 × magnification.

MCF7 cells, both BRD2 and BRD4 are important for cell growth, whereas in U2OS cells, only BRD4 seems to be involved (Fig. 4A).

Discussion

Here we describe the development of a new method (CROSS) to identify writers and erasers of epigenetic marks. As a proof of principle, we identified the SET/MLL complexes to be involved in the methylation of histone H3K4, as well as the lysine demethylases KDM5A and KDM5C in the removal of H3K4me3, but not H3K4me1 or -me2. Similarly, we used CROSS to identify the writers and erasers of H3K27 methylation (PRC2 and KDM6A/UTX, respectively). Finally, we show that CROSS can be used to identify chromatin proteins that are essential for growth in cancer cell lines.

Histone H3K4 Methylation

Using CROSS, we identified the interaction partners of the writers involved in H3K4 methylation. The H3K4 methyltransferases themselves, however, do not show reduced global methylation levels when depleted in cells. A single methyltransferase complex (COMPASS/ Set1) catalyzes H3K4 methylation in yeast, while in mammals, at least six different COMPASS-related complexes are known to methylate H3K4. Possibly each of these methyltransferase complexes catalyzes only a minority of the total amount of H3K4 methylation in the cell lines that we used, and CROSS is not sensitive enough to detect such moderate reductions in global modification levels. In support of this, the knockdown of subunits that are shared by multiple methyltransferase complexes such as WDR82 and CXXC1 (SET1A/ SET1B complexes) resulted in a detectable reduction of global H3K4me3 levels. These observations are also in agreement with observations from the Shilatifard laboratory (72). Interestingly, we found that knockdown of RNF40 results in reduced H3K4me3 levels in U2OS (Fig. 2A). This is in agreement with the observation that ubiquitination of H2B is required for H3K4 trimethylation (286). Thus, CROSS can reveal crosstalk between different histone modifications. Recently, it was found that subunits of the Drosophila methyltransferase Trr complex (the MLL3/4 homolog) are involved in the monomethylation of H3K4 in chromatin areas with an enhancer-like signature (74). However, we did not find any effect of the MLL3/4-specific subunits (KDM6A/UTX, PA1, or PTIP) on the global levels of monomethylation. Still, removal of these subunits using knockdown by siRNA might have effects at specific genomic locations but not on the global levels of H3K4me1.

Histone H3K27 Methylation

Trimethylation of histone H3K27 is mediated by the PRC2 complex. How monomethylation is established is less clear (296). We used CROSS to identify the writers and erasers of H3K27me1. In U2OS, knocking down PRC2 components EZH2, SUZ12, and EED results in reduced H3K27me1 levels (Fig. 3). In contrast, knockdown of these proteins in MCF7 has little effect on the levels of H3K27me1. Interestingly, knocking down two H3K9 methyltransferases, EHMT1/EHMT2, reduced H3K27me1 in MCF7 cells (Fig. 3). In the deconvolution of these data (transfecting the single oligonucleotides instead of the gene pool of four siRNAs), these observations could not be confirmed (data not shown). Further experiments are therefore required to investigate these observations.

Interestingly, our experiments revealed that U2OS and MCF7 have markedly different levels of H3K27me3, even though EZH2 protein levels in both cell lines are similar (Fig. 1A,B). Although at this point, the molecular mechanism underlying this difference is unclear, aberrant H3K27me3 levels can have a profound impact on cell proliferation. This is evident from recent publications that link altered PRC2 activity to cancer (113). In hormone-independent growing prostate cancer cells, EZH2 appears to have an activating, PRC2-independent role (297). Also, H3K27 mutations were linked to gliomas (298). Given these findings and the recent development of highly specific EZH2 inhibitors (299,300), PRC2 in general and EZH2 in particular will remain a major focus, both in fundamental and clinical research.

CROSS to Identify Targets for Epidrugs

During normal development, the cell fate switch from proliferation to differentiation is driven by epigenetic alterations. It is therefore not surprising that epigenetic regulators,

when mutated, are important drivers of cancer. Unlike genetic mutations, altered epigenetic profiles are in principle all reversible, which makes them attractive targets for the development of so-called epidrugs (279,280). Our CROSS setup to screen for growth effects in U2OS and MCF7 (Fig. 4A) could be expanded to a larger panel of cell lines corresponding to different tumor stages and also different tumor origins. Such experiments could lead to the identification of novel targets for drugs.

CROSS in Perspective: Limitations and Strength

CROSS analysis on H3K4 and H3K27 methylation identified the known major complexes targeting these marks (Figs. 2A and 3). Inspection of the data reveals that the dispersion in the MCF7 CROSS data is larger than in U2OS (Suppl. Figs. S5 and S6). However, statistical analysis of the data using limma (301) confirmed the hits identified by robust *z* score analysis (data not shown).

On the basis of our experiments, we estimate that a knockdown efficiency of >50% is required to have an effect on global histone modification (Suppl. Fig. S11) and therefore to be picked up by CROSS. We expect that CROSS will be a powerful tool to identify the writers and erasers for known and novel chromatin marks (*e.g.*, demethylases for H3K79 and H4K20 and recently reported novel histone modifications (302)). Furthermore, CROSS can be applied to histone phosphorylation during mitosis by monitoring the amount of phosphorylation on either H3S10 or H3T3 after siRNA knockdown. In a proof-of-principle experiment, we found that knockdown of the mitotic kinase Haspin reduces H3T3ph as expected (Suppl. Fig. S12).

Currently, the limitation of our experiments is the use of a dedicated library of 529 genes based on known domains and complexes involved in writing, reading, and erasing histone and DNA modifications, as used for proof of principle of CROSS. Clearly, this library is biased toward "known" histone and chromatin biology, and opportunities lie in the use of a genome-wide siRNA library. Such experiments could reveal the interplay between chromatin and metabolism, for example. This is particularly important given the recently emerging strong link between chromatin, metabolism, and cancer. For example, it was recently shown that threonine levels in embryonic stem cells directly influenceH3K4me3 levels (303). Furthermore, metabolites that are part of the TCA cycle such as succinate directly affect the enzymatic activity of lysine demethylases (304,305). Clearly, additional crosstalk between epigenetics and seemingly unrelated cellular pathways and processes are yet to be found, and our CROSS method will be a powerful tool to facilitate such discoveries.

Acknowledgments

We thank Wienand Omta and Arne Smits for help and useful discussions on mathematics and R scripts, Simona Antonova for technical assistance, and the whole Timmers/Vermeulen group for useful discussions. We thank Jan Hendrickx and Didier Allaer from Diagenode (Belgium) for providing antibodies and sharing with us the H3K27ac antibody before it was commercially available. I-BET151 was a kind gift of Dr. Michael D. Woodrow and Dr. Rab Prinjha from GlaxoSmithKline (United Kingdom). Anti-ASH2L was a kind gift of Professor Winship Herr (UNIL, Switzerland).

Supplemental Materials and Methods

qPCR analysis of knockdown

Cells were transfected with siRNA as described in the Materials and Methods. 72 h after transfection total RNA was isolated using RNeasy kit (Qiagen) including a DNase treatment step. 500 ng of total RNA was used for cDNA synthesis (Superscript II, Invitrogen) using random oligo primers. Expression of b-Actin, GAPDH, MLL1, MLL2, INO80, INO80B, INO80E, ASH2L, BRD1, BRD2, BRD3 and BRD4 were analyzed by quantitative PCR and normalized as indicated.

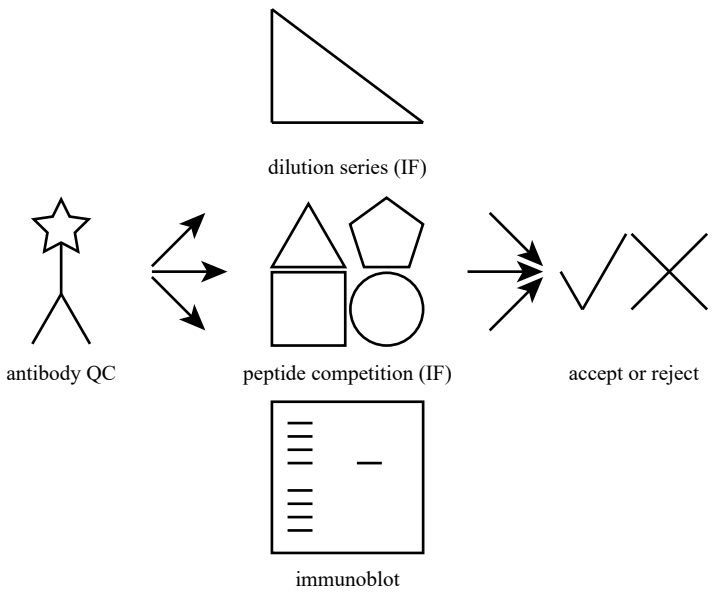


Figure S1. Antibody quality control setup. Antibody quality control consists of three parts. 1: A dilution series is performed on fixed cells using an immunofluorescence assay. 2: Peptide competition is performed on fixed cells to indicate antibody specificity. 3: Immunoblot analysis is performed on in-house synthesized biotin-labelled peptides to indicate antibody specificity.

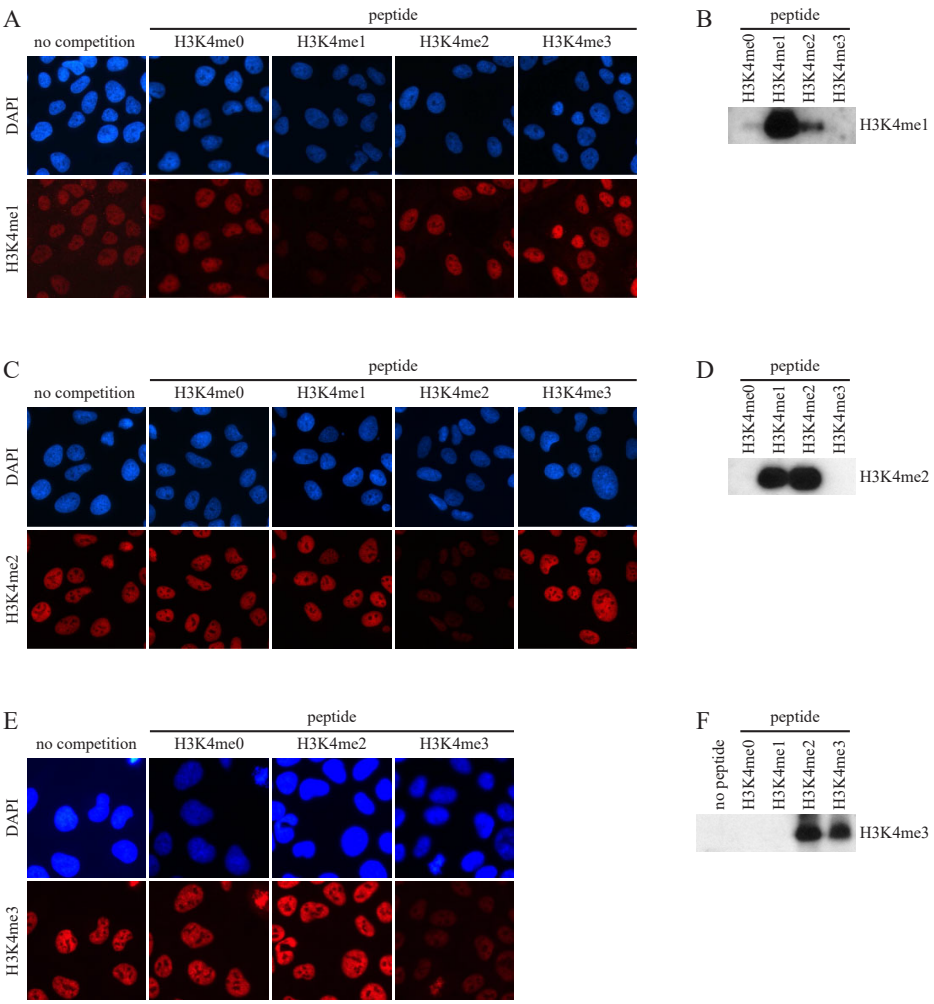


Figure S2. H3K4 methylation antibody control in U2OS cells. (A, C, E) Peptide competition assay for (A) H3K4me1, (C) H3K4me2 and (E) H3K4me3 antibodies, peptides are indicated on the top. (B, D, F) Immunoblot for (B) H3K4me1, (D) H3K4me2 and (F) H3K4me3 antibodies, peptides are indicated on the top. Immunoblots were stained by the indicated antibodies.

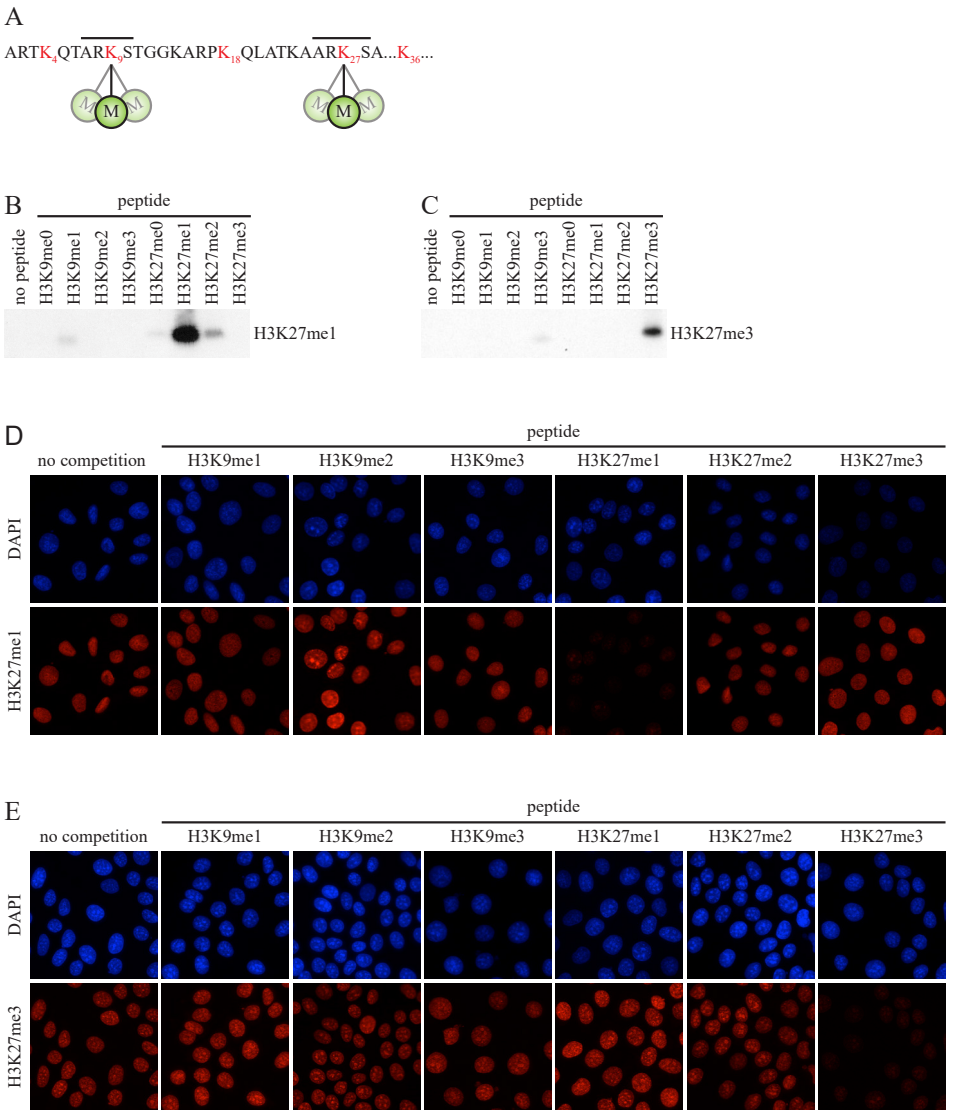


Figure S3. H3K27 methylation antibody control in MCF7 cells. (A) Amino acid sequence of the histone 3 tail. Red Lysines indicate possible methylation sites. M in green circle indicates possible methylation of Lysines at site 9 and 27. Black horizontal bars indicate identical sites within the histone 3 tail. (B, C) Immunoblot for (B) H3K27me1 and (C) H3K27me3 antibodies. Peptides are indicated on the top. Immunoblots were stained by the indicated antibodies. (D, E) Peptide competition assay for (D) H3K27me1 and (E) H3K27me3 antibodies. Peptides are indicated on the top.

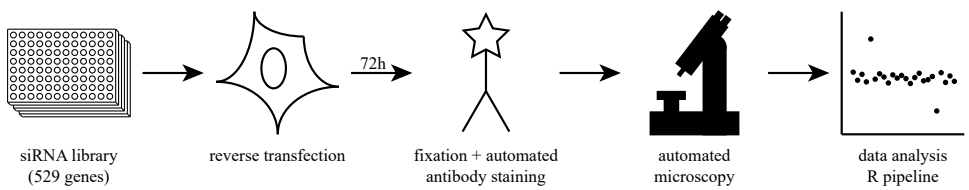


Figure S4. Schematic representation of CROSS pipeline. Cells are reverse transfected using a predefined siRNA library containing 529 genes. After 72 hours the cells are fixed and stained using a robotic platform. Automated microscopy is used to measure global nuclear intensity levels. Data analysis is performed using an R script.

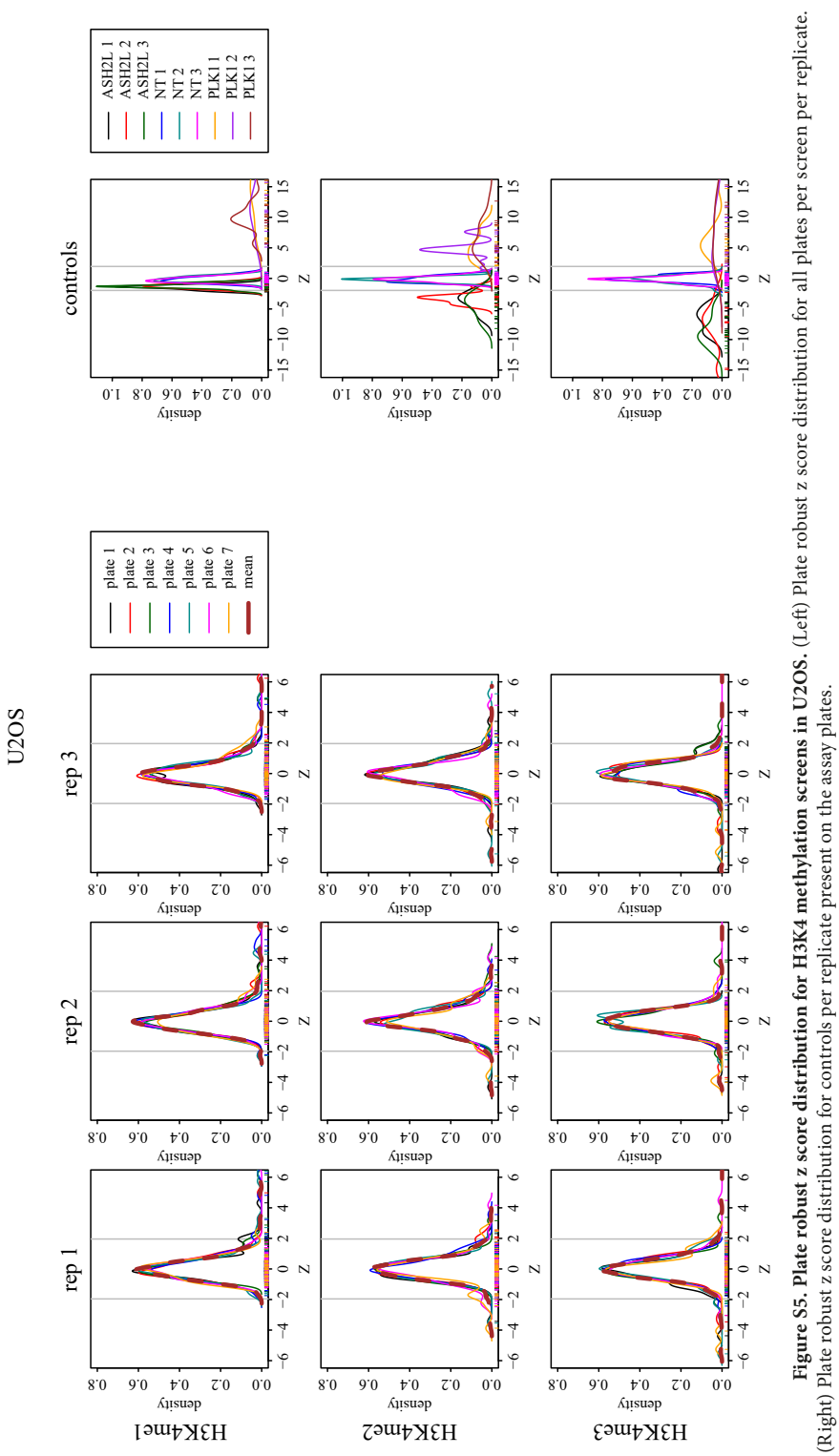


Figure S5. Plate robust z score distribution for H3K4 methylation screens in U2OS. (Left) Plate robust z score distribution for all plates per screen per replicate. (Right) Plate robust z score distribution for controls per replicate present on the assay plates.

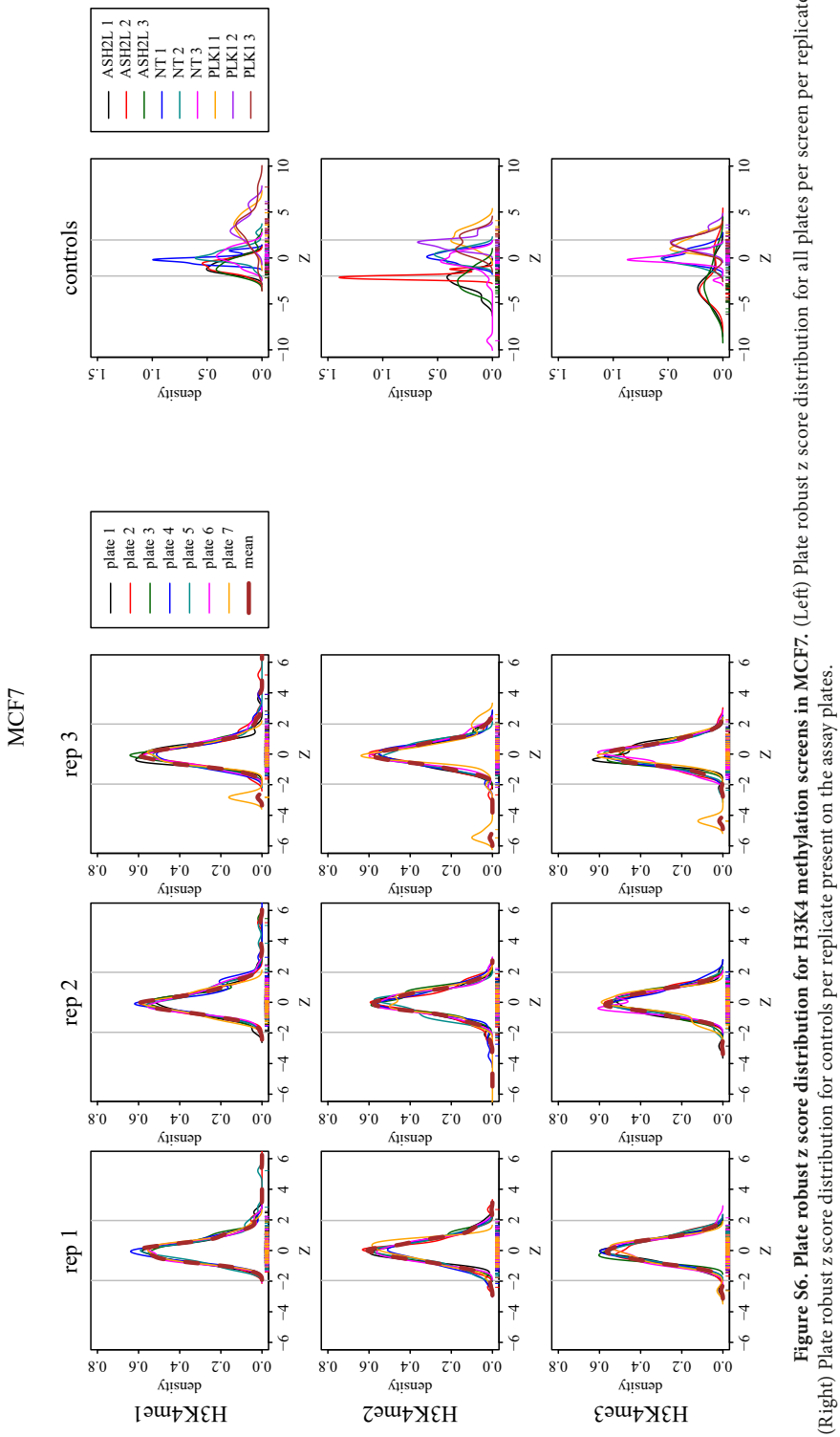


Figure S6. Plate robust z score distribution for H3K4 methylation screens in MCF7. (Left) Plate robust z score distribution for all plates per screen per replicate. (Right) Plate robust z score distribution for controls per replicate present on the assay plates.

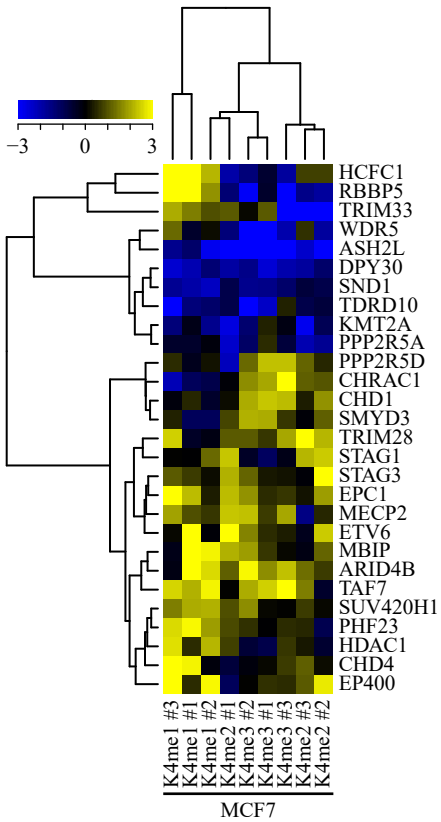


Figure S7. Hierarchical clustering of mono, di and trimethylation of H3K4 in MCF7. Robust z scores were calculated using an R script (see methods section). Replicates are indicated on the bottom and identified hits are on the right. Coloring (ascending from blue to yellow via black) is based on the robust z score for the indicated antibody signal.

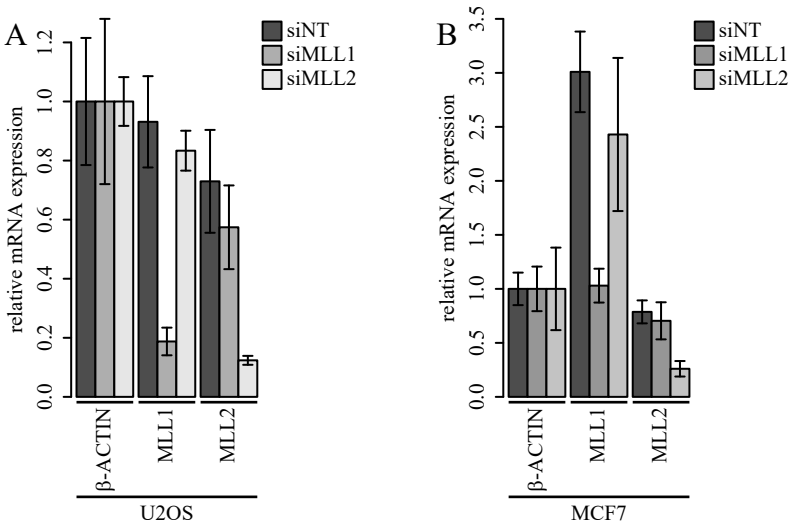


Figure S8. Knockdown analysis of MLL subunits using RT-qPCR. cDNA was prepared from U2OS (A) or MCF7 (B) cells transfected with siRNA targeting MLL1, MLL2 or control siRNA. Relative mRNA levels of MLL1, MLL2 and β-Actin were analyzed by quantitative PCR and were corrected for β-Actin.

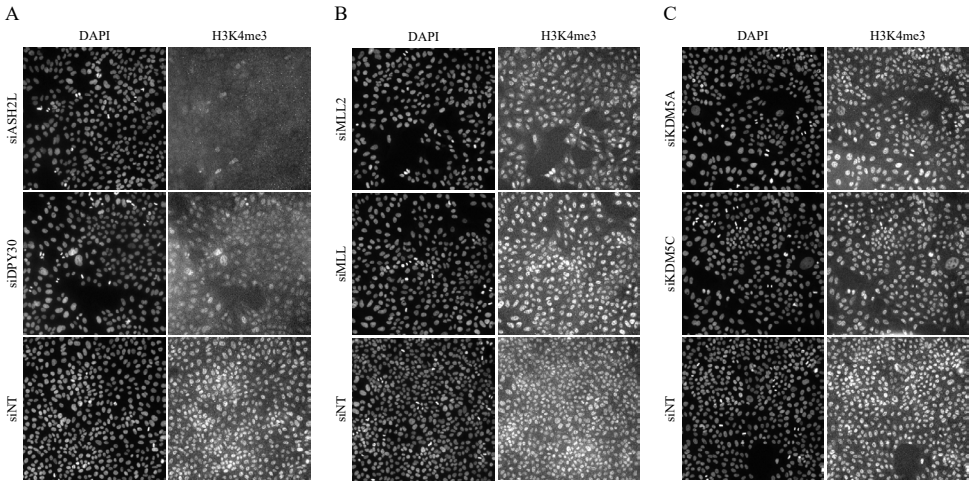


Figure S9. U2OS H3K4me3 CROSS representative Arrayscan images. Depicted are images extracted from the automated microscope. siRNA transfection as indicated. DAPI and H3K4me3 channels are shown.

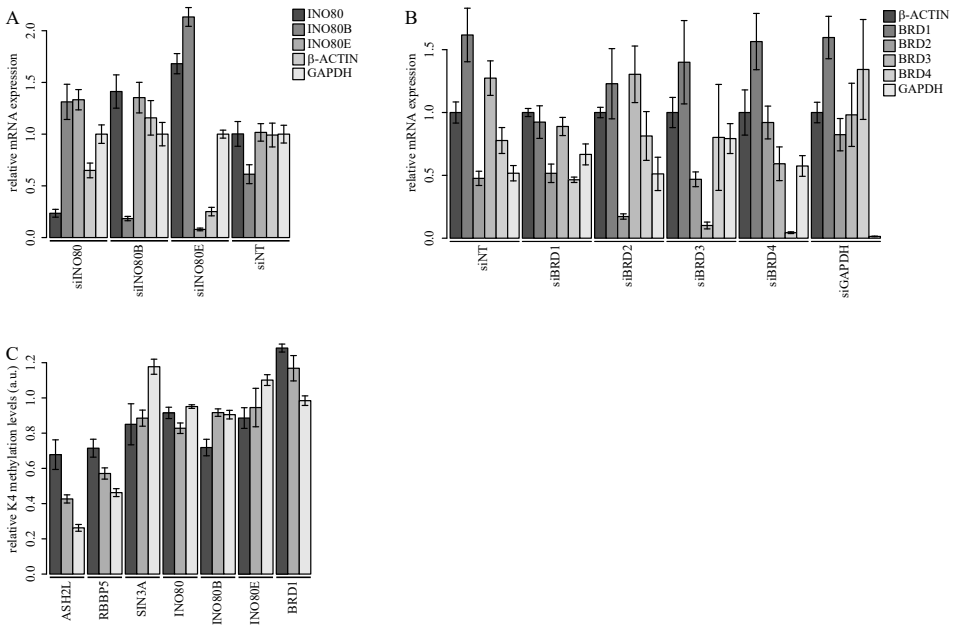


Figure S10. Knockdown analysis of INO80 and BRD subunits using RT-qPCR. (A) cDNA was prepared from U2OS cells transfected with siRNA targeting INO80, INO80B, INO80E or control siRNA. Relative mRNA levels of INO80, INO80B, INO80E, GAPDH and β-Actin were analyzed by quantitative PCR and were corrected for GAPDH. (B) cDNA was prepared from U2OS cells transfected with siRNA targeting BRD1, BRD2, BRD3, BRD4, GAPDH or control siRNA. Relative mRNA levels of BRD1, BRD2, BRD3, BRD4, GAPDH and β-Actin were analyzed by quantitative PCR and were corrected for β-Actin. (C) Relative H3K4 methylation levels of a selection of genes from the U2OS H3K4 methylation screens. Levels were corrected by the median of the plate on which the values were collected from.

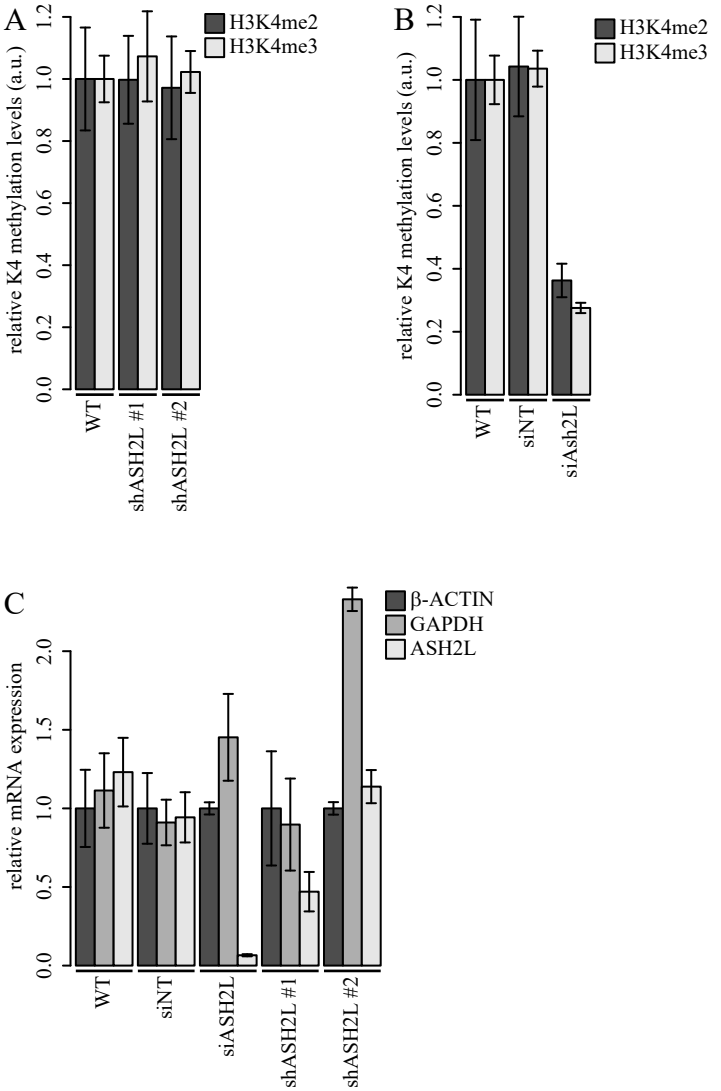


Figure S11. Comparison of global H3K4 levels in siRNA transient knockdown and shRNA stable knockdown cell lines. (A) Immunofluoresence assay of H3K4 methylation levels of shRNA knockdown cell lines. (B) Immunofluoresence assay of H3K4 methylation levels of siRNA transient knockdown cells (C) cDNA was prepared from either U2OS cells transiently transfected with siRNA targeting ASH2L or control siRNA, or from U2OS infected with pLKO.1-ASH2L shRNA viral particles (stable cell lines, line #1 and line #2). Relative mRNA levels of ASH2L, GAPDH and β-Actin were analyzed by quantitative PCR and were corrected for β-Actin.

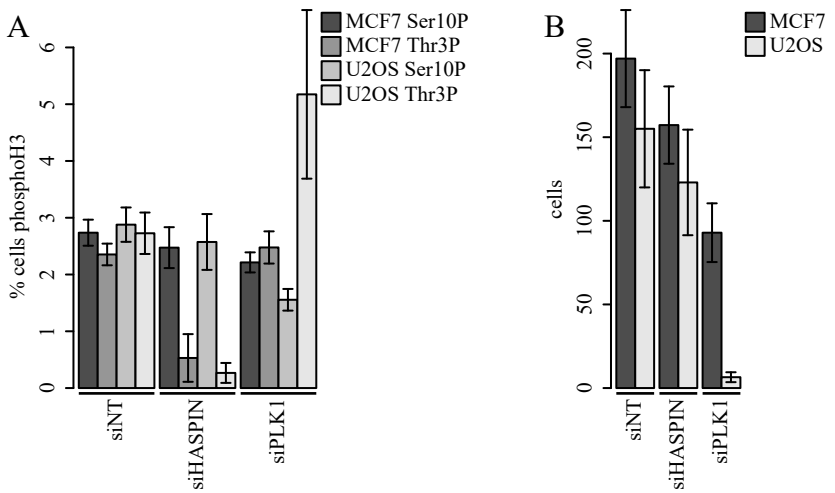


Figure S12. Comparison of global phosphorylation levels of Thr3 and Ser10 upon transient Haspin knockdown. (A) Immunofluoresence assay of phosphorylation levels. Indicated are the percentage of positive cells for the phosphomark (either H3S10ph or H3T3ph) (B) Cells per measured field.

Table S1: List of 529 siRNA targets used for CROSS.

Plate	Row	Col	ThermoGeneName	AllGeneNames	UniprotGeneName	UniprotID	HGNC
1	A	2	ACACA	ACACA ACAC ACC1 ACCA	ACACA_HUMAN	Q13085	ACACA
1	B	2	ACTA2	ACTA2 ACTSA ACTVS GIG46	ACTA_HUMAN	P62736	ACTA2
1	C	2	ASF1B	ASF1B	ASF1B_HUMAN	Q9NVP2	ASF1B
1	D	2	ASH1L	ASH1L KIAA1420 KMT2H	ASH1L_HUMAN	Q9NR48	ASH1L
1	E	2	BRCA1	BRCA1 RNF53	BRCA1_HUMAN	P38398	BRCA1
1	F	2	BRD1	BRD1 BRL BRPF2	BRD1_HUMAN	O95696	BRD1
1	G	2	C6orf173	CENPW C6orf173 CUG2	CENPW_HUMAN	Q5EE01	CENPW
1	H	2	TIP120A	CAND1 KIAA0829 TIP120 TIP120A	CAND1_HUMAN	Q86VP6	CAND1
1	A	3	ACTB	ACTB	ACTB_HUMAN	P60709	ACTB
1	B	3	BAF53A	ACTL6A BAF53 BAF53A INO80K	ACL6A_HUMAN	O96019	ACTL6A
1	C	3	ASH2L	ASH2L ASH2L1	ASH2L_HUMAN	Q9UBL3	ASH2L
1	D	3	ASXL1	ASXL1 KIAA0978	ASXL1_HUMAN	Q8IXJ9	ASXL1
1	E	3	BRD2	BRD2 KIAA9001 RING3	BRD2_HUMAN	P25440	BRD2
1	F	3	BRD3	BRD3 KIAA0043 RING3L	BRD3_HUMAN	Q15059	BRD3
1	G	3	CARM1	CARM1 PRMT4	CARM1_HUMAN	Q86X55	CARM1
1	H	3	CBX1	CBX1 CBX	CBX1_HUMAN	P83916	CBX1
1	A	4	ACTL6	ACTL6B ACTL6 BAF53B	ACL6B_HUMAN	O94805	ACTL6B
1	B	4	ACTR5	ACTR5 ARP5	ARP5_HUMAN	Q9H9F9	ACTR5
1	C	4	ATAD2	ATAD2 L16 PRO2000	ATAD2_HUMAN	Q6PL18	ATAD2
1	D	4	KIAA1240	ATAD2B KIAA1240	ATD2B_HUMAN	Q9ULI0	ATAD2B
1	E	4	BRD4	BRD4 HUNK1	BRD4_HUMAN	O60885	BRD4
1	F	4	BRD7	BRD7 BP75 CELTIX1	BRD7_HUMAN	Q9NPI1	BRD7
1	G	4	CBX2	CBX2	CBX2_HUMAN	Q14781	CBX2
1	H	4	CBX4	CBX4	CBX4_HUMAN	O00257	CBX4
1	A	5	ACTR6	ACTR6 CDA12	ARP6_HUMAN	Q9GZN1	ACTR6
1	B	5	ACTR8	ACTR8 ARP8 INO80N	ARP8_HUMAN	Q9H981	ACTR8
1	C	5	ATF7IP	ATF7IP MCAF MCAF1	MCAF1_HUMAN	Q6VMQ6	ATF7IP
1	D	5	ATF7IP2	ATF7IP2 MCAF2	MCAF2_HUMAN	Q5U623	ATF7IP2
1	E	5	BRD8	BRD8 SMAP SMAP2	BRD8_HUMAN	Q9H0E9	BRD8
1	F	5	BRD9	BRD9 UNQ3040/PRO9856	BRD9_HUMAN	Q9H8M2	BRD9
1	G	5	CBX6	CBX6	CBX6_HUMAN	O95503	CBX6
1	H	5	CBX7	CBX7	CBX7_HUMAN	O95931	CBX7
1	A	6	AIRE	AIRE APECED	AIRE_HUMAN	O43918	AIRE
1	B	6	AKAP1	AKAP1 AKAP149 PRKA1	AKAP1_HUMAN	Q92667	AKAP1
1	C	6	ATRX	ATRX RAD54L XH2	ATRX_HUMAN	P46100	ATRX
1	D	6	SCA7	ATXN7 SCA7	ATX7_HUMAN	O15265	ATXN7
1	E	6	BRDT	BRDT	BRDT_HUMAN	Q58F21	BRDT
1	F	6	BRPF1	BRPF1 BR140	BRPF1_HUMAN	P55201	BRPF1
1	G	6	CBX8	CBX8 PC3 RC1	CBX8_HUMAN	Q9HC52	CBX8
1	H	6	LOC112869	CCDC101 SGF29	SGF29_HUMAN	Q96ES7	CCDC101
1	A	7	ANP32E	ANP32E	AN32E_HUMAN	Q9BTT0	ANP32E
1	B	7	APBB1	APBB1 FE65 RIR	APBB1_HUMAN	O00213	APBB1
1	C	7	DKFZP761G2113	ATXN7L3	AT7L3_HUMAN	Q14CW9	ATXN7L3

1	D	7	KIAA1447	BAHCC1 BAHD2 KIAA1447	BAHC1_HUMAN	Q9P281	BAHCC1
1	E	7	BRPF3	BRPF3 KIAA1286	BRPF3_HUMAN	Q9ULD4	BRPF3
1	F	7	C21ORF107	BRWD1 C21orf107 WDR9	BRWD1_HUMAN	Q9NSI6	BRWD1
1	G	7	CDY1	CDY1 CDY1A	CDY1_HUMAN	Q9Y6F8	CDY1
1	H	7	CDY2	CDY2A CDY2	CDY2_HUMAN	Q9Y6F7	CDY2A
1	A	8	APITD1	APITD1 CENPS FAAP16 MHF1	CENPS_HUMAN	Q8N2Z9	APITD1
1	B	8	APP	APP A4 AD1	A4_HUMAN	P05067	APP
1	C	8	BAHD1	BAHD1 KIAA0945	BAHD1_HUMAN	Q8TBE0	BAHD1
1	D	8	BAZ1A	BAZ1A ACF1 WCRF180 HSPC317	BAZ1A_HUMAN	Q9NRL2	BAZ1A
1	E	8	BRODL	BRWD3	BRWD3_HUMAN	Q6RI45	BRWD3
1	F	8	BUB1	BUB1 BUB1L	BUB1_HUMAN	O43683	BUB1
1	G	8	CDYL	CDYL CDYL1	CDYL1_HUMAN	Q9Y232	CDYL
1	H	8	CDYL2	CDYL2	CDYL2_HUMAN	Q8N8U2	CDYL2
1	A	9	ARID1A	ARID1A BAF250 BAF250A C1orf4 OSA1 SMARCF1	ARI1A_HUMAN	O14497	ARID1A
1	B	9	ARID1B	ARID1B BAF250B DAN15 KIAA1235 OSA2	ARI1B_HUMAN	Q8NFD5	ARID1B
1	C	9	BAZ1B	BAZ1B WBSC10 WBSCR10 WBSCR9 WSTF	BAZ1B_HUMAN	Q9UIG0	BAZ1B
1	D	9	BAZ2A	BAZ2A KIAA0314 TIP5	BAZ2A_HUMAN	Q9UIF9	BAZ2A
1	E	9	C14ORF106	MIS18BP1 C14orf106 KIAA1903 KNL2 M18BP1	M18BP_HUMAN	Q6P0N0	MIS18BP1
1	F	9	C14ORF169	NO66 C14orf169 MAPJD	NO66_HUMAN	Q9H6W3	C14orf169
1	G	9	CECR2	CECR2 KIAA1740	CECR2_HUMAN	Q9BXF3	CECR2
1	H	9	CENPA	CENPA	CENPA_HUMAN	P49450	CENPA
1	A	10	ARID2	ARID2 KIAA1557	ARID2_HUMAN	Q68CP9	ARID2
1	B	10	ARID4A	ARID4A RBBP1 RBP1	ARI4A_HUMAN	P29374	ARID4A
1	C	10	BAZ2B	BAZ2B KIAA1476	BAZ2B_HUMAN	Q9UIF8	BAZ2B
1	D	10	BLM	BLM RECQ2 RECQL3	BLM_HUMAN	P54132	BLM
1	E	10	MGC4606	PAGR1 C16orf53 PA1	PAGR1_HUMAN	Q9BTK6	PAGR1
1	F	10	C20ORF20	MRGBP C20orf20	MRGBP_HUMAN	Q9NV56	MRGBP
1	G	10	CENPB	CENPB	CENPB_HUMAN	P07199	CENPB
1	H	10	CENPC1	CENPC1 CENPC ICEN7	CENPC_HUMAN	Q03188	CENPC1
1	A	11	ARID4B	ARID4B BRCAA1 RBBP1L1 RBP1L1 SAP180	ARI4B_HUMAN	Q4LE39	ARID4B
1	B	11	ASF1A	ASF1A CGI-98 HSPC146	ASF1A_HUMAN	Q9Y294	ASF1A
1	C	11	BMI1	BMI1 PCGF4 RNF51	BMI1_HUMAN	P35226	BMI1
1	D	11	FALZ	BPTF FAC1 FALZ	BPTF_HUMAN	Q12830	BPTF
1	E	11	C21ORF45	MIS18A C21orf45 C21orf46 FASP1	MS18A_HUMAN	Q9NYP9	MIS18A
1	F	11	FLJ37953	TYW5 C2orf60	TYW5_HUMAN	A2RUC4	TYW5
1	G	11	CENPH	CENPH ICEN35	CENPH_HUMAN	Q9H3R5	CENPH
1	H	11	FSHPRH1	CENPI FSHPRH1 ICEN19 LRPR1	CENPI_HUMAN	Q92674	CENPI
2	A	2	FKSG14	CENPK ICEN37 FKSG14	CENPK_HUMAN	Q9BS16	CENPK
2	B	2	DJ383J4.3	CENPL C1orf155 ICEN33	CENPL_HUMAN	Q8N0S6	CENPL
2	C	2	CYP11A1	CYP11A1 CYP11A	CP11A_HUMAN	P05108	CYP11A1
2	D	2	CREBBP	CREBBP CBP	CBP_HUMAN	Q92793	CREBBP
2	E	2	ELP3	ELP3	ELP3_HUMAN	Q9H9T3	ELP3
2	F	2	E(Y)2	ENY2 DC6	ENY2_HUMAN	Q9NPA8	ENY2
2	G	2	GSG2	GSG2	HASP_HUMAN	Q8TF76	GSG2

2	H	2	GTF2I	GTF2I BAP135 WBSCR6	GTF2I_HUMAN	P78347	GTF2I
2	A	3	C22ORF18	CENPM C22orf18 ICEN39 PANE1	CENPM_HUMAN	Q9NSP4	CENPM
2	B	3	BM039	CENPN C16orf60 ICEN32 BM-309	CENPN_HUMAN	Q96H22	CENPN
2	C	3	CSNK2B_2	CSNK2B CK2N G5A	CSK2B_HUMAN	P67870	CSNK2B_2
2	D	3	CSRP2BP	CSRP2BP	CSR2B_HUMAN	Q9H8E8	CSRP2BP
2	E	3	EP300	EP300 P300	EP300_HUMAN	Q09472	EP300
2	F	3	EP400	EP400 CAGH32 KIAA1498 KIAA1818 TNRC12	EP400_HUMAN	Q96L91	EP400
2	G	3	GTF3C4	GTF3C4	TF3C4_HUMAN	Q9UKN8	GTF3C4
2	H	3	PARP1_1	PARP1 ADPRT PPOL	PARP1_HUMAN	P09874	PARP1_1
2	A	4	MGC11266	CENPO ICEN36 MCM21R	CENPO_HUMAN	Q9BU64	CENPO
2	B	4	RP11-19J3.3	CENPP	CENPP_HUMAN	Q6IPU0	CENPP
2	C	4	CTCF	CTCF	CTCF_HUMAN	P49711	CTCF
2	D	4	CXXC1	CXXC1 CFP1 CGBP PCCX1 PHF18	CXXC1_HUMAN	Q9P0U4	CXXC1
2	E	4	EPC1	EPC1	EPC1_HUMAN	Q9H2F5	EPC1
2	F	4	EPC2	EPC2	EPC2_HUMAN	Q52LR7	EPC2
2	G	4	H1F0	H1F0 H1FV	H10_HUMAN	P07305	H1F0
2	H	4	H2AFZ	H2AFZ H2AZ	H2AZ_HUMAN	P0C0S5	H2AFZ
2	A	5	C6ORF139	CENPQ C6orf139	CENPQ_HUMAN	Q7L2Z9	CENPQ
2	B	5	FLJ13111	CENPT C16orf56 ICEN22	CENPT_HUMAN	Q96BT3	CENPT
2	C	5	DHX9	DHX9 DDX9 LKP NDH2	DHX9_HUMAN	Q08211	DHX9
2	D	5	DATF1	DIDO1 C20orf158 DATF1 KIAA0333	DIDO1_HUMAN	Q9BTC0	DIDO1
2	E	5	FLJ20105	ERCC6L PICH	ERC6L_HUMAN	Q2NKX8	ERCC6L
2	F	5	EFO1	ESCO1 EFO1 KIAA1911	ESCO1_HUMAN	Q5FWF5	ESCO1
2	G	5	H3F3A	H3F3A H3.3A H3F3 PP781	H33_HUMAN	P84243	H3F3A
2	H	5	H3F3B	H3F3A H3.3A H3F3 PP781 H3F3B H3.3B	H33_HUMAN	P84243	H3F3B
2	A	6	CHAF1A	CHAF1A CAF CAF1P150	CAF1A_HUMAN	Q13111	CHAF1A
2	B	6	CHD1	CHD1	CHD1_HUMAN	O14646	CHD1
2	C	6	DMAP1	DMAP1 KIAA1425	DMAP1_HUMAN	Q9NPF5	DMAP1
2	D	6	DNMT1	DNMT1 AIM CXXC9 DNMT	DNMT1_HUMAN	P26358	DNMT1
2	E	6	LOC157570	ESCO2	ESCO2_HUMAN	Q56NI9	ESCO2
2	F	6	ESPL1	ESPL1 ESP1 KIAA0165	ESPL1_HUMAN	Q14674	ESPL1
2	G	6	HAT1	HAT1 KAT1	HAT1_HUMAN	O14929	HAT1
2	H	6	HCFC1	HCFC1 HCF1 HFC1	HCFC1_HUMAN	P51610	HCFC1
2	A	7	CHD2	CHD2	CHD2_HUMAN	O14647	CHD2
2	B	7	CHD3	CHD3	CHD3_HUMAN	Q12873	CHD3
2	C	7	DNMT3B	DNMT3B	DNM3B_HUMAN	Q9UBC3	DNMT3B
2	D	7	DOT1L	DOT1L KIAA1814 KMT4	DOT1L_HUMAN	Q8TEK3	DOT1L
2	E	7	ETV6	ETV6 TEL TEL1	ETV6_HUMAN	P41212	ETV6
2	F	7	EZH1	EZH1 KIAA0388	EZH1_HUMAN	Q92800	EZH1
2	G	7	HDAC1	HDAC1 RPD3L1	HDAC1_HUMAN	Q13547	HDAC1
2	H	7	HDAC10	HDAC10	HDA10_HUMAN	Q969S8	HDAC10
2	A	8	CHD4	CHD4	CHD4_HUMAN	Q14839	CHD4
2	B	8	CHD5	CHD5 KIAA0444	CHD5_HUMAN	Q8TDI0	CHD5
2	C	8	DPF1	DPF1 BAF45B NEUD4	DPF1_HUMAN	Q92782	DPF1
2	D	8	DPF2	DPF2 BAF45D REQ UBID4	REQU_HUMAN	Q92785	DPF2

2	E	8	EZH2	EZH2 KMT6	EZH2_HUMAN	Q15910	EZH2
2	F	8	P38IP	SUPT20H C13orf19 FAM48A FP757	SP20H_HUMAN	Q8NEM7	SUPT20H
2	G	8	HDAC11	HDAC11	HDA11_HUMAN	Q96DB2	HDAC11
2	H	8	HDAC2	HDAC2	HDAC2_HUMAN	Q92769	HDAC2
2	A	9	CHD6	CHD6 CHD5 KIAA1335 RIGB	CHD6_HUMAN	Q8TD26	CHD6
2	B	9	CHD7	CHD7 KIAA1416	CHD7_HUMAN	Q9P2D1	CHD7
2	C	9	DPF3	DPF3 BAF45C CERD4	DPF3_HUMAN	Q92784	DPF3
2	D	9	LOC84661	DPY30	DPY30_HUMAN	Q9C005	DPY30
2	E	9	FBXL19	FBXL19 FBL19	FXL19_HUMAN	Q6PCT2	FBXL19
2	F	9	KIAA1333	G2E3 KIAA1333	G2E3_HUMAN	Q7L622	G2E3
2	G	9	HDAC3	HDAC3	HDAC3_HUMAN	O15379	HDAC3
2	H	9	HDAC4	HDAC4 KIAA0288	HDAC4_HUMAN	P56524	HDAC4
2	A	10	CHD8	CHD8 HELSNF1 KIAA1564	CHD8_HUMAN	Q9HCK8	CHD8
2	B	10	CHD9	CHD9 KIAA0308 KISH2 PRIC320 AD-013 x0008	CHD9_HUMAN	Q3L8U1	CHD9
2	C	10	DR1	DR1	NC2B_HUMAN	Q01658	DR1
2	D	10	EED	EED	EED_HUMAN	O75530	EED
2	E	10	P66ALPHA	GATAD2A	P66A_HUMAN	Q86YP4	GATAD2A
2	F	10	P66BETA	GATAD2B KIAA1150	P66B_HUMAN	Q8WXI9	GATAD2B
2	G	10	HDAC5	HDAC5 KIAA0600	HDAC5_HUMAN	Q9UQL6	HDAC5
2	H	10	HDAC6	HDAC6 KIAA0901 JM21	HDAC6_HUMAN	Q9UBN7	HDAC6
2	A	11	CHRAC1	CHRAC1 CHRAC15	CHRC1_HUMAN	Q9NRG0	CHRAC1
2	B	11	CORO2A	CORO2A IR10 WDR2	COR2A_HUMAN	Q92828	CORO2A
2	C	11	EHMT1	EHMT1 EUHMTASE1 GLP KIAA1876 KMT1D	EHMT1_HUMAN	Q9H9B1	EHMT1
2	D	11	EHMT2	EHMT2 BAT8 C6orf30 G9A KMT1C NG36	EHMT2_HUMAN	Q96KQ7	EHMT2
2	E	11	N-PAC	GLYR1 HIBDL NP60	GLYR1_HUMAN	Q49A26	GLYR1
2	F	11	GPS2	GPS2	GPS2_HUMAN	Q13227	GPS2
2	G	11	HDAC7A	HDAC7 HDAC7A	HDAC7_HUMAN	Q8WUI4	HDAC7
2	H	11	HDAC8	HDAC8 HDACL1 CDA07	HDAC8_HUMAN	Q9BY41	HDAC8
3	A	2	HDAC9	HDAC9 HDAC7 HDAC7B HDRP KIAA0744 MITR	HDAC9_HUMAN	Q9UKV0	HDAC9
3	B	2	HDGF	HDGF HMG1L2	HDGF_HUMAN	P51858	HDGF
3	C	2	FLJ20309	INO80D	IN80D_HUMAN	Q53TQ3	INO80D
3	D	2	FLJ90652	INO80E CCDC95	IN80E_HUMAN	Q8NBZ0	INO80E
3	E	2	JMJD1B	KDM3B C5orf7 JHDM2B JMJD1B KIAA1082	KDM3B_HUMAN	Q7LBC6	KDM3B
3	F	2	JMJD2A	KDM4A JHDM3A JMJD2 JMJD2A KIAA0677	KDM4A_HUMAN	O75164	KDM4A
3	G	2	MBD1	MBD1 CXXC3 PCM1	MBD1_HUMAN	Q9UIS9	MBD1
3	H	2	MBD2	MBD2	MBD2_HUMAN	Q9UBB5	MBD2
3	A	3	HDGF2	HDGFRP3 HDGF2 CGI-142	HDGR3_HUMAN	Q9Y3E1	HDGFRP2
3	B	3	PWWP1	HDGFL1 PWWP1	HDGL1_HUMAN	Q5TGJ6	HDGFL1
3	C	3	INTS12	INTS12 PHF22 SBBI22	INT12_HUMAN	Q96CB8	INTS12
3	D	3	ITGB3BP	ITGB3BP CENPR NRIF3	CENPR_HUMAN	Q13352	ITGB3BP
3	E	3	JMJD2B	KDM4B JHDM3B JMJD2B KIAA0876	KDM4B_HUMAN	O94953	KDM4B
3	F	3	JMJD2C	KDM4C GASC1 JHDM3C JMJD2C KIAA0780	KDM4C_HUMAN	Q9H3R0	KDM4C

3	G	3	MBD3	MBD3	MBD3_HUMAN	O95983	MBD3
3	H	3	MBD3L1	MBD3L1 MBD3L	MB3L1_HUMAN	Q8WWY6	MBD3L1
3	A	4	HDGFRP3	HDGFRP3 HDGF2 CGI-142	HDGR3_HUMAN	Q9Y3E1	HDGFRP3
3	B	4	HIF1AN	HIF1AN FIH1	HIF1N_HUMAN	Q9NWT6	HIF1AN
3	C	4	JARID2	JARID2 JMJ	JARD2_HUMAN	Q92833	JARID2
3	D	4	JAZF1	JAZF1 TIP27 ZNF802	JAZF1_HUMAN	Q86VZ6	JAZF1
3	E	4	JMJD2D	KDM4D JHDM3D JMJD2D	KDM4D_HUMAN	Q6B0I6	KDM4D
3	F	4	JARID1A	KDM5A JARID1A RBBP2 RBP2	KDM5A_HUMAN	P29375	KDM5A
3	G	4	MBD3L2	MBD3L2	MB3L2_HUMAN	Q8NHZ7	MBD3L2
3	H	4	MBD4	MBD4 MED1	MBD4_HUMAN	O95243	MBD4
3	A	5	MIZF	HINFP MIZF ZNF743	HINFP_HUMAN	Q9BQA5	HINFP
3	B	5	HIRA	HIRA DGCR1 HIR TUPLE1	HIRA_HUMAN	P54198	HIRA
3	C	5	KIAA1718	JHDM1D KDM7 KIAA1718	KDM7_HUMAN	Q6ZMT4	JHDM1D
3	D	5	JMJD1C	JMJD1C JHDM2C KIAA1380 TRIP8	JHD2C_HUMAN	Q15652	JMJD1C
3	E	5	JARID1B	KDM5B JARID1B PLU1 RBBP2H1	KDM5B_HUMAN	Q9UGL1	KDM5B
3	F	5	JARID1C	KDM5C DXS1272E JARID1C SMCX XE169	KDM5C_HUMAN	P41229	KDM5C
3	G	5	MBD5	MBD5 KIAA1461	MBD5_HUMAN	Q9P267	MBD5
3	H	5	MBD6	MBD6 KIAA1887	MBD6_HUMAN	Q96DN6	MBD6
3	A	6	HJURP	HJURP FAKTS FLEG1 URLC9	HJURP_HUMAN	Q8NCD3	HJURP
3	B	6	SMARCA3	HLTF HIP116A RNF80 SMARCA3 SNF2L3 ZBU1	HLTF_HUMAN	Q14527	HLTF
3	C	6	JMJD4	JMJD4	JMJD4_HUMAN	Q9H9V9	JMJD4
3	D	6	JMJD5	KDM8 JMJD5	KDM8_HUMAN	Q8N371	KDM8
3	E	6	JARID1D	KDM5D HY HYA JARID1D KIAA0234 SMCY	KDM5D_HUMAN	Q9BY66	KDM5D
3	F	6	UTX	KDM6A UTX	KDM6A_HUMAN	O15550	KDM6A
3	G	6	MBIP	MBIP BM-015	MBIP1_HUMAN	Q9NS73	MBIP
3	H	6	MBTD1	MBTD1	MBTD1_HUMAN	Q05BQ5	MBTD1
3	A	7	HR	HR	HAIR_HUMAN	O43593	HR
3	B	7	HSPBAP1	HSPBAP1 PASS1	HBAP1_HUMAN	Q96EW2	HSPBAP1
3	C	7	PTDSR	JMJD6 KIAA0585 PTDSR	JMJD6_HUMAN	Q6NYC1	JMJD6
3	D	7	JMJD7	JMJD7	JMJD7_HUMAN	P0C870	JMJD7
3	E	7	JMJD3	KDM6B JMJD3 KIAA0346	KDM6B_HUMAN	O15054	KDM6B
3	F	7	KIAA0892	MAU2 KIAA0892 SCC4	SCC4_HUMAN	Q9Y6X3	MAU2
3	G	7	MCRS1	MCRS1 INO80Q MSP58	MCRS1_HUMAN	Q96EZ8	MCRS1
3	H	7	MDC1	MDC1 KIAA0170 NFBD1	MDC1_HUMAN	Q14676	MDC1
3	A	8	ING1	ING1	ING1_HUMAN	Q9UK53	ING1
3	B	8	ING1L	ING2 ING1L	ING2_HUMAN	Q9H160	ING2
3	C	8	JMJD8	JMJD8 C16orf20 PP14397	JMJD8_HUMAN	Q96S16	JMJD8
3	D	8	GCN5L2	KAT2A GCN5 GCN5L2 HGCN5	KAT2A_HUMAN	Q92830	KAT2A
3	E	8	KIAA1045	KIAA1045	K1045_HUMAN	Q9UPV7	KIAA1045
3	F	8	L3MBTL	L3MBTL1 KIAA0681 L3MBT L3MBTL	LMBL1_HUMAN	Q9Y468	L3MBTL1
3	G	8	FLJ11730	MEAF6 C1orf149 EAF6	EAF6_HUMAN	Q9HAF1	MEAF6
3	H	8	MECP2	MECP2	MECP2_HUMAN	P51608	MECP2
3	A	9	ING3	ING3 HSPC301	ING3_HUMAN	Q9NXR8	ING3
3	B	9	ING4	ING4 My036	ING4_HUMAN	Q9UNL4	ING4

3	C	9	PCAF	KAT2B PCAF	KAT2B_HUMAN	Q92831	KAT2B
3	D	9	HTATIP	KAT5 HTATIP TIP60	KAT5_HUMAN	Q92993	KAT5
3	E	9	L3MBTL2	L3MBTL2	LMBL2_HUMAN	Q969R5	L3MBTL2
3	F	9	L3MBTL3	L3MBTL3 KIAA1798 MBT1	LMBL3_HUMAN	Q96JM7	L3MBTL3
3	G	9	MEN1	MEN1 SCG2	MEN1_HUMAN	O00255	MEN1
3	H	9	MGEA5	MGEA5 HEXC KIAA0679 MEA5	NCOAT_HUMAN	O60502	MGEA5
3	A	10	ING5	ING5	ING5_HUMAN	Q8WYH8	ING5
3	B	10	INOC1	INO80 INO80A INOC1 KIAA1259	INO80_HUMAN	Q9ULG1	INO80
3	C	10	AOF2	KDM1A AOF2 KDM1 KIAA0601 LSD1	KDM1A_HUMAN	O60341	KDM1A
3	D	10	FBXL11	KDM2A CXXC8 FBL7 FBXL11 JHDM1A KIAA1004	KDM2A_HUMAN	Q9Y2K7	KDM2A
3	E	10	L3MBTL4	L3MBTL4	LMBL4_HUMAN	Q8NA19	L3MBTL4
3	F	10	LBR	LBR	LBR_HUMAN	Q14739	LBR
3	G	10	MINA	MINA MDIG MINA53 NO52	MINA_HUMAN	Q8IUF8	MINA
3	H	10	MIS12	MIS12	MIS12_HUMAN	Q9H081	MIS12
3	A	11	ZNHIT4	INO80B HMGA1L4 PAPA1 ZNHIT4	IN80B_HUMAN	Q9C086	INO80B
3	B	11	C18ORF37	INO80C C18orf37	IN80C_HUMAN	Q6PI98	INO80C
3	C	11	FBXL10	KDM2B CXXC2 FBL10 FBXL10 JHDM1B PCCX2	KDM2B_HUMAN	Q8NHM5	KDM2B
3	D	11	JMJD1A	KDM3A JHDM2A JMJD1 JMJD1A KIAA0742 TSGA	KDM3A_HUMAN	Q9Y4C1	KDM3A
3	E	11	LIN9	LIN9 BARA TGS	LIN9_HUMAN	Q5TKA1	LIN9
3	F	11	MAP3K7	MAP3K7 TAK1	M3K7_HUMAN	O43318	MAP3K7
3	G	11	KLIP1	YIPF3 C6orf109 KLIP1	YIPF3_HUMAN	Q9GZM5	MLF1IP
3	H	11	MLH1	MLH1 COCA2	MLH1_HUMAN	P40692	MLH1
4	A	2	MLL	MLL ALL1 CXXC7 HRX HTRX KMT2A MLL1 TRX1	MLL1_HUMAN	Q03164	KMT2A
4	B	2	MLL2	MLL2 ALR KMT2D MLL4	MLL2_HUMAN	O14686	KMT2D
4	C	2	MYST4	KAT6B KIAA0383 MORF MOZ2 MYST4	KAT6B_HUMAN	Q8WYB5	KAT6B
4	D	2	NAP1L1	NAP1L1 NRP	NP1L1_HUMAN	P55209	NAP1L1
4	E	2	NFX1	NFX1 NFX2	NFX1_HUMAN	Q12986	NFX1
4	F	2	NIPBL	NIPBL IDN3	NIPBL_HUMAN	Q6KC79	NIPBL
4	G	2	H1FX	H1FX	H1X_HUMAN	Q92522	H1FX
4	H	2	FLJ21308	PARP8	PARP8_HUMAN	Q8N3A8	PARP8
4	A	3	MLL3	MLL3 HALR KIAA1506 KMT2C	MLL3_HUMAN	Q8NEZ4	KMT2C
4	B	3	MLL4	WBP7 HRX2 KIAA0304 KMT2B MLL2 MLL4 TRX2	MLL4_HUMAN	Q9UMN6	KMT2B
4	C	3	NAP1L2	NAP1L2 BPX	NP1L2_HUMAN	Q9ULW6	NAP1L2
4	D	3	NAP1L3	NAP1L3 BNAP	NP1L3_HUMAN	Q99457	NAP1L3
4	E	3	NPM1	NPM1 NPM	NPM_HUMAN	P06748	NPM1
4	F	3	NPM2	NPM2	NPM2_HUMAN	Q86SE8	NPM2
4	G	3	BAL	CEL BAL	CEL_HUMAN	P19835	PARP9
4	H	3	PAXIP1L	PAXIP1 PAXIP1L PTIP CAGF28	PAXI1_HUMAN	Q6ZW49	PAXIP1
4	A	4	MLL5	MLL5 KMT2E	MLL5_HUMAN	Q8IZD2	KMT2E
4	B	4	MLLT10	MLLT10 AF10	AF10_HUMAN	P55197	MLLT10
4	C	4	NAP1L4	NAP1L4 NAP2	NP1L4_HUMAN	Q99733	NAP1L4
4	D	4	NAP1L5	NAP1L5 DRLM	NP1L5_HUMAN	Q96NT1	NAP1L5

4	E	4	NPM3	NPM3	NPM3_HUMAN	O75607	NPM3
4	F	4	NPTXR	NPTXR	NPTXR_HUMAN	O95502	NPTXR
4	G	4	PB1	PBRM1 BAF180 PB1	PB1_HUMAN	Q86U86	PBRM1
4	H	4	PCGF1	PCGF1 NSPC1 RNF68	PCGF1_HUMAN	Q9BSM1	PCGF1
4	A	5	MLLT6	MLLT6 AF17	AF17_HUMAN	P55198	MLLT6
4	B	5	MORF4L1	MORF4L1 MRG15 FWP006 HSPC008 HSPC061 PP368	MO4L1_HUMAN	Q9UBU8	MORF4L1
4	C	5	CNAP1	CEP250 CEP2 CNAP1	CP250_HUMAN	Q9BV73	NCAPD2
4	D	5	KIAA0056	NCAPD3 CAPD3 KIAA0056	CNDD3_HUMAN	P42695	NCAPD3
4	E	5	NSD1	NSD1 ARA267 KMT3B	NSD1_HUMAN	Q96L73	NSD1
4	F	5	OGT	OGT	OGT1_HUMAN	O15294	OGT
4	G	5	PCGF2	PCGF2 MEL18 RNF110 ZNF144	PCGF2_HUMAN	P35227	PCGF2
4	H	5	PHC1	PHC1 EDR1 PH1	PHC1_HUMAN	P78364	PHC1
4	A	6	MORF4L2	MORF4L2 KIAA0026 MRGX	MO4L2_HUMAN	Q15014	MORF4L2
4	B	6	HSMPP8	MPHOSPH8 MPP8	MPP8_HUMAN	Q99549	MPHOSPH8
4	C	6	HCAP-G	NCAPG CAPG NYMEL3	CND3_HUMAN	Q9BPX3	CCND3
4	D	6	MTB	NCAPG2 LUZP5	CNDG2_HUMAN	Q86XI2	MT1JP
4	E	6	OIP5	OIP5 MIS18B	MS18B_HUMAN	O43482	OIP5
4	F	6	ORC1L	ORC1 ORC1L PARC1	ORC1_HUMAN	Q13415	ORC1
4	G	6	PHC2	PHC2 EDR2 PH2	PHC2_HUMAN	Q8IXK0	PHC2
4	H	6	PHF1	PHF1 PCL1	PHF1_HUMAN	O43189	PHF1
4	A	7	MSH6	MSH6 GTBP	MSH6_HUMAN	P52701	MSH6
4	B	7	MSL3L1	MSL3 MSL3L1	MS3L1_HUMAN	Q8N5Y2	MSL3
4	C	7	BRRN1	NCAPH BRRN BRRN1 CAPH KIAA0074	CND2_HUMAN	Q15003	NCAPH
4	D	7	NCAPH2	NCAPH2 CAPH2	CNDH2_HUMAN	Q6IBW4	NCAPH2
4	E	7	PARP1_2	PARP1 ADPRT PPOL	PARP1_HUMAN	P09874	PARP1_2
4	F	7	PARP10	PARP10	PAR10_HUMAN	Q53GL7	PARP10
4	G	7	PHF10	PHF10 BAF45A	PHF10_HUMAN	Q8WUB8	PHF10
4	H	7	PHF11	PHF11 BCAP	PHF11_HUMAN	Q9UIL8	PHF11
4	A	8	MTA1	MTA1	MTA1_HUMAN	Q13330	MTA1
4	B	8	MTA2	MTA2 MTA1L1 PID	MTA2_HUMAN	O94776	MTA2
4	C	8	NCOA1_1	NCOA1 BHLHE74 SRC1	NCOA1_HUMAN	Q15788	NCOA1
4	D	8	NCOA1_2	NCOA1 BHLHE74 SRC1	NCOA1_HUMAN	Q15788	NCOA1_2
4	E	8	PARP11	PARP11 C12orf6	PAR11_HUMAN	Q9NR21	PARP11
4	F	8	ZC3HDC1	PARP12 ZC3HDC1	PAR12_HUMAN	Q9H0J9	PARP12
4	G	8	PHF12	PHF12 KIAA1523	PHF12_HUMAN	Q96QT6	PHF12
4	H	8	PHF13	PHF13	PHF13_HUMAN	Q86YI8	PHF13
4	A	9	MTA3	MTA3 KIAA1266	MTA3_HUMAN	Q9BTC8	MTA3
4	B	9	M96	MTF2 PCL2	MTF2_HUMAN	Q9Y483	MTF2
4	C	9	NCOA2	NCOA2 BHLHE75 TIF2	NCOA2_HUMAN	Q15596	NCOA2
4	D	9	NCOA3	NCOA3 AIB1 BHLHE42 RAC3 TRAM1	NCOA3_HUMAN	Q9Y6Q9	NCOA3
4	E	9	KIAA1268	PARP14 BAL2 KIAA1268	PAR14_HUMAN	Q460N5	PARP14
4	F	9	FLJ40597	PARP15 BAL3	PAR15_HUMAN	Q460N3	PARP15
4	G	9	PHF14	PHF14 KIAA0783	PHF14_HUMAN	O94880	PHF14
4	H	9	PHF15	PHF15 JADE2 KIAA0239	JADE2_HUMAN	Q9NQC1	PHF15
4	A	10	MYO18A	MYO18A KIAA0216 MYSPDZ	MY18A_HUMAN	Q92614	MYO18A

4	B	10	MYST1	KAT8 MOF MYST1 PP7073	KAT8_HUMAN	Q9H7Z6	KAT8
4	C	10	NCOA6	NCOA6 AIB3 KIAA0181 RAP250 TRBP	NCOA6_HUMAN	Q14686	NCOA6
4	D	10	NCOR1	NCOR1 KIAA1047	NCOR1_HUMAN	O75376	NCOR1
4	E	10	FLJ20509	PARP16 C15orf30	PAR16_HUMAN	Q8N5Y8	PARP16
4	F	10	PARP2	PARP2 ADPRT2 ADPRTL2	PARP2_HUMAN	Q9UGN5	PARP2
4	G	10	PHF16	PHF16 JADE3 KIAA0215	JADE3_HUMAN	Q92613	PHF16
4	H	10	PHF17_2	PHF17 JADE1 KIAA1807	JADE1_HUMAN	Q6IE81	PHF17_2
4	A	11	MYST2	KAT7 HBO1 HBOa MYST2	KAT7_HUMAN	O95251	KAT7
4	B	11	MYST3	KAT6A MOZ MYST3 RUNXBP2 ZNF220	KAT6A_HUMAN	Q92794	KAT6A
4	C	11	NCOR2	NCOR2 CTG26	NCOR2_HUMAN	Q9Y618	NCOR2
4	D	11	NFRKB	NFRKB INO80G	NFRKB_HUMAN	Q6P4R8	NFRKB
4	E	11	ADPRTL3	PARP3 ADPRT3 ADPRTL3	PARP3_HUMAN	Q9Y6F1	PARP3
4	F	11	PARP4	PARP4 ADPRTL1 KIAA0177 PARPL	PARP4_HUMAN	Q9UKK3	PARP4
4	G	11	PHF17_1	PHF17 JADE1 KIAA1807	JADE1_HUMAN	Q6IE81	PHF17_1
4	H	11	PHF19	PHF19 PCL3	PHF19_HUMAN	Q5T6S3	PHF19
5	A	2	PHF2	PHF2 KIAA0662	PHF2_HUMAN	O75151	PHF2
5	B	2	PHF20	PHF20 C20orf104 GLEA2 HCA58 NZF TZP	PHF20_HUMAN	Q9BVI0	PHF20
5	C	2	PPP2R5E	PPP2R5E	2A5E_HUMAN	Q16537	PPP2R5E
5	D	2	PRDM1	PRDM1 BLIMP1	PRDM1_HUMAN	O75626	PRDM1
5	E	2	PRMT7	PRMT7 KIAA1933	ANM7_HUMAN	Q9NVM4	PRMT7
5	F	2	HRMT1L4	PRMT8 HRMT1L3 HRMT1L4	ANM8_HUMAN	Q9NR22	PRMT8
5	G	2	RNF17	RNF17 TDRD4	RNF17_HUMAN	Q9BXT8	RNF17
5	H	2	RNF2	RNF2 BAP1 DING HIPI3 RING1B	RING2_HUMAN	Q99496	RNF2
5	A	3	PHF21A	PHF21A BHC80 KIAA1696 BM-006	PF21A_HUMAN	Q96BD5	PHF21A
5	B	3	PHF21B	PHF21B KIAA1661	PF21B_HUMAN	Q96EK2	PHF21B
5	C	3	PRDM11	PRDM11 PFM8	PRD11_HUMAN	Q9NQV5	PRDM11
5	D	3	PRDM12	PRDM12 PFM9	PRD12_HUMAN	Q9H4Q4	PRDM12
5	E	3	PSIP1	PSIP1 DFS70 LEDGF PSIP2	PSIP1_HUMAN	O75475	PSIP1
5	F	3	KIAA1935	PWWP2A KIAA1935 MST101	PWP2A_HUMAN	Q96N64	PWWP2A
5	G	3	RNF20	RNF20 BRE1A	BRE1A_HUMAN	Q5VTR2	RNF20
5	H	3	RNF40	RNF40 BRE1B KIAA0661	BRE1B_HUMAN	O75150	RNF40
5	A	4	PHF23	PHF23	PHF23_HUMAN	Q9BUL5	PHF23
5	B	4	PHF3	PHF3 KIAA0244	PHF3_HUMAN	Q92576	PHF3
5	C	4	PRDM13	PRDM13 PFM10	PRD13_HUMAN	Q9H4Q3	PRDM13
5	D	4	PRDM14	PRDM14	PRD14_HUMAN	Q9GZV8	PRDM14
5	E	4	PWWP2	PWWP2B PWWP2	PWP2B_HUMAN	Q6NUJ5	PWWP2B
5	F	4	PYGO1	PYGO1	PYGO1_HUMAN	Q9Y3Y4	PYGO1
5	G	4	RNMT	RNMT KIAA0398	MCES_HUMAN	O43148	RNMT
5	H	4	FLJ10581	RNMTL1 HC90	RMTL1_HUMAN	Q9HC36	RNMTL1
5	A	5	PHF6	PHF6 KIAA1823	PHF6_HUMAN	Q8IWS0	PHF6
5	B	5	PHF7	PHF7	PHF7_HUMAN	Q9BWX1	PHF7
5	C	5	PRDM15	PRDM15 C21orf83 ZNF298	PRD15_HUMAN	P57071	PRDM15
5	D	5	PRDM16	PRDM16 KIAA1675 MEL1 PFM13	PRD16_HUMAN	Q9HAZ2	PRDM16
5	E	5	PYGO2	PYGO2 PP7910	PYGO2_HUMAN	Q9BRQ0	PYGO2
5	F	5	RAD21	RAD21 HR21 KIAA0078 NXP1	RAD21_HUMAN	O60216	RAD21

5	G	5	HBXAP	RSF1 HBXAP XAP8	RSF1_HUMAN	Q96T23	RSF1
5	H	5	RUVBL1	RUVBL1 INO80H NMP238 TIP49 TIP49A	RUVB1_HUMAN	Q9Y265	RUVBL1
5	A	6	PHF8	PHF8 KIAA1111 ZNF422	PHF8_HUMAN	Q9UPP1	PHF8
5	B	6	PHIP	PHIP WDR11	PHIP_HUMAN	Q8WWQ0	PHIP
5	C	6	PRDM2	PRDM2 KMT8 RIZ	PRDM2_HUMAN	Q13029	PRDM2
5	D	6	PRDM4	PRDM4 PFM1	PRDM4_HUMAN	Q9UKN5	PRDM4
5	E	6	PIR51	RAD51AP1 PIR51	R51A1_HUMAN	Q96B01	RAD51AP1
5	F	6	RAD54B	RAD54B	RA54B_HUMAN	Q9Y620	RAD54B
5	G	6	RUVBL2	RUVBL2 INO80J TIP48 TIP49B CGI-46	RUVB2_HUMAN	Q9Y230	RUVBL2
5	H	6	SAP130	SAP130	SP130_HUMAN	Q9H0E3	SAP130
5	A	7	KIAA1542	PHRF1 KIAA1542	PHRF1_HUMAN	Q9P1Y6	PHRF1
5	B	7	PLK1	PLK1 PLK	PLK1_HUMAN	P53350	PLK1
5	C	7	PRDM5	PRDM5 PFM2	PRDM5_HUMAN	Q9NQX1	PRDM5
5	D	7	PRDM7	PRDM7 PFM4	PRDM7_HUMAN	Q9NQW5	PRDM7
5	E	7	RAD54L	RAD54L RAD54A	RAD54_HUMAN	Q92698	RAD54L
5	F	7	RAI1	RAI1 KIAA1820	RAI1_HUMAN	Q7Z5J4	RAI1
5	G	7	SCMH1	SCMH1	SCMH1_HUMAN	Q96GD3	SCMH1
5	H	7	SCML2	SCML2	SCML2_HUMAN	Q9UQR0	SCML2
5	A	8	POLE3	POLE3 CHRAC17	DPOE3_HUMAN	Q9NRF9	POLE3
5	B	8	POLE4	POLE4	DPOE4_HUMAN	Q9NR33	POLE4
5	C	8	PRDM8	PRDM8 PFM5	PRDM8_HUMAN	Q9NQV8	PRDM8
5	D	8	PRDM9	PRDM9 PFM6	PRDM9_HUMAN	Q9NQV7	PRDM9
5	E	8	RARA	RARA NR1B1	RARA_HUMAN	P10276	RARA
5	F	8	RBBP4	RBBP4 RBAP48	RBBP4_HUMAN	Q09028	RBBP4
5	G	8	SERPIND1	SERPIND1 HCF2	HEP2_HUMAN	P05546	SERPIND1
5	H	8	SETD1A	SETD1A KIAA0339 KMT2F SET1 SET1A	SET1A_HUMAN	O15047	SETD1A
5	A	9	PPP2CA	PPP2CA	PP2AA_HUMAN	P67775	PPP2CA
5	B	9	PPP2R1B	PPP2R1B	2AAB_HUMAN	P30154	PPP2R1B
5	C	9	PRG4	LPPR2 PRG4	LPPR2_HUMAN	Q96GM1	PRG4
5	D	9	HRMT1L2	PRMT1 HMT2 HRMT1L2 IR1B4	ANM1_HUMAN	Q99873	PRMT1
5	E	9	RBBP5	RBBP5 RBQ3	RBBP5_HUMAN	Q15291	RBBP5
5	F	9	RBBP6	RBBP6 P2PR PACT RBQ1 My038	RBBP6_HUMAN	Q7Z6E9	RBBP6
5	G	9	HYPB	SETD2 HIF1 HYPB KIAA1732 KMT3A SET2 HSPC069	SETD2_HUMAN	Q9BYW2	SETD2
5	H	9	C14ORF154	SETD3 C14orf154	SETD3_HUMAN	Q86TU7	SETD3
5	A	10	PPP2R5A	PPP2R5A	2A5A_HUMAN	Q15172	PPP2R5A
5	B	10	PPP2R5B	PPP2R5B	2A5B_HUMAN	Q15173	PPP2R5B
5	C	10	HRMT1L1	PRMT2 HMT1 HRMT1L1	ANM2_HUMAN	P55345	PRMT2
5	D	10	HRMT1L3	PRMT3 HRMT1L3	ANM3_HUMAN	O60678	PRMT3
5	E	10	RBBP7	RBBP7 RBAP46	RBBP7_HUMAN	Q16576	RBBP7
5	F	10	CHC1	RCC1 CHC1	RCC1_HUMAN	P18754	RCC1
5	G	10	C21ORF18	SETD4 C21orf18 C21orf27	SETD4_HUMAN	Q9NVD3	SETD4
5	H	10	SETD5	SETD5 KIAA1757	SETD5_HUMAN	Q9C0A6	SETD5
5	A	11	PPP2R5D	PPP2R5D	2A5D_HUMAN	Q14738	PPP2R5D
5	B	11	PPP2R5C	PPP2R5C KIAA0044	2A5G_HUMAN	Q13362	PPP2R5C
5	C	11	HRMT1L6_2	PRMT6 HRMT1L6	ANM6_HUMAN	Q96LA8	PRMT6_2

5	D	11	HRMT1L6_3	PRMT6 HRMT1L6	ANM6_HUMAN	Q96LA8	PRMT6_3
5	E	11	RERE	RERE ARG ARP ATN1L KIAA0458	RERE_HUMAN	Q9P2R6	RERE
5	F	11	RING1	RING1 RNF1	RING1_HUMAN	Q06587	RING1
5	G	11	FLJ21148	SETD6	SETD6_HUMAN	Q8TBK2	SETD6
5	H	11	SET7	SETD7 KIAA1717 KMT7 SET7 SET9	SETD7_HUMAN	Q8WTS6	SETD7
6	A	2	SET8	SETD8 KMT5A PRSET7 SET07 SET8	SETD8_HUMAN	Q9NQR1	SETD8
6	B	2	SETDB1	SETDB1 KIAA0067 KMT1E	SETB1_HUMAN	Q15047	SETDB1
6	C	2	SMARCA4	SMARCA4 BAF190A BRG1 SNF2B SNF2L4	SMCA4_HUMAN	P51532	SMARCA4
6	D	2	SMARCA5	SMARCA5 SNF2H WCRF135	SMCA5_HUMAN	O60264	SMARCA5
6	E	2	SMYD5	SMYD5 RAI15	SMYD5_HUMAN	Q6GMV2	SMYD5
6	F	2	SND1	SND1 TDRD11	SND1_HUMAN	Q7KZF4	SND1
6	G	2	TADA2L	TADA2A TADA2L KL04P	TAD2A_HUMAN	O75478	TADA2A
6	H	2	MGC21874	TADA2B ADA2B	TAD2B_HUMAN	Q86TJ2	TADA2B
6	A	3	SETDB2	SETDB2 C13orf4 CLLD8 KMT1F	SETB2_HUMAN	Q96T68	SETDB2
6	B	3	SETMAR	SETMAR	SETMR_HUMAN	Q53H47	SETMAR
6	C	3	SMARCAL1	SMARCAL1 HARP	SMAL1_HUMAN	Q9NZC9	SMARCAL1
6	D	3	SMARCB1	SMARCB1 BAF47 INI1 SNF5L1	SNF5_HUMAN	Q12824	SMARCB1
6	E	3	SP100	SP100	SP100_HUMAN	P23497	SP100
6	F	3	SP110	SP110	SP110_HUMAN	Q9HB58	SP110
6	G	3	TADA3L	TADA3 ADA3 TADA3L	TADA3_HUMAN	O75528	TADA3
6	H	3	TAF1	TAF1 BA2R CCG1 CCGS TAF2A	TAF1_HUMAN	P21675	TAF1
6	A	4	SFMBT1	SFMBT1 RU1	SMBT1_HUMAN	Q9UHJ3	SFMBT1
6	B	4	SFMBT2	SFMBT2 KIAA1617	SMBT2_HUMAN	Q5VUG0	SFMBT2
6	C	4	SMARCC1	SMARCC1 BAF155	SMRC1_HUMAN	Q92922	SMARCC1
6	D	4	SMARCC2	SMARCC2 BAF170	SMRC2_HUMAN	Q8TAQ2	SMARCC2
6	E	4	SP140	SP140 LYSP100	LY10_HUMAN	Q13342	SP140
6	F	4	SHARP	SPEN KIAA0929 MINT SHARP	MINT_HUMAN	Q96T58	SPEN
6	G	4	TAF10	TAF10 TAF2A TAF2H TAFII30	TAF10_HUMAN	Q12962	TAF10
6	H	4	TAF12	TAF12 TAF15 TAF2J TAFII20	TAF12_HUMAN	Q16514	TAF12
6	A	5	SGOL1	SGOL1 SGO1	SGOL1_HUMAN	Q5FBB7	SGOL1
6	B	5	SGOL2	SGOL2	SGOL2_HUMAN	Q562F6	SGOL2
6	C	5	SMARCD1	SMARCD1 BAF60A	SMRD1_HUMAN	Q96GM5	SMARCD1
6	D	5	SMARCD2	SMARCD2 BAF60B PRO2451	SMRD2_HUMAN	Q92925	SMARCD2
6	E	5	SRCAP	SRCAP KIAA0309	SRCAP_HUMAN	Q6ZRS2	SRCAP
6	F	5	SS18	SS18 SSXT SYT	SSXT_HUMAN	Q15532	SS18
6	G	5	TAF1L	TAF1L	TAF1L_HUMAN	Q8IZX4	TAF1L
6	H	5	TAF2	TAF2 CIF150 TAF2B	TAF2_HUMAN	Q6P1X5	TAF2
6	A	6	SHMT2	SHMT2	GLYM_HUMAN	P34897	SHMT2
6	B	6	SHPRH	SHPRH KIAA2023	SHPRH_HUMAN	Q149N8	SHPRH
6	C	6	SMARCD3	SMARCD3 BAF60C	SMRD3_HUMAN	Q6STE5	SMARCD3
6	D	6	SMARCE1	SMARCE1 BAF57	SMCE1_HUMAN	Q969G3	SMARCE1
6	E	6	STAG1	STAG1 SA1	STAG1_HUMAN	Q8WVM7	STAG1
6	F	6	STAG2	STAG2 SA2	STAG2_HUMAN	Q8N3U4	STAG2
6	G	6	TAF3	TAF3	TAF3_HUMAN	Q5VWG9	TAF3
6	H	6	TAF4	TAF4 TAF2C TAF2C1 TAF4A TAFII130 TAFII135	TAF4_HUMAN	O00268	TAF4

6	A	7	SIN3A	SIN3A	SIN3A_HUMAN	Q96ST3	SIN3A
6	B	7	SIRT1	SIRT1 SIR2L1	SIR1_HUMAN	Q96EB6	SIRT1
6	C	7	SMC1L1	SMC1A DXS423E KIAA0178 SB1.8 SMC1 SMC1L1	SMC1A_HUMAN	Q14683	SMC1A
6	D	7	SMC1L2	SMC1B SMC1L2	SMC1B_HUMAN	Q8NDV3	SMC1B
6	E	7	STAG3	STAG3	STAG3_HUMAN	Q9UJ98	STAG3
6	F	7	STK31	STK31 SGK396	STK31_HUMAN	Q9BXU1	STK31
6	G	7	TAF5	TAF5 TAF2D	TAF5_HUMAN	Q15542	TAF5
6	H	7	TAF5L	TAF5L PAF65B	TAF5L_HUMAN	O75529	TAF5L
6	A	8	SIRT2	SIRT2 SIR2L SIR2L2	SIR2_HUMAN	Q8IXJ6	SIRT2
6	B	8	SIRT3	SIRT3 SIR2L3	SIR3_HUMAN	Q9NTG7	SIRT3
6	C	8	SMC2L1	SMC2 CAPE SMC2L1 PRO0324	SMC2_HUMAN	O95347	SMC2
6	D	8	CSPG6	SMC3 BAM BMH CSPG6 SMC3L1	SMC3_HUMAN	Q9UQE7	SMC3
6	E	8	SUPT3H	SUPT3H SPT3	SUPT3_HUMAN	O75486	SUPT3H
6	F	8	STAF65γ	SUPT7L KIAA0764	ST65G_HUMAN	O94864	SUPT7L
6	G	8	TAF6	TAF6 TAF2E TAFII70	TAF6_HUMAN	P49848	TAF6
6	H	8	TAF6L	TAF6L PAF65A	TAF6L_HUMAN	Q9Y6J9	TAF6L
6	A	9	SIRT4	SIRT4 SIR2L4	SIR4_HUMAN	Q9Y6E7	SIRT4
6	B	9	SIRT5	SIRT5 SIR2L5	SIR5_HUMAN	Q9NXA8	SIRT5
6	C	9	SMC4L1	SMC4 CAPC SMC4L1	SMC4_HUMAN	Q9NTJ3	SMC4
6	D	9	SMNDC1	SMNDC1 SMNR SPF30	SPF30_HUMAN	O75940	SMNDC1
6	E	9	SUV39H1	SUV39H1 KMT1A SUV39H	SUV91_HUMAN	O43463	SUV39H1
6	F	9	SUV39H2	SUV39H2 KMT1B	SUV92_HUMAN	Q9H5I1	SUV39H2
6	G	9	TAF7	TAF7 TAF2F TAFII55	TAF7_HUMAN	Q15545	TAF7
6	H	9	TAF9L	TAF9B TAF9L	TAF9B_HUMAN	Q9HBM6	TAF9B
6	A	10	SIRT6	SIRT6 SIR2L6	SIR6_HUMAN	Q8N6T7	SIRT6
6	B	10	SIRT7	SIRT7 SIR2L7	SIR7_HUMAN	Q9NRC8	SIRT7
6	C	10	SMYD1	SMYD1	SMYD1_HUMAN	Q8NB12	SMYD1
6	D	10	SMYD2	SMYD2 KMT3C	SMYD2_HUMAN	Q9NRG4	SMYD2
6	E	10	SUV420H1	SUV420H1 KMT5B CGI-85	SV421_HUMAN	Q4FZB7	SUV420H1
6	F	10	SUV420H2	SUV420H2 KMT5C PP7130	SV422_HUMAN	Q86Y97	SUV420H2
6	G	10	TARBP1	TARBP1 TRM3 TRP185	TARB1_HUMAN	Q13395	TARBP1
6	H	10	TBL1X	TBL1X TBL1	TBL1X_HUMAN	O60907	TBL1X
6	A	11	SMARCA1	SMARCA1 SNF2L SNF2L1	SMCA1_HUMAN	P28370	SMARCA1
6	B	11	SMARCA2	SMARCA2 BAF190B BRM SNF2A SNF2L2	SMCA2_HUMAN	P51531	SMARCA2
6	C	11	SMYD3	SMYD3 ZMYND1 ZNFN3A1	SMYD3_HUMAN	Q9H7B4	SMYD3
6	D	11	SMYD4	SMYD4 KIAA1936	SMYD4_HUMAN	Q8IYR2	SMYD4
6	E	11	JJAZ1	SUZ12 CHET9 JJAZ1 KIAA0160	SUZ12_HUMAN	Q15022	SUZ12
6	F	11	STAF42	TADA1 TADA1L	TADA1_HUMAN	Q96BN2	TADA1
6	G	11	TBL1XR1	TBL1XR1 IRA1 TBLR1	TBL1R_HUMAN	Q9BZK7	TBL1XR1
6	H	11	TCF19	TCF19 SC1	TCF19_HUMAN	Q9Y242	TCF19
7	A	2	TCF20	TCF20 KIAA0292 SPBP	TCF20_HUMAN	Q9UGU0	TCF20
7	B	2	TDRD1	TDRD1	TDRD1_HUMAN	Q9BXT4	TDRD1
7	C	2	C6ORF75	TRMT11 C6orf75 MDS024	TRM11_HUMAN	Q7Z4G4	TRMT11
7	D	2	TRRAP	TRRAP PAF400	TRRAP_HUMAN	Q9Y4A5	TRRAP
7	E	2	GAS41	YEATS4 GAS41	YETS4_HUMAN	O95619	YEATS4

7	F	2	YY1	YY1 INO80S	TYY1_HUMAN	P25490	YY1
7	G	2	Buffer	0	0	0	Buffer
7	H	2	Buffer	0	0	0	Buffer
7	A	3	TDRD10	TDRD10 UNQ9380/PRO34205	TDR10_HUMAN	Q5VZ19	TDRD10
7	B	3	FLJ13072	TDRD12 ECAT8	TDR12_HUMAN	Q587J7	TDRD12
7	C	3	TTF2	FOXE1 FKHL15 FOXE2 TITF2 TTF2	FOXE1_HUMAN	O00358	TTF2
7	D	3	NICE-4	UBAP2L KIAA0144 NICE4	UBP2L_HUMAN	Q14157	UBAP2L
7	E	3	ZBTB2	ZBTB2 KIAA1483 ZNF437	ZBTB2_HUMAN	Q8N680	ZBTB2
7	F	3	ZBTB33	ZBTB33 KAISO ZNF348	KAISO_HUMAN	Q86T24	ZBTB33
7	G	3	Buffer	0	0	0	Buffer
7	H	3	Buffer	0	0	0	Buffer
7	A	4	TDRD3	TDRD3	TDRD3_HUMAN	Q9H7E2	TDRD3
7	B	4	TDRD5	TDRD5	TDRD5_HUMAN	Q8NAT2	TDRD5
7	C	4	UBE2A	UBE2A RAD6A	UBE2A_HUMAN	P49459	UBE2A
7	D	4	UBE2B	UBE2B RAD6B	UBE2B_HUMAN	P63146	UBE2B
7	E	4	ZCWPW1	ZCWPW1	ZCPW1_HUMAN	Q9H0M4	ZCWPW1
7	F	4	LOC152098	ZCWPW2	ZCPW2_HUMAN	Q504Y3	ZCWPW2
7	G	4	Buffer	0	0	0	Buffer
7	H	4	Buffer	0	0	0	Buffer
7	A	5	TDRD6	TDRD6	TDRD6_HUMAN	O60522	TDRD6
7	B	5	TDRD7	TDRD7 PCTAIRE2BP	TDRD7_HUMAN	Q8NHU6	TDRD7
7	C	5	UCHL5	UCHL5 UCH37 AD-019 CGI-70	UCHL5_HUMAN	Q9Y5K5	UCHL5
7	D	5	UHRF1	UHRF1 ICBP90 NP95 RNF106	UHRF1_HUMAN	Q96T88	UHRF1
7	E	5	ZMYND11	ZMYND11 BS69	ZMY11_HUMAN	Q15326	ZMYND11
7	F	5	PRKCBP1	ZMYND8 KIAA1125 PRKCBP1 RACK7	PKCB1_HUMAN	Q9ULU4	ZMYND8
7	G	5	Buffer	0	0	0	Buffer
7	H	5	Buffer	0	0	0	Buffer
7	A	6	TDRD9	TDRD9 C14orf75	TDRD9_HUMAN	Q8NDG6	TDRD9
7	B	6	TDRKH	TDRKH TDRD2	TDRKH_HUMAN	Q9Y2W6	TDRKH
7	C	6	UHRF2	UHRF2 NIRF RNF107	UHRF2_HUMAN	Q96PU4	UHRF2
7	D	6	USP22	USP22 KIAA1063 USP3L	UBP22_HUMAN	Q9UPT9	USP22
7	E	6	ZZZ3	ZZZ3	ZZZ3_HUMAN	Q8IYH5	ZZZ3
7	F	6	Buffer	0	0	0	Buffer
7	G	6	Buffer	0	0	0	Buffer
7	H	6	Buffer	0	0	0	Buffer
7	A	7	TERF2IP	TERF2IP DRIP5 RAP1 PP8000	TE2IP_HUMAN	Q9NYB0	TERF2IP
7	B	7	TFPT	TFPT INO80F	TFPT_HUMAN	P0C1Z6	TFPT
7	C	7	UTY	UTY	UTY_HUMAN	O14607	UTY
7	D	7	TCFL1	VPS72 TCFL1 YL1	VPS72_HUMAN	Q15906	VPS72
7	E	7	Buffer	0	0	0	Buffer
7	F	7	Buffer	0	0	0	Buffer
7	G	7	Buffer	0	0	0	Buffer
7	H	7	Buffer	0	0	0	Buffer
7	A	8	THUMPD2	THUMPD2 C2orf8	THUM2_HUMAN	Q9BTF0	THUMPD2
7	B	8	KIAA1856	TNRC18 CAGL79 KIAA1856	TNC18_HUMAN	O15417	TNRC18

7	C	8	WDHD1	WDHD1 AND1	WDHD1_HUMAN	O75717	WDHD1
7	D	8	WDR5	WDR5 BIG3	WDR5_HUMAN	P61964	WDR5
7	E	8	Buffer	0	0	0	Buffer
7	F	8	Buffer	0	0	0	Buffer
7	G	8	Buffer	0	0	0	Buffer
7	H	8	Buffer	0	0	0	Buffer
7	A	9	TP53BP1	TP53BP1	TP53B_HUMAN	Q12888	TP53BP1
7	B	9	DNMT2	TRDMT1 DNMT2	TRDMT_HUMAN	O14717	TRDMT1
7	C	9	WDR82	WDR82 TMEM113 WDR82A UNQ9342/ PRO34047	WDR82_HUMAN	Q6UXN9	WDR82
7	D	9	WFDC5	WFDC5 PRG5 WAP1 UNQ3036/PRO9842	WFDC5_HUMAN	Q8TCV5	WFDC5
7	E	9	Buffer	0	0	0	Buffer
7	F	9	Buffer	0	0	0	Buffer
7	G	9	Buffer	0	0	0	Buffer
7	H	9	Buffer	0	0	0	Buffer
7	A	10	TIF1	TRIM24 RNF82 TIF1 TIF1A	TIF1A_HUMAN	O15164	TRIM24
7	B	10	TRIM28	TRIM28 KAP1 RNF96 TIF1B	TIF1B_HUMAN	Q13263	TRIM28
7	C	10	WHSC1	WHSC1 KIAA1090 MMSET NSD2 TRX5	NSD2_HUMAN	O96028	WHSC1
7	D	10	WHSC1L1	WHSC1L1 NSD3 DC28	NSD3_HUMAN	Q9BZ95	WHSC1L1
7	E	10	Buffer	0	0	0	Buffer
7	F	10	Buffer	0	0	0	Buffer
7	G	10	Buffer	0	0	0	Buffer
7	H	10	Buffer	0	0	0	Buffer
7	A	11	TRIM33	TRIM33 KIAA1113 RFG7 TIF1G	TRI33_HUMAN	Q9UPN9	TRIM33
7	B	11	KIAA0298	TRIM66 C11orf29 KIAA0298	TRI66_HUMAN	O15016	TRIM66
7	C	11	XRCC1	XRCC1	XRCC1_HUMAN	P18887	XRCC1
7	D	11	FLJ10201	YEATS2 KIAA1197	YETS2_HUMAN	Q9ULM3	YEATS2
7	E	11	Buffer	0	0	0	Buffer
7	F	11	Buffer	0	0	0	Buffer
7	G	11	Buffer	0	0	0	Buffer
7	H	11	Buffer	0	0	0	Buffer

Table S1: List of 529 siRNA targets used for CROSS. The custom siRNA library targeting human chromatin-associated proteins, including writers, erasers, readers and cofactors of known chromatin modifying complexes. In columns are given the ThermoGeneName (Gene name given by Thermo Scientific), AllGeneNames (All known gene names), UniProtGeneName (Gene name given by Uniprot), UniprotID (Uniprot Identification number) and HGNC (Official gene name by HGNC).

Chapter 3

Quantitative proteomics of the SMAD (SMA and MAD related) transcription factor family identifies Importin 5 as a Bone Morphogenetic Protein receptor-SMAD specific Importin.

Roy Baas[1], Ayestha Sijm[1], Hetty A.A.M. van Teeffelen[1], Robert van Es[1], Harmjan R. Vos[1] and H.Th. Marc Timmers[1]

Adapted from:
J Biol Chem. 2016 291, 24121-24132
doi: 10.1074/jbc.M116.748582

[1]Molecular Cancer Research and Stem Cells, Regenerative Medicine Center, Center for Molecular Medicine, University Medical Center Utrecht, Uppsalalaan 8, 3584 CT Utrecht, The Netherlands

Abstract

Gene-specific transcription factors (GSTFs) control gene transcription by DNA binding and specific protein complex recruitment, which regulates promoter accessibility for transcription initiation by RNA polymerase II. Mutations in the GSTFs *SMA and MAD related 2 (SMAD2)* and *SMAD4* are frequently associated with colon and rectal carcinomas. These proteins play an important role in Transforming Growth Factor-β (TGF-β) and Bone Morphogenetic Protein (BMP) signaling pathways controlling cell fate and proliferation.

To study the protein interactome of the SMAD protein family we generated a quantitative proteomics pipeline that allows for inducible expression of GFP-tagged SMAD proteins followed by affinity purification and quantitative mass spectrometry analysis. Data are available via ProteomeXchange with identifier PXD004529. The nuclear importin IPO5 was identified as a novel interacting protein of SMAD1. Overexpression of IPO5 in various cell lines specifically increases nuclear localization of BMP receptor activated SMADs (R-SMADs) confirming a functional relationship between IPO5 and BMP but not TGF-β R-SMADs. Finally, we provide evidence that variation in length of the lysine stretch of the nuclear localization sequence (NLS) is a determinant for importin specificity.

Introduction

Transcription regulation is a tightly controlled process that is influenced by the binding of GSTFs to specific DNA elements (306). These GSTF families are highly heterogeneous, spanning several hundreds of proteins like the C2H2-type zinc fingers (307) to the eight members of the SMAD protein family. Binding of GSTFs to DNA elements mediates recruitment of multi-subunit co-activator complexes like SET/MLL, Mediator, TFIID, or the BAF remodeler, each of which can specifically affect the chromatin microenvironment resulting in alteration of gene transcriptional states. Cancer genome-wide association studies (GWAS) have shown that GSTFs and chromatin-associated proteins are often mutated in various solid tumors types (252-254). In particular, mutations in *SMAD2* and *SMAD4* have been frequently associated with colon and rectal carcinomas (5.7% and 9.8% respectively) (252). The eight SMAD family members can be divided into the R-SMADs (SMAD1/2/3/5/9), I-SMADs (SMAD6/7) and Co-SMAD (SMAD4) (Fig. 1A). R-SMADs can additionally be subdivided into BMP (SMAD1/5/9) and TGF-β (SMAD2/3) subclasses depending on the ligand activating.

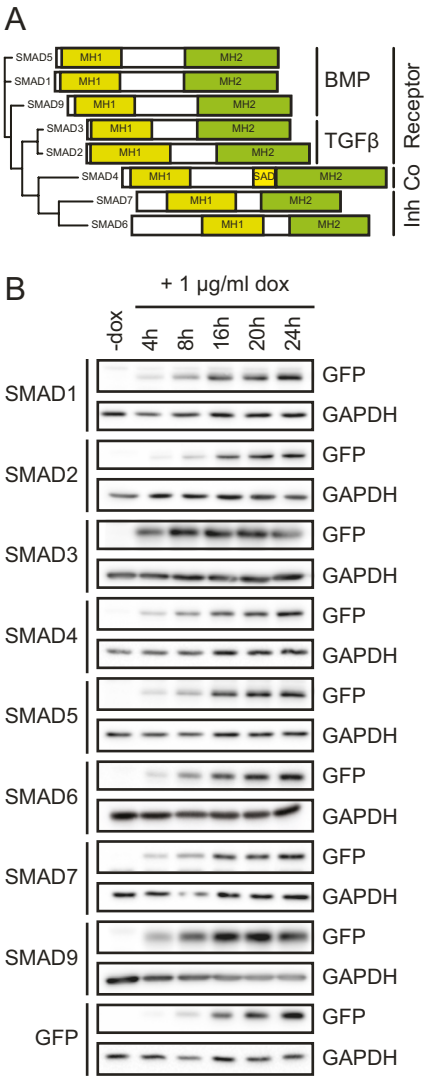


Figure 1. Generation of GFP-SMAD cell lines. (A) SMAD protein phylogeny. Clustering analysis and systemic representation of SMAD proteins with indicated domains (MH1; MH2; PY; SAD). SMAD categorization is indicated on the right. (B) SMAD cell line time series after doxycycline induction was analyzed on immunoblot. HeLa cell lines were induced with 1 µg/ml doxycycline for 4, 8, 16, 20 or 24 hours prior to harvesting or harvested without induction. Blots were probed using GFP or GAPDH antibodies.

SMAD signaling starts with the activation of the Activin receptor-like kinase (ALK) type II receptors by the binding of the BMP or TGF-β cytokines. Activation of type II receptors recruits and phosphorylates ALK type I receptors, which in turn activate the R-SMADs by phosphorylating the SSXS residues in the C terminal domain (308-310). In basal conditions the SMADs shuttle constantly between the cytoplasm and the nucleus (183) via interaction with the nucleoporins (159). After receptor-mediated R-SMAD activation, an import protein is required for their nuclear translocation. The activated R-SMADs recruit SMAD4 and the complex is transported to the nucleus via different import proteins like IPO7 and IPO8 (179). These importins recognize a conserved stretch of charged lysine residues that acts as a nuclear localization signal (NLS). An NLS is present in the MAD homology 1 (MH1) domain of the R-SMADs, but differs in sequence in the Co-SMAD and I-SMADs

(182,183,311). Gene activation by SMADs also induces a negative feedback loop through induced expression of I-SMAD target genes SMAD6 and SMAD7, which negatively affect SMAD signaling (221,223,312,313).

To study the SMAD GSTF protein family in a comprehensive manner we generated a quantitative proteomics pipeline for human cells. Using this method we identified IPO5 as a novel SMAD1 interactor. Overexpression of IPO5 increases nuclear localization of BMP but not TGF-β R-SMADs, which suggested a functional relationship between BMP signaling and IPO5. Mutation experiments extending the lysine stretch of the NLS of SMAD3 shows that the length of the lysine stretch is involved in differentiating IPO5 and IPO7 as importins for the BMP and TGF-β subgroup of SMAD transcription factors.

Results

GFP-SMAD cell line generation

In order to study the interactome of the human SMAD proteins, stable doxycycline-inducible HeLa cell lines containing GFP-SMAD fusion proteins or GFP without fusion partner as control were generated using the Flp-In T-Rex system (Fig. S1). To characterize expression of the GFP-fusion proteins, a doxycycline induction time series was performed. Immunoblot shows maximal protein expression after 16 hours with no further increase at 20 or 24 hours induction time points (Fig. 1B). SMAD protein activity is regulated via post-translational modifications affecting protein localization and functional SMADs are expected to shuttle between cytoplasm and nucleus. Indeed, confocal microscopy shows that all GFP-SMAD can be detected both in the cytoplasm and in the nucleus (Fig. 2) suggesting that localization of the fusion proteins is not impaired by the N-terminal GFP tag. Subsequently, we tested SMAD activation evaluating the phosphorylation status of SMAD2 upon TGF-β stimulation using phosphor-specific antibodies. Phosphorylated GFP-SMAD2 and endogenous SMAD2 was detected in treated and untreated conditions suggesting TGF-β pathway is (partially) active under normal HeLa cell culture conditions (Fig. 3A). To confirm this observation, the SB431542 inhibitor of TGF-β signaling (314) was added in various concentrations to the HeLa GFP-SMAD2 cell line one hour prior to harvesting (Fig. 3B). A decrease in phosphorylation of both GFP-SMAD2 and endogenous SMAD2 was observed

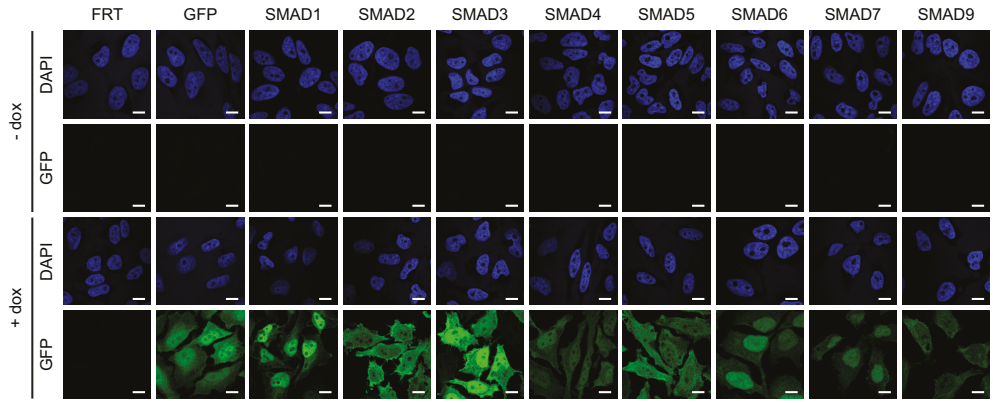
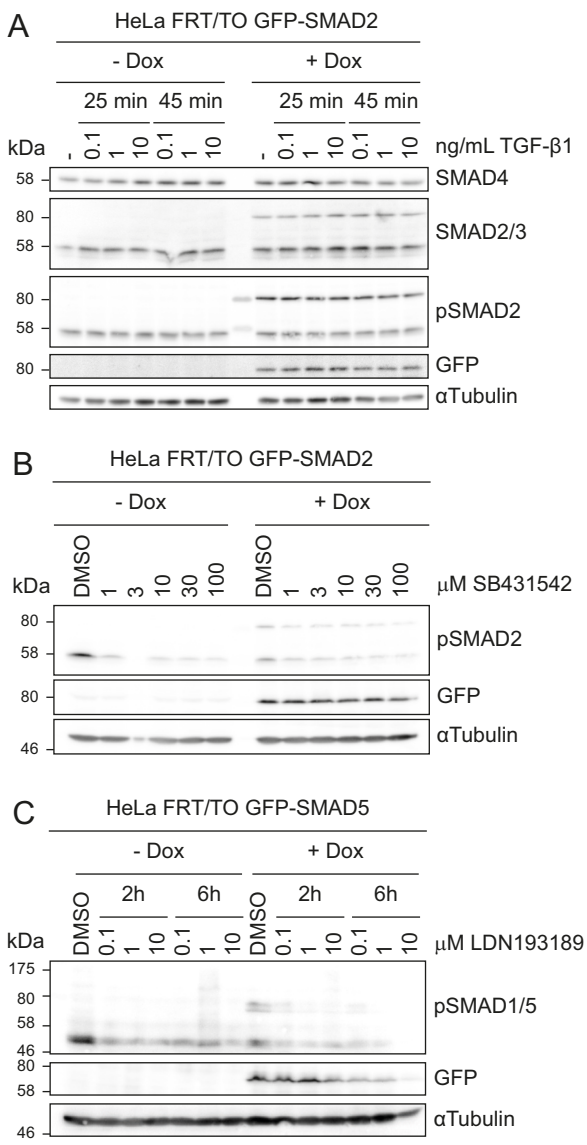


Figure 2. Analysis of GFP-SMAD expression. Doxycycline induction test for confocal microscopy shows GFP-fusion protein localization. HeLa cell lines were seeded on coverslips and induced with 1 µg/ml doxycycline for 16 hours prior to fixing and analyzed using confocal microscopy.



compared to untreated control samples confirming active TGF-β signaling under normal cell culture conditions. To determine whether BMP signaling is also activated under normal HeLa cell culture conditions, we treated GFP-SMAD5 cells with the BMP signaling inhibitor LDN193189. Again, phosphorylation levels of endogenous and GFP-SMAD5 were reduced compared to control conditions (Fig. 3C), suggesting active BMP signaling. Overall, partially active TGF-β and BMP signaling pathways are consistent with detection of GFP-SMAD proteins both in the cytoplasm and in the nucleus (Fig. 2). These data show that we generated stable inducible HeLa cell lines expressing functional GFP-SMAD proteins.

SMAD interactome identification using mass spectrometry

Localization of GFP-SMAD proteins to both the cytoplasm and nucleus suggests different pools of proteins with different biological activity and/or interaction partners. To systematically study these different SMAD pools we prepared cytoplasmic and nuclear extracts from the GFP-SMAD cell lines after 16 – 20 hours of doxycycline induction (Fig. 4A).

Figure 3. GFP-SMAD fusion constructs are responsive to phosphorylation. (A) TGF-β stimulation titration on HeLa GFP-SMAD2 expressing cells on two time points. Cells were seeded and induced with doxycycline for 16 hours where indicated. Cells were induced with TGF-β 0.1, 1 or 10 ng/ml for 25 or 45 minutes prior to harvesting. Samples were analyzed on immunoblot using SMAD4, SMAD2/3, phospho-SMAD2, GFP and αTubulin antibodies. (B) SB431542 inhibitor titration on HeLa GFP-SMAD2 expressing cells. Cells were seeded and induced with doxycycline for 16 hours where indicated. Inhibitor was added with 1, 3, 10, 30 or 100 μM 60 minutes prior to harvesting. Samples were analyzed on immunoblot using phospho-SMAD2, GFP and αTubulin antibodies. (C) LDN193189 inhibitor titration on HeLa GFP-SMAD5 expressing cells on two time points. Cells were seeded and induced with doxycycline for 16 hours where indicated. Inhibitor was added with 0.1, 1 or 10 μM 60 minutes prior to harvesting. Samples were analyzed on immunoblot using phospho-SMAD1/5, GFP and αTubulin antibodies.

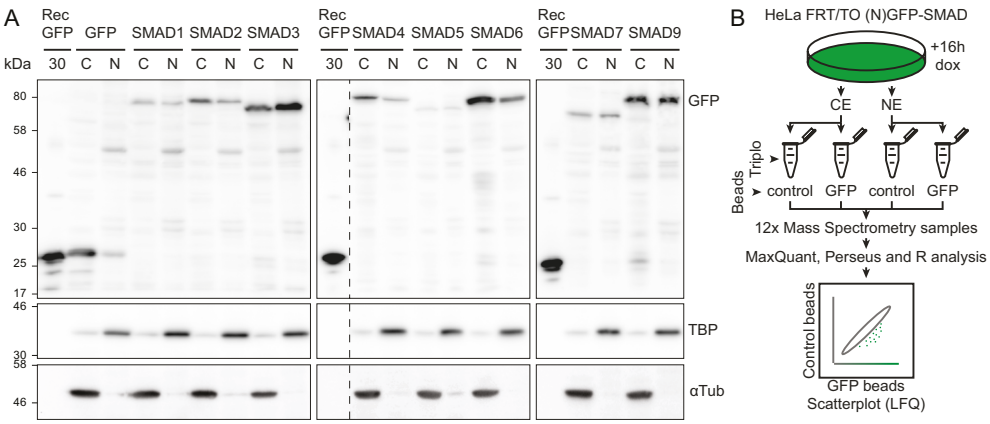


Figure 4. Analysis of nuclear and cytoplasmic extracts and mass spectrometry approach. (A) Immunoblot containing 50 µg input protein originating from cytoplasmic and nuclear extracts were analyzes using GFP-antibodies. Recombinant GFP (30 ng) was run in parallel for reference. Samples were analyzed using GFP, TBP and αTubulin antibodies. Dashed line indicates removal of lanes unrelated to this experiment on blot. Rec GFP = recombinant GFP; C = cytoplasmic fraction; N = nuclear fraction. (B) Schematic representation of experimental workflow. Cytoplasmic and nuclear extracts were prepared from HeLa GFP-fusion protein expressing cells. GFP pulldown or mock pulldown experiments were performed on the extracts in triplo. Samples were measured by mass spectrometry and raw data was analyzed using MaxQuant, Perseus and visualized in R.

Immunoblot analysis of cytoplasmic and nuclear fractions confirmed the GFP-fluorescence results of Fig. 2. Blots were developed with antibodies against the nuclear TATA-binding protein (TBP) or cytoplasmic tubulin as controls. We found that GFP-SMAD1, SMAD5 and SMAD7 are expressed to lower levels as the other GFP-SMADs. To determine the interactome of SMAD proteins cytoplasmic and nuclear extracts were subjected to GFP-affinity purification using nanobeads (Fig. 4B). All purifications were performed in triplicate and measured by MS/MS analysis using a Velos mass spectrometer. In total, 108 affinity purifications were performed and all data files were simultaneously analyzed using MaxQuant. Correlation plots show very good reproducibility between all samples (Fig. S2). A clear correlation is observed between samples within the same cell line for the extracts (cytoplasmic versus nucleus), while no clear correlation is observed between extracts from separate GFP-SMAD cell lines.

The data was visualized using scatterplots, which were generated with the median data obtained for each triplicate set. Hereby, the affinity-purified samples were plotted as x values against its corresponding control situation as the y values. When using this data representation a protein cloud is expected on the diagonal representing proteins that non-specifically bind to the beads (Fig. 5 and 6). Proteins that bind to the GFP-tagged bait protein will be more abundant in the affinity purification with the GFP-binding beads and therefore, they will be shifted from the diagonal towards the x-axis. In certain situations, for example when a protein is only expressed in a low amount, proteins will only be identified when enriched using the GFP-binder beads but not the control beads. To include these cases a constant value was imputed for missing data points to plot values that only contain a value for one bead type. When the results are plotted the data forms a horizontal line for proteins identified only on GFP binding beads and a vertical line for proteins only identified on control beads. Proteins present on the horizontal line are therefore also regarded as enriched, but the fold enrichment cannot be calculated.

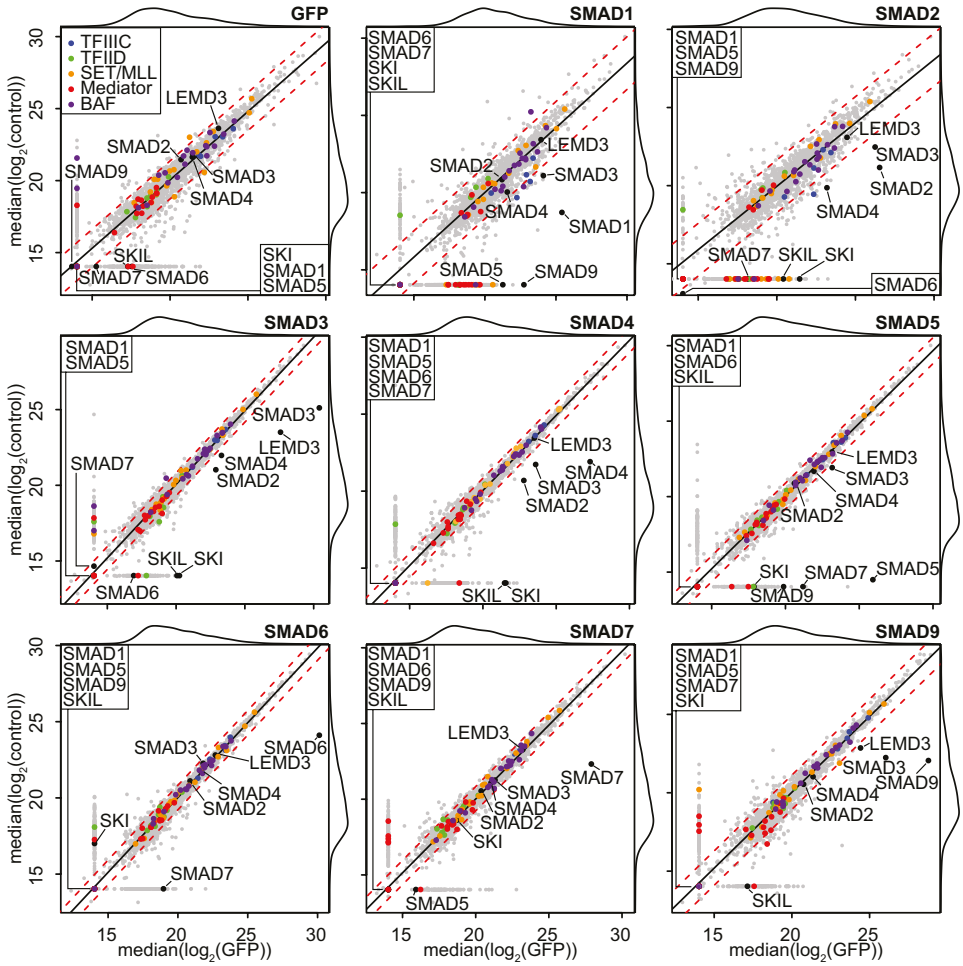


Figure 5. Mass spec analysis for nuclear fractions. Scatterplot representations for nuclear fraction affinity purifications. GFP affinity purifications (x-axis) were plotted against control purifications (y-axis). Subunits from TFIIIC (blue), TFIID (green), SET/MLL (orange), Mediator (red) and BAF (purple) complexes are highlighted using their respective colors and SMADs, SKI, SKIL and LEMD3 are indicated. The continuous black line and dotted red lines represent a linear model based on the data with the respective 95% confidence interval. A distribution of all proteins without an imputed value is indicated in the top and right sides of the scatterplot.

As validation of the mass spec data from the nuclear fraction, we examined SMAD enrichment in different SMAD baits (Fig. 5). As expected, the SMAD2/3/4 proteins interact with each other in the nucleus. Furthermore, the TGF-β inhibitory proteins SKI and SKIL are present in SMAD2/3/4 affinity purifications. These two proteins are known direct interactors of the SMAD2/3/4 complex (315,316). In addition and as expected, we identified the R-SMAD inhibitor LEMD3 (MAN1) (227,228) as a SMAD3 interactor. Interestingly, LEMD3 is not enriched in SMAD4 affinity purifications. This indicates that LEMD3 and SMAD4 are mutually exclusive binders for SMAD2/3 and fit the current view that LEMD3 is a repressor of R-SMAD activation (227,228). LEMD3 is not enriched in SMAD1/2/5 while this interaction has been described before (228). This could be a HeLa specific or a transcriptional status effect. A new interaction of LEMD3 with SMAD9 is also identified and complements

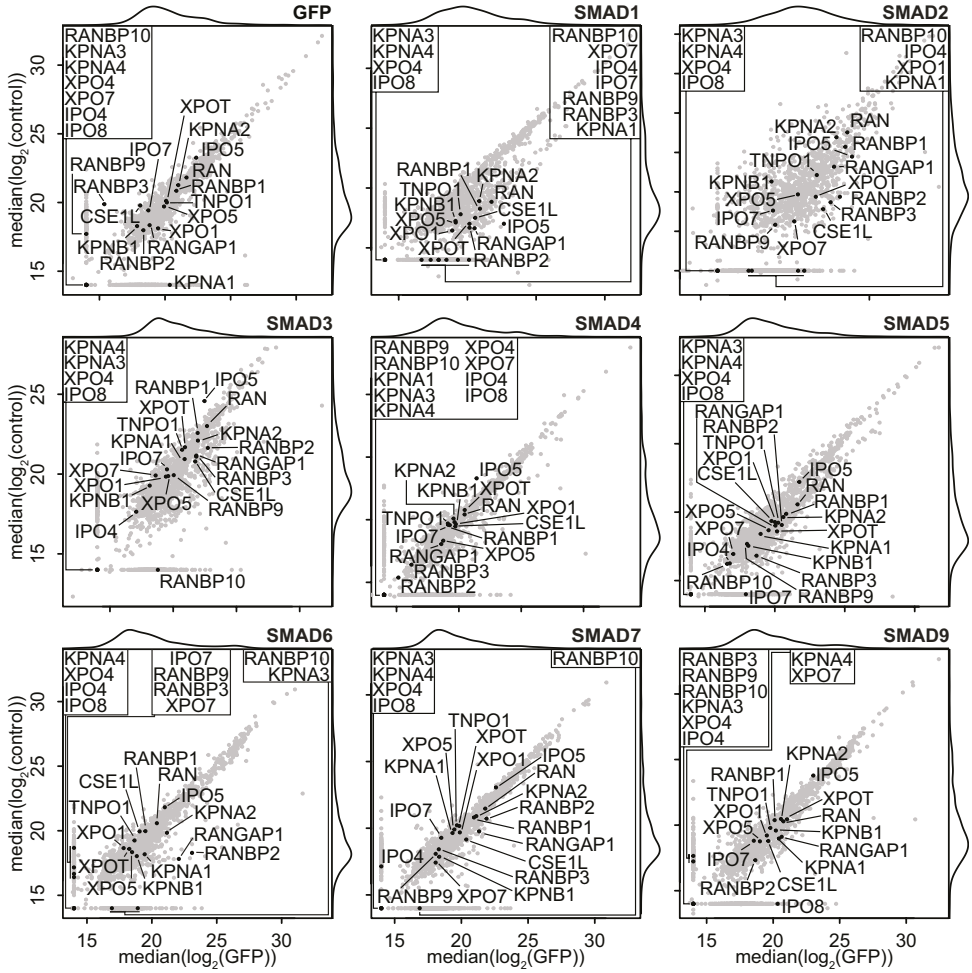


Figure 6. Mass spec analysis for cytoplasmic fractions. Scatterplot representations for cytoplasmic fraction affinity purifications. GFP affinity purifications (x-axis) were plotted against control purifications (y-axis). Identified transport proteins are highlighted. A distribution of all proteins without an imputed value is indicated in the top and right sides of the scatterplot.

the current view on LEMD3 as a repressor of R-SMAD activation. Next, several complexes that have different functions in the nucleus were visualized for enrichment in the nuclear fractions (Fig. 5 and Table S1). Enrichment for several of the TFIID, SET1/MLL and Mediator subunits were observed in SMAD1 and SMAD2 while no enrichment was observed in other SMADs or the GFP control cell line. The TFIID, SET1/MLL and Mediator complexes are known for their activating role in pol II-mediated transcription. As a control we examined the pol III transcription factor TFIIIC and found that it was not enriched in any SMAD cell line, nor was the BAF chromatin regulatory complex. This suggests that SMAD1 and SMAD2 are employing specific transcription co-activator complexes to stimulate pol II-mediated transcription.

To examine the quantitative proteomics results for the cytoplasmic fraction the data were also represented as scatterplots. Apart from the expected background distribution a

second protein distribution, absent in the nuclear fraction was observed (Fig. S3A). Gene ontology (GO) term analysis shows enrichment for unfolded protein binding and protein biosynthesis (Fig. S3B). Possibly the interactors were identified due to incomplete synthesis of the GFP-SMAD proteins, although no discrete partial GFP-SMAD proteins are observed by immunoblotting (Fig. 4A). Alternatively, the GFP-SMAD fusion protein might be incorrectly folded. It is important to note that this secondary distribution is not observed in the affinity purifications of the GFP control samples from the cytoplasmic fraction.

SMAD proteins require transport proteins to shuttle between the cytoplasm and nucleus. To investigate which transport proteins associate with the different SMADs all transport proteins were identified in the cytoplasmic fraction plots. For the different SMADs various transport proteins were found enriched, while in the GFP control cell line most transport proteins were identified in the background diagonal. SMAD4, which needs an R-SMAD as partner to translocate to the nucleus, does not contain any enriched transport proteins. SMAD1 shows enrichment for several transport proteins like the previously

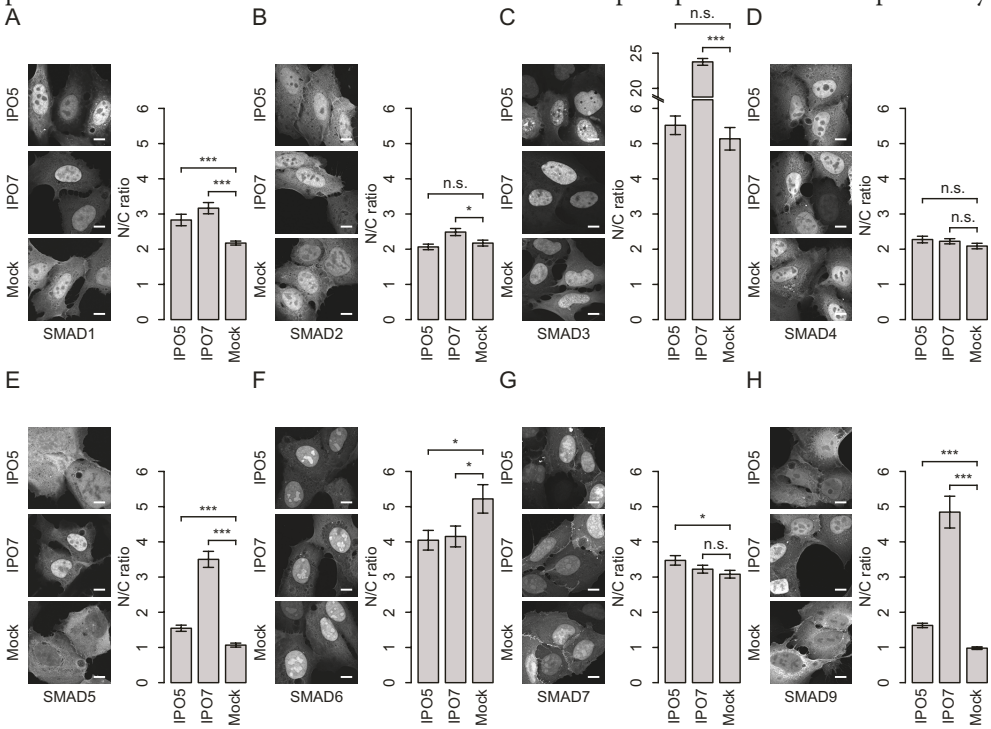


Figure 7. Overexpression of IPO5 stimulates nuclear localization of BMP- but not of TGFβ-activated RSMADs. (A-H) U-2 OS cells were seeded on coverslips and transfected using a combination of GFP-SMADs with FLAG-IPO5, FLAG-IPO7 or mock. Cells were fixed 48 hours after transfection and analyzed using confocal microscopy. Nuclear and cytoplasmic intensities were quantified using ImageJ. Shown is combined data of n = 2 biological replicates with varying technical replication n specified per set. Bars represent average nuclear/cytoplasmic (N/C) ratio, error bars represent standard error of the mean. Significant differences are indicated for various significance levels (*** = $p < 0.001$; ** = $0.001 < p < 0.01$; * = $0.01 < p < 0.05$; n.s. = not significant). Scale bar = 10µm. (A) SMAD1 N/C ratios. Technical replication n = 55, 61, 74. (B) SMAD2 N/C ratios. Technical replication n = 72, 67, 67. (C) SMAD3 N/C ratios. Technical replication n = 61, 71, 62. (D) SMAD4 N/C ratios. Technical replication n = 70, 59, 70. (E) SMAD5 N/C ratios. Technical replication n = 46, 45, 41. (F) SMAD6 N/C ratios. Technical replication n = 47, 55, 46. (G) SMAD7 N/C ratios. Technical replication n = 62, 58, 66. (H) SMAD9 N/C ratios. Technical replication n = 64, 55, 64.

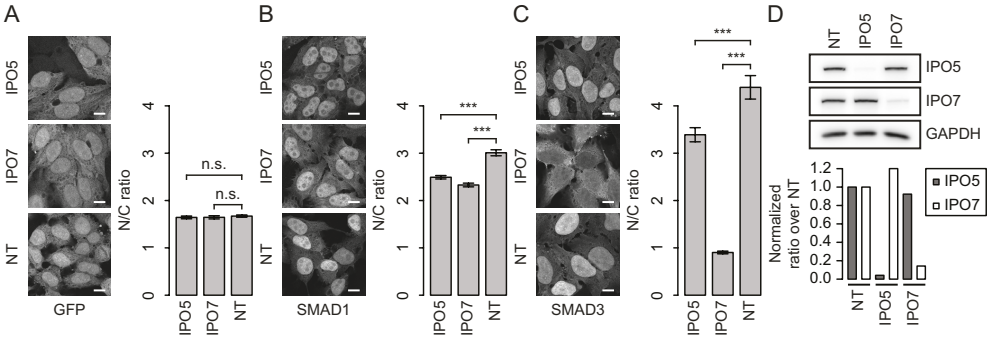


Figure 8. Knockdown of IPO5 reduces nuclear localization of SMAD1. (A-C) HeLa cell lines were transfected with siRNA against non-targeting, IPO5 or IPO7. Cells were induced with 1 µg/ml doxycycline for 16 hours prior to fixing and analyzed 72 hours after siRNA transfection using confocal microscopy. Nuclear and cytoplasmic GFP-intensities were quantified using ImageJ. A representative experiment is shown. Bars represent average nuclear/cytoplasmic (N/C) ratio, error bars represent standard error of the mean. Significant differences are indicated for various significance levels (*** = $p < 0.001$; ** = $0.001 < p < 0.01$; * = $0.01 < p < 0.05$; n.s. = not significant). Scale bar = 10µm. (A) GFP N/C ratios. Technical replication n = 63, 62, 69. (B) SMAD1 N/C ratios. Technical replication n = 70, 80, 64. (C) SMAD3 N/C ratios. Technical replication n = 57, 59, 58. (D) Knockdown efficiency of IPO5 and IPO7 in non-induced HeLa GFP cells. Cells were reverse transfected with siNT, siIPO5 or siIPO7 and collected after 72 hours. Samples were analyzed on immunoblot and bands were quantified in ImageJ. Band intensities were corrected for GAPDH expression and normalized to siNT.

reported IPO7 (179), but also displays a strong enrichment of IPO5 (Fig. 6).

Confirmation of Importin5 as a novel interactor of SMAD1

Control of subcellular localization is important for the biological function of SMAD proteins. IPO7 overexpression was shown previously to force nuclear translocation of SMAD1/3/4 proteins (179). To investigate the function of IPO5 as a SMAD interactor we coexpressed all GFP-SMAD fusion proteins with IPO5 or IPO7 in human U-2 OS osteosarcoma cells to measure the ratio of nuclear over cytoplasmic (N/C) GFP intensities. As expected, overexpression of IPO7 stimulates nuclear localization of the receptor activated SMADs 1, 2, 3, 5 and 9 but not the inhibitory SMADs 6 and 7 or the common SMAD4 (Fig. 7). Strikingly, overexpression of IPO5 forced nuclear localization of the BMP subclass R-SMADs 1, 5 and 9 but not the TGF-β subclass R-SMADs 2 and 3 suggesting IPO5 subclass specificity. IPO5- and IPO7-dependent nuclear localization was confirmed for SMAD1 and SMAD3 in the human breast adenocarcinoma cell line MCF7 (Fig. S4). Knockdown of IPO5 or IPO7 using siRNA in HeLa GFP-SMAD stable cell lines gave inconclusive results (data not shown) except for SMAD1 and 3, which accumulated reproducibly in the cytoplasm (Fig. 8), Compared to IPO7, knockdown of IPO5 has a small effect on GFP-SMAD3. IPO5 has been shown to interact with the RanBP1 and RanBP2 proteins (155,166). Knockdown of RanBP1 or of RanBP2 using siRNAs was performed in HeLa cells expressing GFP-SMAD1 or GFP-SMAD3 to investigate their subcellular distribution. We found N/C ratios of these SMAD proteins are reduced significantly upon RanBP2 but not RanBP1 knockdown (Fig. 9A-D). Since IPO5 and IPO7 facilitate nuclear localization of BMP R-SMAD proteins, expression of IPO5 and IPO7 were examined in context of BMP activation. Both IPO5 and IPO7 levels remained constant upon BMP stimulation or inhibition in U-2 OS cells. As expected pSMAD1/5 levels are responsive to activation and repression of BMP signaling (Fig. 9E). Taken together, the overexpression and knockdown results support involvement of IPO5 in the subcellular distribution of a subset of SMAD proteins.

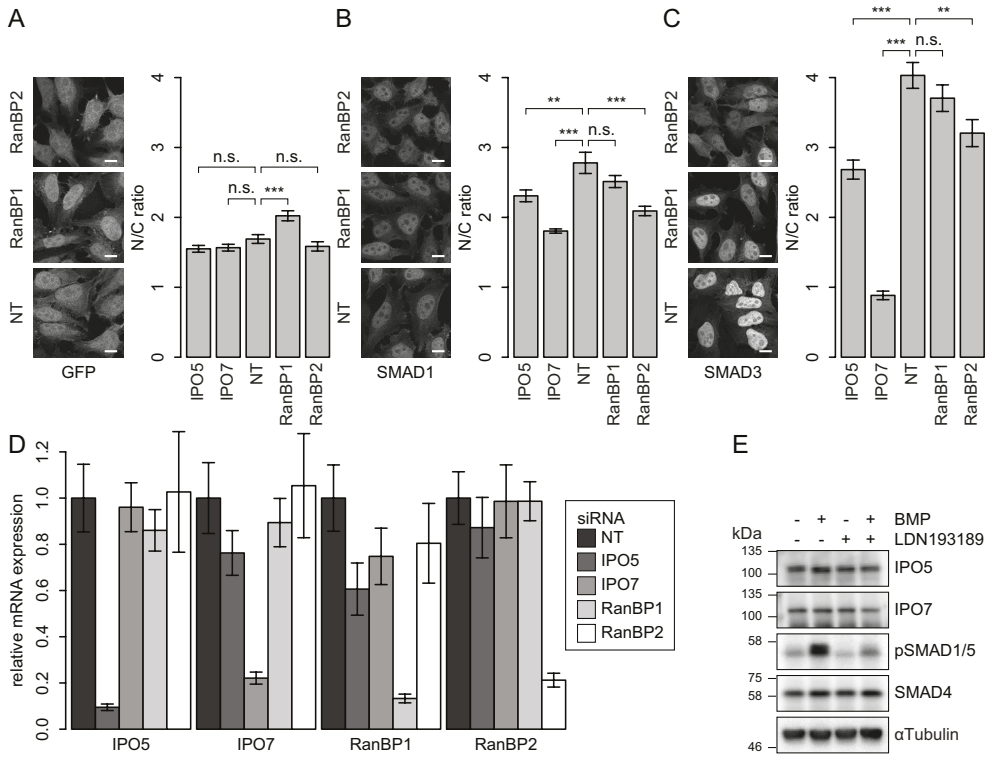


Figure 9. Knockdown of RanBP2 but not RanBP1 reduces nuclear localization of SMAD1 and 3. (A-C) HeLa cell lines were transfected with siRNA against non-targeting, IPO5, IPO7, RanBP1 or RanBP2. Cells were induced with 1 µg/ml doxycycline for 16 hours prior to fixing and analyzed 72 hours after siRNA transfection using confocal microscopy. Nuclear and cytoplasmic GFP-intensities were quantified using ImageJ. Bars represent average nuclear/cytoplasmic (N/C) ratio, error bars represent standard error of the mean. Significant differences are indicated for various significance levels (*** = $p < 0.001$; ** = $0.001 < p < 0.01$; * = $0.01 < p < 0.05$; n.s. = not significant). Scale bar = 10µm. (A) GFP N/C ratios. Technical replication n = 50, 48, 44, 49, 40. (B) SMAD1 N/C ratios. Technical replication n = 45, 63, 43, 55, 56. (C) SMAD3 N/C ratios. Technical replication n = 52, 39, 59, 37, 49. (D) Knockdown efficiency of non-targeting, IPO5, IPO7, RanBP1 and RanBP2 in non-induced HeLa GFP cells. Cells were reverse transfected with siNT, siIPO5, siIPO7, siRanBP1 or siRanBP2 and collected after 72 hours. Samples were analyzed using qRT-PCR. Bars represent average expression corrected for ACTIN and normalized to the siNT control. Error bars represent standard deviation. Technical replication n = 3 for all samples. (E) IPO5 and IPO7 levels do not change upon BMP stimulation. U-2 OS cells were treated with 100 nM LDN193189 inhibitor for 24 h and subsequently stimulated with 10 ng/ml BMP for 6 h as indicated. Immunoblot was probed using IPO5, IPO7, pSMAD1/5, SMAD4 and Tubulin antibodies.

An extended NLS lysine stretch is required for IPO5 nuclear import.

The nuclear localization sequence (NLS) of SMAD proteins consists of an N-terminal basic motif and it is localized in the MH1 domain. R-SMADs contain the canonical KKLKK sequence. However, BMP R-SMADs carry a basic stretch extended by two additional lysines (KK) compared to TGF-β R-SMADs (-T) (Fig. 10A). To investigate the molecular mechanism by which IPO5 distinguishes between the BMP and TGF-β R-SMADs, the NLS of SMAD3 (SMAD3 wt) was replaced by the NLS containing a long lysine stretch (SMAD3 T45K 45insK). Overexpression of SMAD3 T45K 45insK with IPO5 stimulates nuclear localization, which is not observed in the SMAD3 wt control (Fig. 10B). These results suggest that IPO5-mediated nuclear localization is dependent on an extended lysine stretch present in the NLS of the BMP induced R-SMADs.

Discussion

SMAD signaling is involved in many different cellular processes.

In this study interactome datasets for all SMAD proteins were generated using affinity purification for HeLa cells followed by mass spec analysis. The nuclear transport protein IPO5 was identified as a novel interactor for BMP-regulated SMAD1 and using confocal microscopy a functional link was shown between IPO5 and the BMP regulated R-SMADs but not TGF-β regulated R-SMADs. By extending the NLS lysine stretch in SMAD3 we found SMAD3 nuclear localization sensitive to IPO5, suggesting that the length of the lysine stretch in the NLS is responsible for differentiation between the BMP and TGF-β regulated SMADs.

SMAD proteins shuttle continuously between cytoplasm and nucleus, which is independent of a receptor activation signal (183). The export of SMAD4 is known to be CRM1-mediated, while export of R-SMADs is CRM1-independent. Phosphorylation of R-SMAD followed by SMAD4 recruitment causes nuclear accumulation of SMADs, which

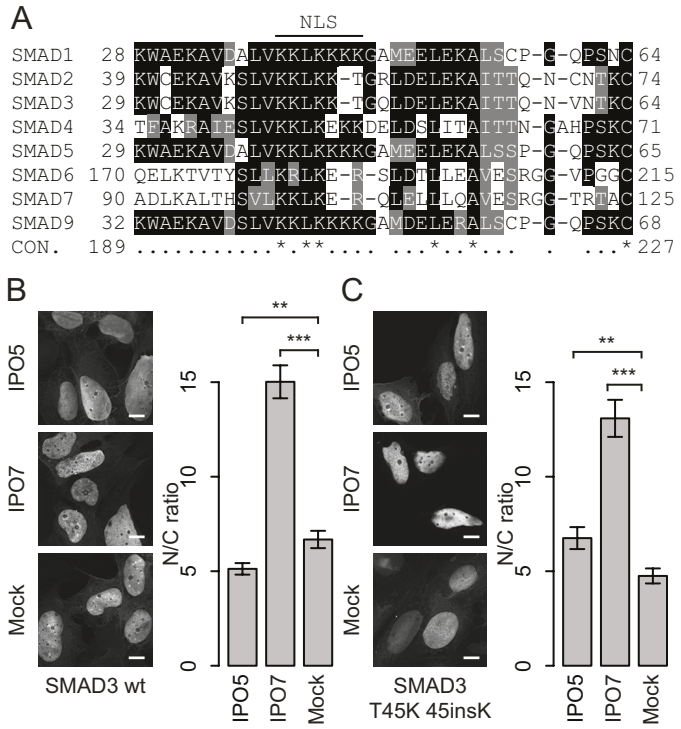


Figure 10. Extension of the SMAD3 NLS allows for IPO5 mediated nuclear localization. (A) Multiple sequence alignment for the SMADs. Clustering was performed using Clustal Omega. NLS sequence is indicated above. Con. = consensus. (B-C) U-2 OS cells were seeded on coverslips and transfected using a combination of GFP-SMAD3 wt or T45K 45insK with FLAG-IPO5, FLAG-IPO7 or mock. Cells were fixed 48 hours after transfection and analyzed using confocal microscopy. Nuclear and cytoplasmic intensities were quantified using ImageJ. Shown is combined data of n = 2 biological replicates with varying technical replication n specified per set. Bars represent average nuclear/cytoplasmic (N/C) ratio, error bars represent standard error of the mean. Significant differences are indicated for various significance levels (*** = $p < 0.001$; ** = $0.001 < p < 0.01$; * = $0.01 < p < 0.05$; n.s. = not significant). Scale bar = 10μm. (B) SMAD3 wt N/C ratios. Technical replication n = 66, 97, 82. (C) SMAD3 T45K 45insK N/C ratios. Technical replication n = 70, 92, 82.

result in reduced nuclear export. Transcription factors like TAZ (317) have been shown to be responsible for SMAD nuclear accumulation, although the exact mechanism is unknown. In the nucleus, phosphorylation of R-SMAD complexes is gradually reduced causing dissociation of SMAD4 followed by export from the nucleus. This constant SMAD nuclear/cytoplasmic cycling is an important control mechanism for constant monitoring of the receptor activation status. This results in a quick adaptation of nuclear SMAD accumulation, when the signal is present and to reduce to steady state levels when the receptors become inactive. Initially, transport proteins importin α1 and β1 (181) have been connected with SMAD signaling. Later studies implied IPO7 and IPO8 in signal dependent and independent transport of SMAD1/3/4 involving NLS found in the MH1 domain. Our results corroborate an important role of IPO7 in subcellular distribution of SMAD1, 3, 5 and 9 (Fig. 7). IPO7 seems to be the major contributor of SMAD nuclear import, while the import mediated by IPO5 is restricted. Nuclear import of the R-SMADs by IPO7 does not select for a specific type of R-SMAD (Fig 7). Systemically testing of IPO5 against all SMADs in this study shows selectivity of IPO5 for BMP activated R-SMADs, which is determined by the length of the lysine stretch of the NLS. Selectivity of the IPO5 protein for the cargo allows for regulation of signaling, thus adding a new player in controlling the subcellular distribution of SMAD proteins.

The SMAD activation pathway is a signaling cascade that gains complexity in regulation over different levels in the cell (312,318-323). Over 30 different molecules including TGF-β and BMP can be found in the human body that belong to the TGF-β superfamily (324). These molecules can signal to a combination of seven type I and five type II receptors. Since not all of these receptors are present in each cell type this makes TGF-β and BMP signaling very context-dependent. It is important to note that we did not identify any TGF-β or BMP receptor in GFP-SMAD purifications from the cytoplasmic fractions, which could be due to the low efficiency of membrane protein extraction in our procedures.

The N-terminal location of the GFP moiety in the SMAD fusion constructs was chosen to maintain the C-terminal SSXS residues required for R-SMAD phosphorylation and protein activation (308-310). This resulted in fusion constructs that could be phosphorylated upon stimulation of their receptors using their respective ligands (Fig. 3). However, the possibility still remains that interaction partners have been missed due to N-terminal tagging of SMAD proteins. In our experiments we confirmed binding of the transcriptional repressor LEMD3 to SMAD3 in the nucleus and provide evidence that LEMD3 and SMAD4 binding are mutually exclusive (Fig. 5). Interestingly, no LEMD3 binding to other R-SMADs is observed while SKI and SKIL are identified in SMAD2/3 and 4, which indicates a high specificity in repressor interactions

Activated SMAD proteins stimulate gene transcription, but the exact nature of their nuclear coactivators has not been determined in much detail. We provide evidence that Mediator, TFIID and SET/MLL but not BAF complex could participate in transcription activation by the SMAD1 and SMAD2 proteins. The observation of specific co-activators for specific SMAD family members warrants a further study into the downstream transcription activation pathways for TGF-β and BMP signaling.

Experimental procedures

Plasmids

PCR was performed on plasmids containing SMAD cDNAs using primers containing attB1 or attB2 sites to obtain Gateway compatible constructs (Table 1) (325). FLAG-SMAD2 and FLAG-SMAD4 cDNAs (326) were generous gifts from dr. Peter ten Dijke (Leiden University, Leiden, The Netherlands). IMAGE clones 3546771, 6094328, 3906006, 3632931, 30915321 and 8143763 were purchased for SMAD 1, 3, 5, 6, 7 and 9 respectively from Dharmacon (Lafayette, CO). PCR products were recombined with pDON201 in a Gateway BP reaction to obtain pENTR vectors according to manufacturer's instructions (Thermo Fischer Scientific, Waltham, MA). All pENTR vectors were confirmed using Sanger sequencing. Sequencing showed SMAD1 to be a V140L variant compared to the refseq sequence. All other SMADs were identical to the refseq. To obtain HeLa FRT compatible expression vectors pENTR vectors containing SMAD variants were recombined with Gateway-compatible pCDNA5/FRT/TO/N-GFP in an LR reaction according to manufacturer's instructions (Thermo Fischer Scientific, Waltham, MA). SMAD3 mutagenesis PCR was performed using mutagenesis primers in Table 1. pCDNA3.1-FLAG-IPO5 (327) was a generous gift from dr. Yi-Shuian Huang (Institute of Biomedical Sciences, Taipei, Taiwan). pME18S-FLAG-hIPO7 (pTIB435) (328) was a generous gift from dr. Tohru Itoh (The University of Tokyo, Tokyo, Japan).

Cell culture

HeLa cells containing the Flp Recombination Target (FRT) site and Tet repressor were grown in Dulbecco's Modified Eagle's Medium (DMEM) containing 4,5 g/L glucose

Table 1. Primers used for cloning

Primer	Sequence
SMAD1 Fw	GGGG ACA AGT TTG TAC AAA AAA GCA GGC TTC AAT GTG ACA AGT TTA TTT TCC TTT AC
SMAD2 Fw	GGGG ACA AGT TTG TAC AAA AAA GCA GGC TTC TCG TCC ATC TTG CCA TTC AC
SMAD3 Fw	GGGG ACA AGT TTG TAC AAA AAA GCA GGC TTC TCG TCC ATC CTG CCT TTC AC
SMAD4 Fw	GGGG ACA AGT TTG TAC AAA AAA GCA GGC TTC GAC AAT ATG TCT ATT ACG AAT AC
SMAD5 Fw	GGGG ACA AGT TTG TAC AAA AAA GCA GGC TTC ACG TCA ATG GCC AGC TTG
SMAD6 Fw	GGGG ACA AGT TTG TAC AAA AAA GCA GGC TTC TTC AGG TCC AAA CGC TCG
SMAD7 Fw	GGGG ACA AGT TTG TAC AAA AAA GCA GGC TTC TTC AGG ACC AAA CGA TCT GC
SMAD9 Fw	AA AAA GCA GGC TTC CAC TCC ACC ACC CCC ATC AGC TCC CTC TTC TCC
ADAPT Fw	GGGG ACA AGT TTG TAC AAA AAA GCA GGC T
SMAD1 Rv	GGGG ACC ACT TTG TAC AAG AAA GCT GGG TC CTA AGA TAC AGA TGA AAT AGG ATT ATG AG
SMAD2 Rv	GGGG ACC ACT TTG TAC AAG AAA GCT GGG TC CTA TGA CAT GCT TGA GCA ACG CAC
SMAD3 Rv	GGGG ACC ACT TTG TAC AAG AAA GCT GGG TC CTA AGA CAC ACT GGA ACA GCG GAT G
SMAD4 Rv	GGGG ACC ACT TTG TAC AAG AAA GCT GGG TC CTA GTC TAA AGG TTG TGG GTC TG
SMAD5 Rv	GGGG ACC ACT TTG TAC AAG AAA GCT GGG TC CTA TGA AAC AGA AGA TAT GGG GTT CAG AG
SMAD6 Rv	GGGG ACC ACT TTG TAC AAG AAA GCT GGG TC CTA TCT GGG GTT GTT GAG GAG G
SMAD7 Rv	GGGG ACC ACT TTG TAC AAG AAA GCT GGG TC CTA CCG GCT GTT GAA GAT GAC C
SMAD9 Rv	AG AAA GCT GGG TC CTA AGA CAC TGA AGA AAT GGG GTT ATG TGG AGA GCC
ADAPT Rv	GGGG ACC ACT TTG TAC AAG AAA GCT GGG T
SMAD3 T45K 45insK FW	GGT CAA GAA ACT CAA GAA GAA GAA AGG GCA GCT GGA CGA GCT GG
SMAD3 T45K 45insK RV	CCA GCT CGT CCA GCT GCC CTT TCT TCT TCT TGA GTT TCT TGA CC

supplemented with 10% v/v fetal bovine serum, 10 mM L-glutamine, 100 U/mL penicillin and 100 U/mL streptomycin. HeLa FRT parental cell line growth medium was supplemented with 5 µg/mL blasticidin S (InvivoGen, San Diego, CA) and 200 µg/mL zeocin (Invitrogen, Carlsbad, CA) to select for the FRT site and Tet repressor. To generate cell lines containing stable integration of GFP-tagged proteins, cells were seeded in culture medium without blasticidin S and zeocin antibiotics 8 hours prior to transfection. Stable doxycycline-inducible HeLa cell lines were created by co-transfecting the pOG44 Flp recombinase expression vector along with pCDNA5/FRT/TO expression vectors into HeLa FRT cells using polyethyleneimine (PEI). 48 hours after transfection growth medium was supplemented with 250 µg/mL hygromycin B (Roche Diagnostics, Mannheim, Germany) and 5 µg/mL blasticidin S to select for expression vector integration and Tet repressor.

Immunoblot Analysis

Cells were seeded in 6-wells plates at a density of 30,000 cells/well. Cells were induced 16 to 20 hours prior to harvesting by addition of 1 µg/mL doxycycline to the culture media unless otherwise stated. Stimulation with TGF-β1 (#8915LC, Cell Signaling) or rhBMP-2 (355-BEC, R&D systems, Minneapolis, MN) or inhibition with SB431542 (#616461, Calbiochem) or LDN193189 (SML0559, Sigma- Aldrich) was performed by addition of the listed concentrations and duration prior to harvesting. Cells were lysed in sample buffer (160 mM Tris-HCl pH 6.8, 4% SDS, 20% glycerol, 0.05% bromophenol blue) and equal amounts of protein were run on a 10% SDS-PAGE and transferred onto PVDF membrane. The membrane was developed with the appropriate antibodies and ECL and scanned using ImageQuant LAS 4000.

Antibodies

Antibodies were used from the following sources: GAPDH (clone 6C5, mAb374, Millipore), GFP (gift from Geert Kops), IPO5 (H-4, sc-17802, Santa Cruz), IPO7 (ab99273, Abcam), phospho-SMAD1/5 (D5B10, #11971, Cell Signaling), phospho-SMAD2 (138D4, #3108, Cell Signaling), SMAD2/3 (N-19, sc-6032, Santa Cruz), SMAD4 (#9515, Cell Signaling), TBP (20C7 in-house hybridoma (329)), α-Tubulin (DM1A, CP06, Calbiochem).

Nuclear extract preparation and GFP affinity purification

Cytoplasmic and nuclear extracts were prepared as described (268) with the exception that DTT was omitted from buffer C. Protein concentration was determined using Bradford assay as described before (268). GFP affinity purification was essentially performed as described (268) with buffer C lacking DTT. Sample binding was performed for 30 minutes on a rotating wheel at 4°C using 1 mg input protein Peptides were eluted from the beads by 2 h trypsin incubation in elution buffer (100 mM Tris-HCl pH 7.5, 2 M urea, 10 mM DTT). Eluate was collected and beads were eluted for a second time. Eluates were combined and trypsin digested overnight. Tryptic digests were desalted using Stage-tips (330).

Mass spectrometry

Tryptic peptides were separated using an EASY-nLC 1000 (Thermo Fischer Scientific, Waltham, MA) mounted with a C18 column operating online with an LTQ Orbitrap Velos. An acetonitrile gradient (0% - 80%) was applied over 2 h to elute peptides from the column. Peptides were ionized using ESI. The top 15 most abundant peptides were selected for fragmentation using CID for MS/MS. Proteins were identified from the raw data using

MaxQuant version 1.5.2.8 using the Uniprot human FASTA database. Perseus version 1.5.1.6 was used to filter for contaminants and reverse hits. LFQ values were log2 transformed and groups based on cell line, extract type and bead type were defined. A filter was applied to remove contaminants and proteins that could be identified using a reversed database which are very likely false-positive identifications. Then a filter was applied that identified proteins that were present at least 3 out of 3 times within at least one combination of conditions (bead type, extract and cell line) resulting in a total of 3649 selected proteins after Perseus filtering. The median of each group was taken and data was plotted using R version 3.2.2 (331) with Rstudio version 0.99.486.

siRNA transfection

Cells were transfected in a reverse transfection protocol using Hiperfect (Qiagen, Valencia, CA) according to the manufacturer's instructions. In short: 50 µL of a SMARTpool of four siRNAs (25 nM each) was dispensed in a 12-wells plate (Costar, 3513, Corning, New York, NY) containing 15 mm diameter cover glasses (Paul Marienfeld GmbH & Co. KG, Lauda Königshofen, Germany). Transfection mix (4 µL Hiperfect in total volume of 150 µL Opti-MEM [Invitrogen, Carlsbad, CA] per well) was added followed by 30 minutes incubation at room temperature. Meanwhile, cells were counted and 30,000 cells/well were added in 800 µL. Cells were fixed 72 h after transfection.

Transient transfection

U-2 OS osteosarcoma cells were counted and seeded 30,000 cells/well in a 12-wells plate containing 15 mm diameter cover glasses. Cells were left overnight to attach and expression vectors were transfected using PEI. Cells were fixed 48 hours after transfection.

Microscopy slide preparation

Cover glasses containing treated cells were washed once with phosphate-buffered saline (PBS). Cells were fixed by incubation with 4% formaldehyde in PBS for 20 minutes at room temperature. Cells were incubated with DAPI solution (2 mg/L 4',6-diamidino-2-phenylindole [Sigma-Aldrich, St. Louis, MO] in 50 mM Tris-HCl pH 7.4, 100 mM NaCl) to stain nuclei. Cover glasses were mounted onto microscope slides (Thermo Fischer Scientific, Waltham, MA) using Immu-Mount (Invitrogen, Carlsbad, CA) and left to dry overnight at 4°C.

Confocal microscopy

Images were taken using a Zeiss LSM 510 META confocal microscope with PLAN Apochromat Oil-DIC 63x/1.4 objective. Samples were scanned using a 405 nm Violet Diode Laser for DAPI staining and a 488 nm Argon Laser for GFP. Microscope gain and offset settings were kept constant within experiments that compared a fusion protein in different conditions. Nuclear and cytoplasmic intensities were measured using ImageJ version 1.50d. Plots and statistics were generated using R version 3.2.1 with Rstudio version 0.99.473.

Expression analysis qRT-PCR

Total RNA was isolated from siRNA treated cells 72 h after transfection using RNeasy kit (Qiagen, Valencia, CA) including a DNase treatment using RNase-Free DNase Set (Qiagen, Valencia, CA) according to manufacturer's instructions. 500 ng of total RNA was used for cDNA synthesis (Superscript II, Invitrogen, Carlsbad, CA) using random primers.

A 25 µL qRT-PCR reaction was performed in a CFX Connect Real-Time PCR Detection System (Bio-Rad, Hercules, CA) using iQ SYBR Green Supermix (Bio-Rad, Hercules, CA) Primers used for qRT-PCR can be found in Table 2.

Table 2. Primers used for qRT-PCR

Primer	Sequence
Actin fw	AGAAAATCTGGCACCACACC
ACTIN rv	AGAGGCGTACAGGGATAGCA
IPO5 Fw	TGCTCTAGATCGAATGGCTTG
IPO5 rv	ATGCCGGTATTTCCAGTCAG
IPO7 Fw	GTGAACAGGGATGTACCTAATGAA
IPO7 rv	ATGTAAGGCCCACTTCTTGC
RanBP1 Fw	CCAATAGTTTCTCTTCCTGAGCA
RanBP1 Rv	GAGGCAAATCGGAACAGTTT
RanBP2 Fw	TGTAGTGATACTGATGAAGACAATGG
RanBP2 Rv	TTGTGCTAGTTATTTCTTCTGTCTGAG

Acknowledgements

The authors would like to thank Proteins at Work (http://www.netherlandsproteomicscentre.nl/paw) for providing usage and support of their mass spectrometry facility. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (332) partner repository with the dataset identifier PXD004529. Dr. Peter ten Dijke (Leiden University, Leiden, The Netherlands), Dr. Yi-Shuian Huang (Institute of Biomedical Sciences, Taipei, Taiwan) and Dr. Tohru Itoh (The University of Tokyo, Tokyo, Japan) are thanked for providing plasmids. Furthermore, the authors would like to thank Dr. Simona Capponi and Ana Rita Lourenço, MSc for their critical reading of the manuscript and all the members of the Timmers group for their helpful discussions and suggestions. This work was financially supported by an NWO-Life Sciences grant (821.02.012) to H.Th.M.T.

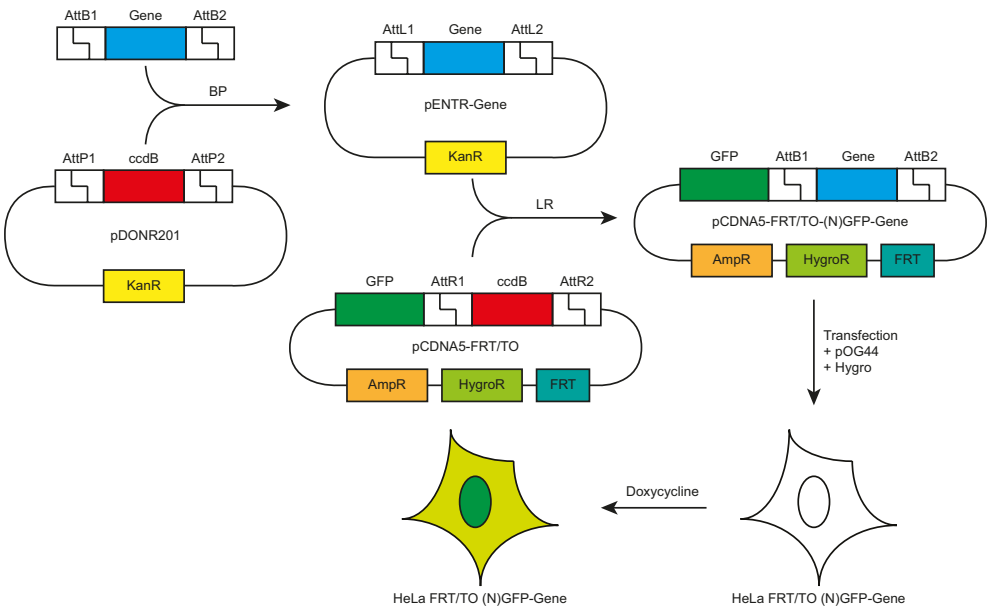


Figure S1. Gateway cloning and HeLa FRT/TO stable cell line generation strategy. Schematic representation of the used Gateway cloning strategy. The cDNA of the gene of interest was used in PCR to obtain Gateway-compatible fragments. BP and LR reactions were performed to obtain pENTR and pCDNA5 vectors respectively. Transformation of pCDNA5 vectors with pOG44 into HeLa FRT parental cell lines allows for stable integration of the pCDNA5 vector. Doxycycline induction allows for expression of the GFP-fusion construct. (AmpR = ampicillin resistance; KanR = kanamycin resistance; HygroR = hygromycin resistance; FRT = FLP Recombination Target; TO = tet operator 2; ccdB = control of cell death B).

Figure S2. Correlation plots for mass spectrometry experiments per cell line. Correlation is shown in numbers and visual representation (dots) for the different extracts (Cyto = cytoplasmic or Nucl = nuclear) and bead type (GFP or control) for the different replicates.

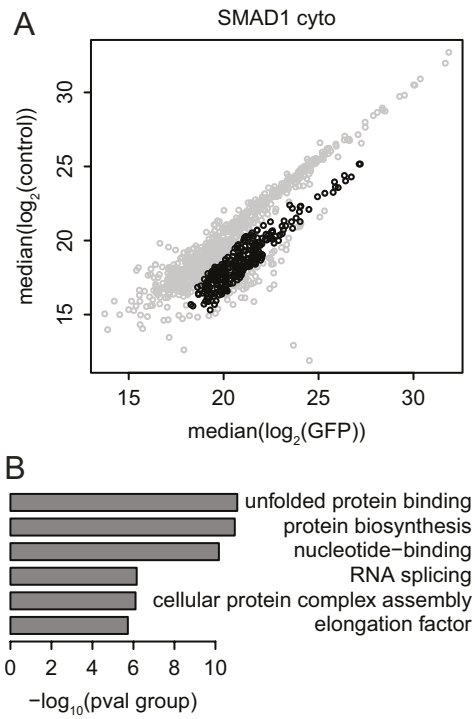


Figure S3. A secondary distribution for cytoplasmic affinity purifications. (A) Scatterplot for affinity purified cytoplasmic SMAD1. (B) DAVID GO term analysis was performed on the black highlighted dots in A.

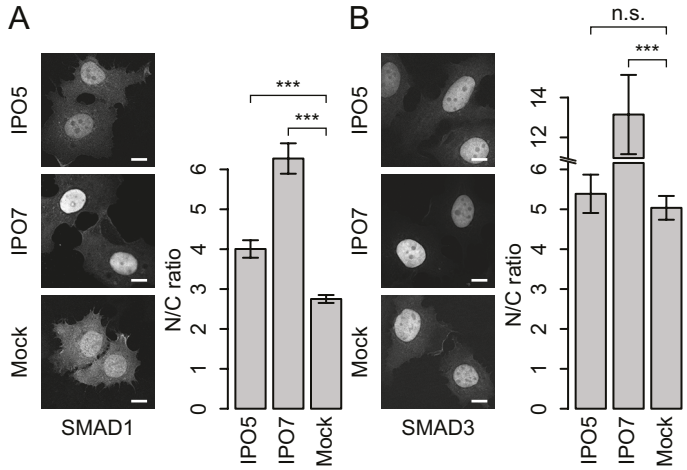


Figure S4. Overexpression of IPO5 stimulates nuclear localization of SMAD1 but not of SMAD3 in MCF7. (A-B) MCF7 cells were seeded on coverslips and transfected using a combination of GFP-SMAD1 or 3 with FLAG-IPO5, FLAG-IPO7 or mock. Cells were fixed 48 hours after transfection and analyzed using confocal microscopy. Nuclear and cytoplasmic intensities were quantified using ImageJ. Bars represent average nuclear/cytoplasmic (N/C) ratio, error bars represent standard error of the mean. Significant differences are indicated for various significance levels (*** = $p < 0.001$; ** = $0.001 < p < 0.01$; * = $0.01 < p < 0.05$; n.s. = not significant). Scale bar = 10μm. (A) SMAD1 N/C ratios. Technical replication n = 45, 46, 48. (B) SMAD3 N/C ratios. Technical replication n = 46, 45, 45.

Chapter 4

The Mixed Lineage Leukemia 4 (MLL4) methyltransferase complex is a mediator of Transforming Growth Factor beta (TGF-β)-activated gene transcription.

Roy Baas[1], Hetty A.A.M. van Teeffelen[1], Sjoerd J.D. Tjalsma[1] and H.Th. Marc Timmers[1]

Manuscript submitted to *Transcription*

[1]Molecular Cancer Research and Stem Cells, Regenerative Medicine Center, Center for Molecular Medicine, University Medical Center Utrecht, Uppsalalaan 8, 3584 CT Utrecht, The Netherlands

Abstract

SMA and MAD related (SMAD)-mediated Transforming Growth Factor-β (TGF-β) signalling is required for various cellular processes. The activated heterotrimeric SMAD protein complexes can associate with nuclear proteins like the histone acetyltransferases p300, PCAF and the Mixed Lineage Leukemia 4 (MLL4) subunit Pax Transactivation domain-Interacting Protein (PTIP) to regulate gene transcription.

Here we investigate the functional role of PTIP and PTIP Interacting protein 1 (PA1) in relation to TGF-β activated SMAD signaling. To study SMAD-mediated signaling we immunoprecipitated PTIP and PA1 and identified the TGF-β and not BMP-specific SMADs as interacting proteins. Knockdown experiments of MLL4 and the subunits PA1 and PTIP confirm TGF-β specific genes to be regulated by the MLL4 complex, linking TGF-β signaling to the MLL4 methyltransferase complex.

Introduction

Transforming Growth Factor-β (TGF-β) and Bone Morphogenetic Protein (BMP) signaling transduction routes in the cell are implicated in various cellular processes. Both signal transduction routes start with the activation of a transmembrane receptor pair located at the plasma membrane. Receptor activation causes phosphorylation of downstream receptor-activated proteins called SMA and MAD related (SMADs) (308-310). These Receptor activated SMADs (R-SMADs) in turn bind to SMAD4 as a heterotrimeric protein complex and regulate gene transcription. Depending on the activating signaling ligand different R-SMAD proteins become activated. BMP causes phosphorylation of SMAD1/5/9 while TGF-β causes phosphorylation of SMAD2/3. The SMAD mediated signaling routes are regulated by the SMAD6 and SMAD7 negative regulators.

Activated SMAD complexes can be bound by transcription factors and co-localize to specific genomic locations (317,333). Furthermore, SMAD complexes can bind to the coactivator complex p300 and the PCAF histone acetyltransferases (334-336). Recruitment of p300 or the PCAF subunit of the SAGA and ATAC complexes (337) stimulates gene transcription via the acetylation of histone tails in proximity of the complex. This causes the DNA to be less tightly wrapped around the acetylated histones and makes the DNA more accessible for basal transcription factors. Additionally, proteins containing double bromodomains or certain plant homeodomains (PHD) fingers can recognize and bind these acetylated histones (30,338-340). By altering the local chromatin environment the transcriptional status of the cell is changed and local transcription is promoted.

Other histone modifications that stimulate transcription are H3K4 mono- and trimethylation. H3K4 trimethylation is present on the promoters of active genes and is a binding site for the TFIID subunit TAF3 (68). H3K4 monomethylation is found at active enhancer sites. The deposition of H3K4 methylation is facilitated by the six SET1/MLL methyltransferase complex types present in human cells. SET domain-containing protein 1A (SET1A) and SET1B complexes are responsible for the bulk H3K4 trimethylation (72) while Mixed Lineage Leukemia 3 (MLL3) and MLL4 complexes are maintaining mono-methylation at enhancer sites (74,75).

These SET1/MLL complexes all contain the WDR5, RBBP5, ASH2L and DPY30 proteins that make up the WRAD core, which is always present (76). Other interacting proteins differ between the different SET1 and MLL complexes and are sometimes absent. Additional to the WRAD core, the MLL3 and MLL4 complexes are composed of Pax Transactivation domain-Interacting Protein (PTIP) and PTIP Interacting protein 1 (PA1) (76,77,341). Apart from interacting in the MLL3 and MLL4 methyltransferases, PTIP and PA1 also form a protein complex that can exist and act independently during DNA damage and IgH class switching (76,342,343). SMAD2 was previously identified as an interaction partner of PTIP (76,344). However the link to the MLL3 and MLL4 methyltransferases has not been clarified yet.

Here we study the relationship between SMAD-mediated signaling in the context of the gene regulatory function of PTIP. In a comprehensive approach we determined the interaction of PTIP and PA1 proteins with all SMADs to identify a link between the TGF-β SMADs and PTIP and PA1. Gene expression analysis following PTIP and PA1 knockdown confirms PTIP and PA1 to be required for TGF-β gene activation. Expression analysis after

MLL3 and MLL4 knockdown shows this activation to be MLL4 but not MLL3 dependent. This links SMAD signaling to the MLL4 methyltransferase complex.

Results

PTIP and PA1 bind to TGF-β activated SMADs

Interaction of PTIP and SMAD2 has been identified and studied before (344) and was confirmed previously using an unbiased proteomics approach (76). However, SMAD2 is only one of the eight SMADs of this transcription factor family. SMADs proteins resemble to each other (345) and therefore more SMADs could potentially interact with PTIP. To systemically identify the SMAD proteins binding to PTIP HEK 293T cells were co-transfected with HA-tagged SMAD and GFP-PTIP expression plasmids and affinity purification was performed via the GFP-tag on PTIP. Immunoblotting with the HA-antibody revealed presence of the TGF-β activated SMAD2, 3 and 4 in the GFP purifications, while BMP stimulated SMAD1 and 5 were absent (Fig. 1A). Interestingly, SMAD9 and MENIN are also present in low levels. PTIP and PA1 are known to be subunits of the MLL3 and MLL4 methyltransferase complexes and as a PTIP can form a heterodimer with PA1. To examine interaction with PA1 HEK 293T cells were co-transfected with HA-SMADs and GFP-PA1. GFP-affinity purification of the transfected cell lysates also showed interaction with the TGF-β SMADs (Fig. 1B). Interestingly, SMAD6 and SMAD7 also bind to PA1 while this binding is not observed with PTIP. Again SMAD9 and MENIN are present at low levels. Overall, these experiments indicate that the TGF-β SMADs but not the BMP SMADs interact directly with PTIP and PA1.

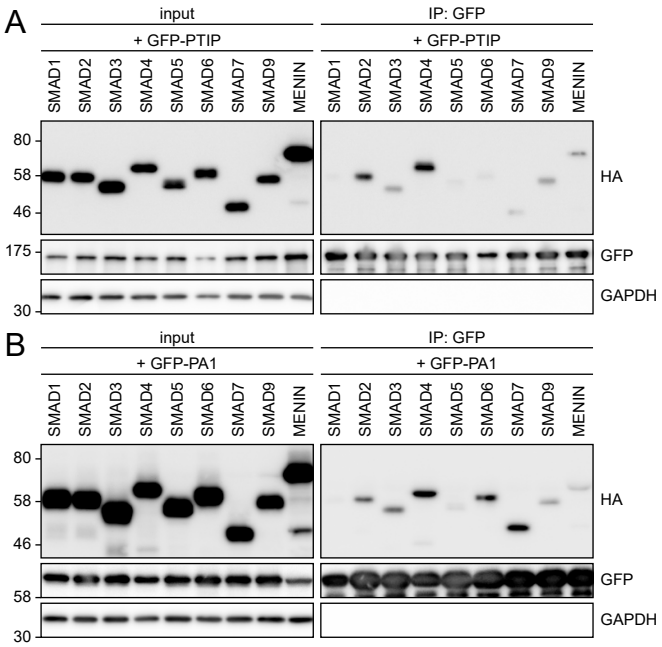


Figure 1. TGF-β activated RSMADs interact with PTIP and PA1. (A-B) HEK 293T cells were co-transfected with HA-tagged SMADs and GFP-tagged PA1 or PTIP expression plasmids. Affinity purification of GFP fusions was performed on whole cell extracts. Samples were analyzed by immunoblotting using HA, GFP and GAPDH antibodies. (A) Co-immunoprecipitation for GFP-PTIP. (B) Co-immunoprecipitation for GFP-PA1. For input control 1.5% of the extract was analyzed.

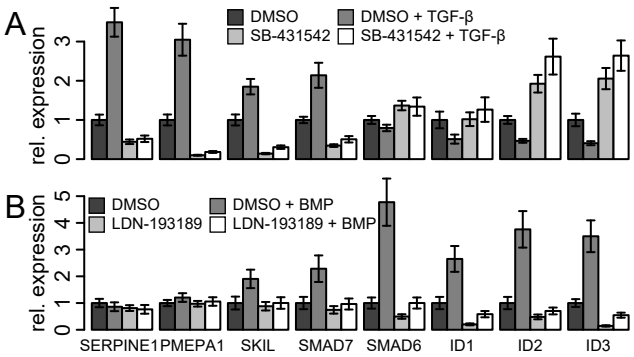


Figure 2. Investigation of TGF-β and BMP target genes in U-2 OS cells. (A-B) Gene expression analysis of TGF-β and BMP target genes. (A) U-2 OS cells were treated with 10 μM SB-431542 inhibitor for 24 h and subsequently stimulated with 10 ng/ml TGF-β for 6 h as indicated. (B) U-2 OS cells were treated with 100 nM LDN-193189 inhibitor for 24 h and subsequently stimulated with 10 ng/ml BMP for 6 h as indicated. (A-B) Samples were analyzed using qRT-PCR. Bars represent average expression corrected for ACTIN and normalized to the siNT control. Error bars represent standard deviation. Technical replication n = 3 for all samples.

Investigation of TGF-β and BMP target gene in U-2 OS

Previous studies have identified TGF-β and BMP signaling target genes in cell systems like the HaCaT keratinocytes and osteoblast BMP inducible C2C12 myoblast cell lines (205-208). The human osteosarcoma derived cell line U-2 OS is easily transfectable and efficient siRNA-mediated knockdown can be obtained (346). To investigate a functional link of TGFβ with PTIP and PA1, the mRNA expression of TGF-β and BMP target genes were tested for TGF-β or BMP sensitivity in U-2 OS cells (Fig. 2A and 2B). TGF-β treatment increased mRNA expression of the *SERPINE1*, *PMEPA1*, *SKIL* and *SMAD7* genes 2 to 4-fold, while inhibition of TGF-β signaling with the small molecule inhibitor SB-431542 reduced target mRNA expression by 50–90% (Fig. 2A). *SMAD6* expression remained unaffected. The well-known BMP target genes *ID1*, *ID2* and *ID3* reduced 50–60% in expression on TGF-β stimulation, while TGF-β inhibition stimulated expression of these genes to approximately 3-fold. BMP stimulation induced gene expression of *SKIL*, *SMAD7*, *SMAD6*, *ID1*, *ID2* and *ID3* 2 – 5 fold, while *SERPINE1* and *PMEPA1* were unaffected (Fig. 2B). Based on the response to TGF-β and BMP stimulation we divided the genes into four groups (Table 1). The first group is TGF-β responsive and consists of genes *SERPINE1* and *PMEPA1*. The second group is both TGF-β and BMP responsive and consists of *SKIL* and *SMAD7*. *SMAD6* is BMP responsive. And the fourth group consists of genes that become activated by BMP, but are repressed on TGF-β stimulation. Expression of the members of the MLL3 or MLL4 methyltransferase complexes is not affected by stimulation or repression of either TGF-β or BMP (Fig. S1A and S1B).

Table 1: Genes and TGF-β and BMP responses

Gene	TGF-β	BMP	Group
SERPINE1	↑	-	TGF-β responsive
PMEPA1	↑	-	TGF-β responsive
SKIL	↑	↑	TGF-β and BMP responsive
SMAD7	↑	↑	TGF-β and BMP responsive
SMAD6	-	↑	BMP responsive
ID1	↓	↑	BMP responsive, TGF-β repressed
ID2	↓	↑	BMP responsive, TGF-β repressed
ID3	↓	↑	BMP responsive, TGF-β repressed

PTIP and PA1 knockdown reduce TGF-β specific gene expression

To investigate the interaction between TGF-β signaling with PTIP and PA1 in context of MLL3 and MLL4 protein complexes the knockdown efficiencies of *PTIP* and *PA1* expression were examined

in U-2 OS using siRNA (Fig. 3). Expression was tested using qRT-PCR for all genes and immunoblot analysis. mRNA expression analysis showed knockdown efficiencies of 70–95% for *PA1*, *PTIP* and *SMAD4* (Fig. 3A). For *MLL3* and *MLL4* a single round of knockdown yielded mRNA knockdown efficiencies of approximately 50% (data not shown). To obtain a better knockdown cells were transfected twice to obtain a knockdown efficiency of 70–90% (Fig. 3B). Immunoblotting confirmed the expression analysis and, in addition, this showed a strong reduction of H3K4me3 levels with *ASH2L* knockdown (Fig 3C).

Next, expression levels of the previously-tested TGF-β and BMP target genes were investigated. *SERPINE1* and *PMEPA1* gene expression was reduced upon *PA1*, *PTIP* and the combination of *PA1* and *PTIP* knockdown confirming a functional link (Fig. 4A). Interestingly, the BMP stimulated genes *ID1*, *ID2* and *ID3* (Fig. 2B) were not affected after *PA1* or *PTIP* knockdown (Fig. 4A), suggesting a specific link between TGF-β signaling and PA1 and PTIP to this subgroup of genes, but not to BMP stimulated genes. *SMAD6* mRNA expression is induced after *PA1* knockdown. However, this does not seem to be linked to PTIP, because knockdown of *PTIP* does not affect expression of *SMAD6*. PA1 and PTIP are present in both the MLL3 and MLL4 complexes, but also as an independent hetero-dimeric complex. To investigate involvement of the MLL3 and/or MLL4 histone methyltransferase complexes, we determined the effects after *MLL3* or *MLL4* siRNA knockdown. Interestingly, knockdown of *MLL4* but not *MLL3* resulted in reduced expression of the *SERPINE1* and *PMEPA1* genes suggesting that PA1 and PTIP mediate TGF-β SMAD transcription as part of the MLL4 complex (Fig. 4B). Knockdown of *MLL4* resulted in induced expression of BMP induced genes *SMAD6*, *ID1*, *ID2* and *ID3*.

Overall, these data indicate that TGF-β mediated transcription activation involves the PA1 and PTIP subunits of the histone methyltransferase complex MLL4.

Discussion

SMAD-mediated signaling culminates in the altered mRNA expression of selected gene sets. Although TGF-β and BMP signaling is heavily studied, relatively little is known about the downstream coactivator complexes that cooperate with SMAD signaling to

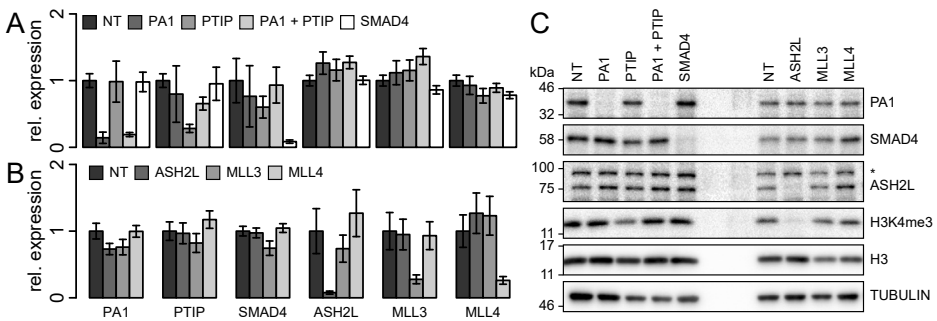


Figure 3. Knockdown efficiency of siRNA treated U-2 OS. (A-C) Cells were reverse transfected with siNT, siPA1, siPTIP, siPA1 + siPTIP or siSMAD4 or with siNT, siASH2L, siMLL3 or siMLL4 and treated again after 24 h and collected after 72 hours. (A-B) Samples were analyzed using qRT-PCR. Bars represent average expression corrected for ACTIN and normalized to the siNT control. Error bars represent standard deviation. Technical replication n = 3 for all samples. (C) Samples were analyzed on immunoblot using PA1, SMAD4, ASH2L, H3K4me3, H3 and Tubulin antibodies. * indicates nonspecific antibody background binding.

facilitate target gene transcription. Both TGF-β and BMP SMAD signaling complexes have been shown to recruit and interact with the coactivating proteins CREB binding protein (CBP), p300 and the PCAF subunits of the SAGA and ATAC histone acetyltransferases to promote target gene transcription (334-336). In contrast, we show that TGF-β but not BMP activated SMADs cooperate with the MLL4 methyltransferase complex suggesting a TGF-β specific target gene response pathway. The SMAD2, 3 and 4 proteins interact directly with the PA1 and PTIP proteins in transient transfection experiments. Knockdown of *PA1*, *PTIP* and *MLL4* result in reduced expression of the TGF-β responsive genes *SERPINE1* and *PMEPA1*.

MLL3 and MLL4 methyltransferase complexes are required for methylation of H3K4 and for demethylation of H3K27 of enhancer elements in the genome (74,75). Although both MLL3and MLL4 proteins methylate enhancer chromatin, the MLL4 protein is essential in early embryonic mouse development, while MLL3 knockout mice die around birth (318). Recently, MLL4 was described to be essential for cell fate transition by recruiting p300 to the enhancer sites, independently of its methyltransferase activity (347). Furthermore, motif analysis showed SMAD2/3 binding sites to be enriched at MLL4 occupied active enhancers (347).

Both SMADs and MLL methyltransferases are often found mutated in various forms of cancers (252). *SMAD2* and *SMAD4* are mutated in 5.7 and 9.8% respectively in colon and rectal carcinomas, while MLL4 is most often (but not exclusively) mutated in bladder urothelial carcinoma and uterine corpus cell carcinomas (7.1 and 8.3% respectively). Furthermore, *MLL4* is mutated in 2.1% of all colorectal cancers. Enhancer profiling in colon cancers has shown thousands sites that either gained or lost K4me1 levels compared to

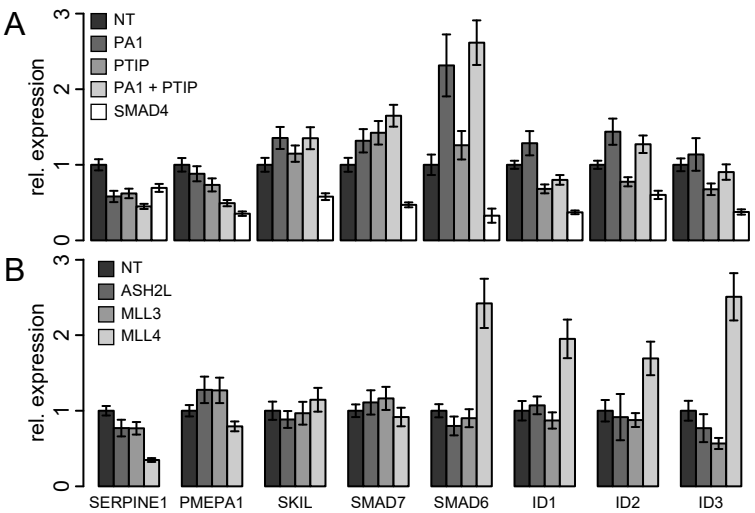


Figure 4. TGF-β target genes are selectively affected by PA1 and PTIP knockdown. (A-B) Gene expression analysis of TGF-β and BMP target genes and SET/MLL complex member genes. (A) U-2 OS cells were reverse transfected with siNT, siPA1, siPTIP, siPA1 + siPTIP or siSMAD4. (B) Cells were reverse transfected with siNT, siASH2L, siMLL3 or siMLL4 and treated again after 24 h to enhance knockdown efficiency. (A-B) Samples were analyzed using qRT-PCR. Bars represent average expression corrected for ACTIN and normalized to the siNT control. Error bars represent standard deviation. Technical replication n = 3 for all samples.

healthy tissues (348). The link between SMAD mediated TGF-β signaling and the MLL4 methyltransferase allows for new opportunities to study and possibly identify treatment for these cancers and warrants further research.

Materials and methods

Plasmids

Gateway compatible pENTR-SMAD vectors (325,345) were recombined with pMT2SM-HA in an LR reaction according to manufacturer's instructions (Thermo Fischer Scientific). GFP-PTIP and GFP-PA1 expression vectors were described before (76).

Cell culture

Cells were grown in Dulbecco's Modified Eagle's Medium (DMEM) containing 1 g/L or 4.5 g/L glucose for U-2 OS or HEK 293T respectively supplemented with 10% v/v fetal bovine serum, 10 mM L-glutamine, 100 U/mL penicillin and 100 U/mL streptomycin.

Stimulation and inhibition of TGF-β or BMP signaling

U-2 OS cells were counted and seeded 150,000 cells per well in a 6-wells plate. Cells were left overnight to attach. Cells were treated with 10 μM SB431542 (#616461, Calbiochem) or 100 nM LDN193189 (SML0559, Sigma- Aldrich) to inhibit TGF-β or BMP signaling respectively. After 24 h the cells were stimulated with 10 ng/mL TGF-β1 (#8915LC, Cell Signaling) or 10 ng/mL rhBMP-2 (355-BEC, R&D systems) for 6 h and collected for expression analysis.

siRNA transfection

U-2 OS cells were transfected using a reverse transfection protocol using Hiperfect (Qiagen) according to the manufacturer's instructions. In short: 125 μL of a SMARTpool of four siRNAs (25 nM each) was dispensed in a 6-wells plate (657165, Greiner bio-one). Transfection mix (10 μL Hiperfect in total volume of 375 μL Opti-MEM [Invitrogen] per well) was added followed by 30 minutes incubation at room temperature. After this 70,000 U-2 OS cells/well were added in 2 mL. Cells were collected 72 h after transfection.

Transient transfection, whole cell extracts and immunoprecipitation – HEK 293T cells were counted and seeded at 150,000 cells per well in a 6-wells plate. Cells were left overnight to attach and expression vectors were transfected using polyethyleneimine (PEI). Cells were washed once with PBS and cell extracts were collected in whole cell extract buffer (50 mM Tris-HCl pH 8.0, 10% glycerol, 100 mM NaCl, 10 mM MgCl2, 0.5 mM DTT, 0.1% NP-40 and 1x protease inhibitor [Complete Protease Inhibitor Cocktail, Roche]). GFP affinity purification was essentially performed as described up to sample elution (268).

Immunoblot analysis

Cells were lysed in sample buffer (160 mM Tris-HCl pH 6.8, 4% SDS, 20% glycerol, 0.05% bromophenol blue) and equal amounts of protein were separated using a 10% or 15% SDS-PAGE and transferred onto PVDF membrane. The membrane was developed with the appropriate antibodies and ECL and scanned using ChemiDoc Touch Imaging System (Bio-Rad).

Antibodies

Antibodies were used from the following sources: ASH2L (kind gift from Winship

Herr), GAPDH (clone 6C5, mAb374, Millipore), GFP (kind gift from Geert Kops), H3 (ab1791; Abcam), H3K4me3 (pAb003-050, Diagenode) (346), HA (3F10-HRP, Roche), PA1 (kind gift from Junjie Chen), SMAD4 (#9515, Cell Signaling), α-Tubulin (DM1A, CP06, Calbiochem).

Expression analysis qRT-PCR

Cells were lysed and total RNA was isolated from cells using RNeasy kit (Qiagen) including a DNase treatment using RNase-Free DNase (Qiagen) according to manufacturer's instructions. 500 ng of total RNA was used for cDNA synthesis (Superscript II, Invitrogen) using random primers. A 25 µL qRT-PCR reaction was performed in a CFX Connect Real-Time PCR Detection System (Bio-Rad) using iQ SYBR Green Supermix (Bio-Rad) Primers used for qRT-PCR can be found in Table 2.

Acknowledgements

The authors would like to thank Simona Antonova, MSc and dr. Koen Dreijerink, MD for their critical reading of the manuscript and all the members of the Timmers group for their helpful discussions and suggestions.

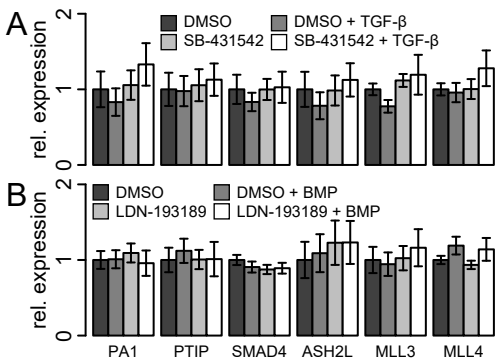


Figure S1. MLL3/4 complex member expression analysis. (A-B) Gene expression analysis of SET/MLL complex member genes. (A) U-2 OS cells were treated with 10 µM SB-431542 inhibitor for 24 h and subsequently stimulated with 10 ng/ml TGF-β for 6 h as indicated. (B) U-2 OS cells were treated with 100 nM LDN-193189 inhibitor for 24 h and subsequently stimulated with 10 ng/ml BMP for 6 h as indicated. (A-B) Samples were analyzed using qRT-PCR. Bars represent average expression corrected for ACTIN and normalized to the siNT control. Error bars represent standard deviation. Technical replication n = 3 for all samples.

Table 2: Primers used for qRT-PCR.

Primer	Sequence
ACTIN Fw	AGAAAATCTGGCACCACACC
ACTIN Rv	AGAGGCGTACAGGGATAGCA
PA1 Fw	CCCAGAGCGAAGAGGAGAGA
PA1 Rv	GTGGGCATGTGTGGTTTTTCC
PTIP Fw	CATCGACCCGCAGGTTATTCA
PTIP Rv	GAGCTTCTCCCACCTCTGGA
ASH2L Fw	TGCTGTGTCTACTAGTGGGAATTT
ASH2L Rv	TTGGCTCCTCGTCCTTTTC
MLL3 Fw	CCACGAAAACAAAGAGGACAG
MLL3 Rv	TGGGTGCTTACACTTACACAAGAT
MLL4 Fw	GTCTTCACTGACGCCTCTCC
MLL4 Rv	GCAGCATGTCAGCCTCTTTT
SMAD4 Fw	CCTGTTCACAATGAGCTTGC
SMAD4 Rv	GCAATGGAACACCAATACTCAG
KDM6A Fw	CTGCACAAGTAAAAGCAACTGTC
KDM6A Rv	CTTTTGGAGATACTGAATAGCATAGC
SERPINE1 Fw	AAGGCACCTCTGAGAACTTCA
SERPINE1 Rv	CCCAGGACTAGGCAGGTG
PMEPA1 Fw	CTGTCTGCACGGTCCTTCAT
PMEPA1 Rv	CCACAGGCATCCTTCTGAGG
SKIL Fw	GAGGCTGAATATGCAGGACAG
SKIL Rv	CTTGCCTATCGGCCTCAG
SMAD7 Fw	TCAAGAGGCTGTGTTGCTGTG
SMAD7 Rv	TGGTTTGAGAAAATCCATCGGGT
SMAD6 Fw	TTCTCAGACGCCAGCATGTC
SMAD6 Rv	GCCCTGAGGTAGGTCGTAGA
ID1 Fw	CAGGGACCTTCAGTTGGAGC
ID1 Rv	AACGCATGCCGCCTCG
ID2 Fw	TGACCACCCTCAACACGGAT
ID2 Rv	CATGAACACCGCTTATTCAGCC
ID3 Fw	GAGTCCCGAGAGGCACTCA
ID3 Rv	CTCGGCTGTCTGGATGGGAA

Addendum A

Interactome of the KDM6A/UTX subunit of the MLL3 and MLL4 histone methyltransferase complexes.

Roy Baas[1], Sjoerd J.D. Tjalsma[1], Robert van Es[1], Harmjan R. Vos[1] and H.Th. Marc Timmers[1]

[1]Molecular Cancer Research and Stem Cells, Regenerative Medicine Center, Center for Molecular Medicine, University Medical Center Utrecht, Uppsalalaan 8, 3584 CT Utrecht, The Netherlands

Abstract

Histone modifications like acetylation, methylation and ubiquitination are closely related to local transcriptional status. These modifications can be placed and removed by their respective catalytic enzymes that can be found in multimeric protein complexes, like the H3K4 methyltransferase complexes MLL3 and MLL4, the Polycomb repressive deubiquitinase (PR-DUB) complex and the less-studied histone deacetylase complex MiDAC. KDM6A/UTX is known to be integrated in the MLL3/4 complexes, as well in the PR-DUB complex. To study the association of UTX with these multimeric protein complexes we generated stable UTX expressing cell lines. Immunoprecipitation followed by quantitative mass spectrometry analysis confirms interaction of UTX with MLL3/4 and PR-DUB complexes. Furthermore, we show interaction of the MiDAC complex with UTX, which suggests a new level of chromatin crosstalk between these three protein complexes.

Introduction

Histone modifications influence local chromatin environment and relate to transcriptional status of nearby genes. Modifications include acetylation, methylation and ubiquitination of various amino acid residues and dependent on the location can have different outcomes. Histones can become acetylated by histone acetyl transferases (HATs). Generally, acetylation of any lysine residue on histone tails result in the loss of positive charge on the respective lysine residue. Wrapping of the negatively charged DNA around these modified histones becomes less tight, resulting in increased accessibility of the DNA for proteins such as the general transcription machinery. Additionally, acetylated lysines are recognized by proteins that carry bromodomains or certain plant homeodomains (PHD) fingers (338-340). For example, the TBP associated factor 1 (TAF1) subunit of the Transcription Factor IID (TFIID) complex contains a double bromodomain that facilitates TFIID recruitment and stimulates transcription (30). Additionally, the DPF3 subunit of the BAF remodeling complex contains a double PHD finger that binds methylated and acetylated H3 and H4 histones (339). The acetylation of histones can be counteracted by histone deacetylases (HDACs), like the NuRD, Sin3 and CoREST complexes (42,43,45). A less understood HDAC protein complex is the mitotic deacetylase (MiDAC) complex, which consists of HDAC1/2, ELMSAN and DNTTIP1 and its role in transcription regulation remains an open question (46-48).

An alternative histone modification related to an active chromatin environment includes methylation of histone 3 on lysine 4 (H3K4). H3K4me1 is found at enhancer elements, while H3K4me3 predominantly marks active transcription start sites. H3K4 methylation is facilitated by the SET domain-containing protein 1A/B (SET1A/B) or Mixed Lineage Leukemia 1-4 (MLL1-4) catalytic subunits of the SET/MLL methyltransferase complexes. Genome-wide bulk H3K4me3 is deposited by the SET1A and SET1B complexes presumably via their interaction with the transcribing RNA polymerase II (72), while MLL1 and MLL2 gene-specifically deposit H3K4me3, for example on promoters of *Hox* genes (73). MLL3 and MLL4 facilitate H3K4 mono-methylation of enhancer elements throughout the genome (74,75). The SET1/MLL protein complexes are composed of the WDR5, RBBP5, ASH2L and DPY30 proteins that make up the core of the complex and are shared between all six different complexes (76). The SET1/MLL complexes can be partially distinguished by the composition of complex-specific proteins that are dedicated to a subset of SET1/MLL complexes. SET1A/B and MLL1/2 share the HCFC1 and OGT proteins and are furthermore composed of CFP1 and WDR82 or MENIN and PSIP respectively (76). MLL3/MLL4 complexes contain PTIP, PA1, NCOA6 and KDM6A/UTX as specific subunits (76,115).

Ubiquitously-transcribed TPR protein on the X chromosome (UTX, official name is KDM6A) is a lysine demethylase that in conjunction with MLL3 and MLL4 facilitates generation of transcription permissive chromatin by removing the repressive H3K27me2/3 mark via its JmjC catalytic domain (115). Deposition of H3K27me3 is catalyzed by the methyltransferase EZH2 of the Polycomb Repressive Complex 2 (PRC2) (349). Subsequently, PRC1 is recruited to the H3K27me3 mark and ubiquitinates H2AK119 (350), resulting in a repressive chromatin environment.

The removal of H2AK119ub1 is performed by the Polycomb repressive deubiquitinase (PR-DUB) complex (351). Deubiquitinating enzymes (DUBs) possess peptidase activity that can release ubiquitin from the lysine residue to which the ubiquitin is linked. The PR-DUB

complexes are multimeric protein complexes, that consists of the catalytically DUB subunit BAP1 and one of three activating ASXL proteins. Furthermore, a copy of MBD5 or MBD6, FOXK1 or FOXK2, KDM1B and OGT may be present (114,268). Interestingly, a single study identified KDM6A/UTX to be associated with BAP1, ASXL2 and HCFC1 proteins of the PR-DUB complex (144).

To further study the role of KDM6A/UTX in the context of multimeric protein complexes we generated stable cell lines containing doxycycline-inducible GFP-UTX fusion constructs. Immunoprecipitation followed by quantitative mass spectrometry analysis confirms KDM6A/UTX to be present in MLL3, MLL4 and PR-DUB protein complexes. Furthermore we show a new interaction of KDM6A/UTX with the MiDAC complex. To our knowledge this interaction has not been described and potentially reveals a new possible mechanism of crosstalk between these three multimeric histone modifying complexes.

Results

GFP-UTX cell line generation and characterization

UTX is known to be integrated in the nuclear MLL3/4 methyltransferase complex (76,115). In a previous microscopy assay, FLAG-UTX was shown to localize to the cytoplasm and a C-terminal nuclear localization signal (NLS) fusion resulted in proper nuclear localization (Ramin Shiekhattar, personal communication). In order to study the interactome of the KDM6A/UTX protein, (for brevity we will use UTX in this section) stable doxycycline-inducible HeLa cell lines containing GFP-UTX or GFP-UTX-NLS fusion constructs were generated using the Flp-In T-REx system (345). To characterize the expression and localization of the fusion proteins cells were analyzed using confocal fluorescence

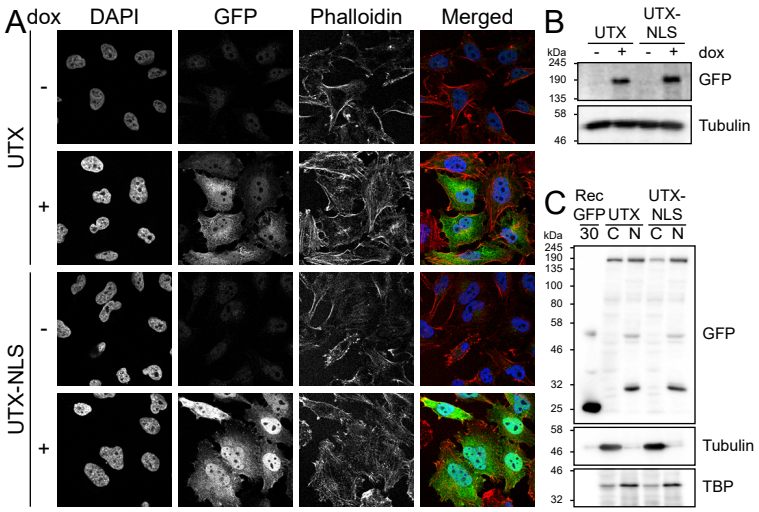


Figure 1. Analysis of GFP-UTX expression. (A) Doxycycline induction test for confocal microscopy shows GFP-fusion protein localization. HeLa cell lines were seeded on coverslips and induced with 1 µg/ml doxycycline for 16 hours prior to fixing. Fixed cells were stained using DAPI and phalloidin-633 and analyzed using confocal microscopy. (B) Doxycycline induction test for immunoblot analysis. HeLa cell lines were induced with 1 µg/ml doxycycline for 16 hours prior to harvesting or harvested without induction. Blots were probed using GFP or αTubulin antibodies. (C) Immunoblot analysis of 50 µg cytoplasmic and nuclear extract derived from doxycycline induced HeLa cell lines. Recombinant GFP (30 ng) was run in parallel for reference. Samples were analyzed by immunoblotting using GFP, TBP and αTubulin antibodies.

microscopy (Fig. 1A). An antibody against GFP was used to enhance signal intensity of the GFP-fusion protein. Cells that were not induced with doxycycline showed only very low cross-reactivity of the antibody. Confocal analysis showed the GFP-UTX fusion protein to be localized in both cytoplasm and nucleus, which confirms other observations (Ramin Shiekhattar, personal communication). In contrast, the GFP-UTX-NLS fusion protein is mainly localized in the cell nucleus with only residual signal remaining in the cytoplasm. This difference in localization could not be explained in fusion protein expression difference as observed by immunoblotting, but rather depends on the NLS (Fig. 1B). Generation and immunoblot analysis of cytoplasmic and nuclear extracts confirms nuclear preference of GFP-UTX-NLS over GFP-UTX (Fig. 1C). For GFP-UTX a band of equal intensity is observed in cytoplasmic and nuclear fraction while for GFP-UTX-NLS the band in the cytoplasmic fraction is clearly reduced. Overall these data show we successfully generated stable inducible HeLa cell lines expressing GFP-UTX and GFP-UTX-NLS.

UTX interactome identification using quantitative mass spectrometry

The difference in localization between GFP-UTX and GFP-UTX-NLS suggests that UTX might be incorporated in different protein complexes. To systematically investigate the interaction partners of UTX mass spectrometry analysis was performed on GFP-affinity-purified cytoplasmic and nuclear fractions of both GFP-UTX and GFP-UTX-NLS fusion constructs. All GFP-affinity purifications were performed in triplicate and control agarose beads without GFP-coupled antibodies were performed in triplicate for all extracts. All 24 affinity purifications were measured using a Velos mass spectrometer and data files were simultaneously analyzed using MaxQuant. Perseus and R were used to generate volcano plots for each extract (Fig. 2).

Affinity-purified GFP-UTX from the cytoplasmic fraction showed interaction with the known SET1/MLL methyltransferase core subunits WDR5, RBBP5, ASH2L and DPY30 (Fig. 2A). Furthermore, the MLL4 and its specific subunit PTIP are enriched. Apart from the MLL4 methyltransferase complex, the known interactor ZNF281 was observed (Fig. 2A) (115).

Interestingly, additional interaction partners are observed when GFP-UTX was purified from the nuclear fraction (Fig. 2B). Interaction with the MLL4 complex and its core subunits are still observed and additionally PA1, NCOA6 and MLL3 are gained. Stoichiometry analysis shows a 3:1:1 binding ratio for WDR5:RBBP5:ASH2L that is consistent throughout the different immunoprecipitations Relative binding of DPY30 is quite variable (Fig. 3A). A ratio of 1:1:0.3 for PA1:PTIP:NCOA6 is observed in the nuclear fractions with no binding of PA1 in the cytoplasmic fraction, as observed before in the volcano plots (Fig. 2A). Apart from the MLL3/MLL4 methyltransferase complexes the previously-described interaction with subunits from the PR-DUB complex was observed. A clear enrichment was observed for PR-DUB subunits ASXL2, BAP1, KDM1B and MBD6 with a relative stoichiometry of 1:1:1:0.5 for BAP1:MBD6:ASXL2:KDM1B was determined. However, other PR-DUB subunits like ASXL1, MBD5, OGT, HCFC1, FOXK1 and FOXK2 are clearly absent, which suggests that UTX may only be present in a subset of variant PR-DUB complexes. Interestingly, new interactions are observed with DNTTIP and ELMSAN, which are two subunits belonging to the MiDAC protein complex.

No MLL3/MLL4 interaction partners were observed in affinity-purifications of GFP-

UTX-NLS from the cytosol, which supports NLS-induced localization to the nucleus (Fig. 2C). The nuclear pool of GFP-UTX-NLS confirms observations made with GFP-UTX for interaction with the MLL3, MLL4, PR-DUB and ELMSAN protein complexes (Fig. 2D). Additionally, interaction with HDAC2, the third subunit of the MiDAC complex is observed. Stoichiometry analysis for the MiDAC complex shows relative binding in a ratio of 1:0.5:0.25 for ELMSAN1:DNTTIP1:HDAC1/2.

To identify the relative integration of UTX into the different MLL3/4, PR-DUB and MiDAC complexes, stoichiometry was compared between the catalytic subunits of each complex relative to the total amount of UTX (Fig. 2D). 30 and 10% of total bound UTX is in complex with MLL3 and MLL4, suggesting unbound UTX is present and overrepresented compared to the bound protein complexes. MLL3/MLL4 complexes comprised the majority of the UTX bound complex whereas BAP1 and HDAC1/2 were present at approximately 300 fold lower stoichiometry.

Overall, these results show the KDM6A/UTX protein assembles in the MLL4 methyltransferase complex in the cytoplasm. Upon nuclear import KDM6A/UTX becomes

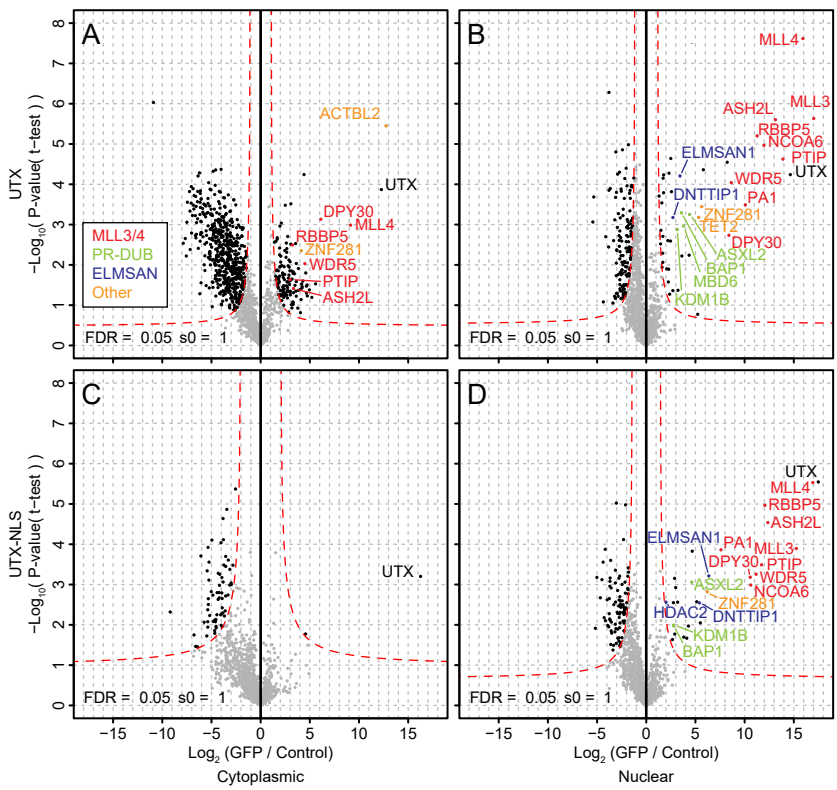


Figure 2. Mass spectrometry analysis for cytoplasmic and nuclear fractions. Volcano plot representation for cytoplasmic and nuclear fraction affinity purifications. GFP affinity purification enrichment (x-axis) were plotted against t-test p-value (y-axis). Subunits of MLL3 and MLL4 complexes (red), PR-DUB complex (green), ELMSAN complex (blue) and other (orange) complexes are highlighted using their respective colors and UTX is indicated. FDR and s0 parameters are indicated in the plots. (A) UTX cytoplasmic extract. (B) UTX nuclear extract. (C) UTX-NLS cytoplasmic extract. (D) UTX-NLS nuclear extract.

integrated as expected in MLL3, MLL4 and PR-DUB complexes in different amounts. In addition, we observe that a small amount of the KDM6A/UTX protein associates with the MiDAC complex.

Discussion

KDM6A/UTX is a histone demethylase for H3K27, which is known to be integrated in multiprotein complexes like MLL3, MLL4 and PR-DUB and, thereby it contributes to transcriptional control. Removal of H3K27 methylation and gain of H3K4 methylation and H3K27 acetylation on enhancers and promoter elements allows recruitment of transcription factors that initiate transcription. In this study interactome datasets were generated for the cytoplasmic and nuclear pools of KDM6A/UTX. Interaction was confirmed for these protein complexes and the MiDAC complex was identified as a novel interaction partner.

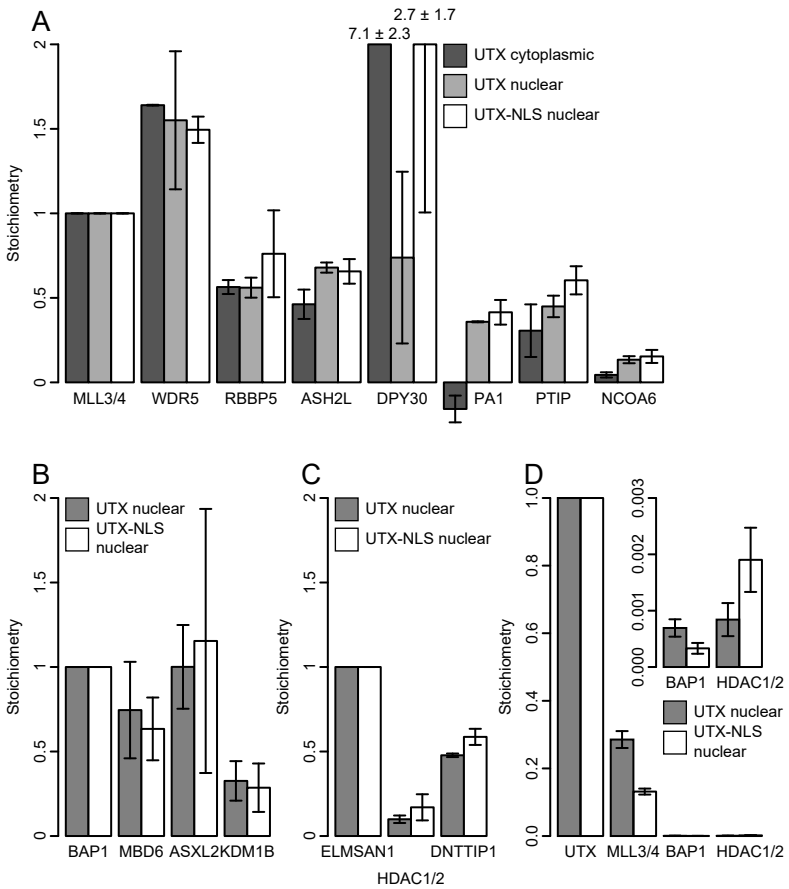


Figure 3. Stoichiometry analysis of UTX incorporated complexes. (A-C) Stoichiometry was determined using iBAQ values. Bars represent average relative binding. Error bars represent standard deviation. Technical replication n = 3 for all samples. (A) Stoichiometry for MLL3 and MLL4 complex members relative to the combined MLL3 and MLL4 proteins. (B) Stoichiometry for PR-DUB complex members relative to BAP1. (C) Stoichiometry for MiDAC complex members relative to ELMSAN1. (D) Stoichiometry for catalytic enzymes of the MLL3/4, PR-DUB and MiDAC complexes relative to UTX.

Control of protein activity is important for regulation of processes in the cell. Regulation can take place in a variety of ways, for example by sequestering proteins in compartments via incorporation in different localized complexes. The localization of GFP-UTX in both cytoplasm and nucleus suggests different pools of protein that could be regulated by incorporation in MLL3 and MLL4 complexes (Fig. 1A). Once, KDM6A/UTX is incorporated the complex can perform a function outside the nucleus or enter the nucleus once required where additional proteins can bind. This allows the multiprotein complex to perform its nuclear function. This is emphasized by the gain of additional interacting pool of proteins in the immunoprecipitation of the nuclear fraction.

The WRD5/RBBP5/ASH2L/DPY30 core of the SET1/MLL complexes was previously identified in a 1:1:1:6 stoichiometry (76). The UTX affinity purifications show a different stoichiometry of 3:1:1 for WDR5:RBBP5:ASH2L and a variable amount of DPY30. Interestingly, WDR5 is engaged in many different protein complexes (76). Therefore, WDR5 needs to be relatively more present in the cell and association of WDR5 in these different complexes could explain the elevated stoichiometry for WDR5. The variation of DPY30 stoichiometry is consistent with previous observed variation of DPY30 in SET1/MLL complexes. Proper quantification is compromised by its small protein size of 11 kDa (76).

Although the incorporation of KDM6A/UTX into the MLL3 and MLL4 complexes is well established in literature, only a single report linked KDM6A/UTX to the PR-DUB complex (144). By affinity purifying KDM6A/UTX we confirmed the incorporation into the PR-DUB complex. The demethylase activity of UTX removes the H3K27me2/3 repressive chromatin modifications and allows the H3K27 site to be acetylated. In conjunction with the deubiquitination activity of BAP1 on H2AK119ub1 this results in a chromatin environment permissive for recruitment of transcription factors activating transcription. Purification of KDM6A/UTX was performed on extracts obtained from cells that were not synchronized for cell cycle. Surprisingly, KDM6A/UTX was identified as an interactor of the MiDAC complex, which has been shown to be upregulated during mitosis (46). The demethylation activity of KDM6A/UTX contributing to activation would be counteracted by reduced transcription due the deacetylase activity of HDAC2 (352). Explanation for the crosstalk between these enzymes might be found in the fact that the MiDAC complex only exists during mitosis. Additional experiments that could involve nocodazole blocking to synchronize and block cells in mitosis are required to elucidate the function of KDM6A/UTX in this protein complex.

Comparison of the relative integration of UTX in the MLL3, MLL4, PR-DUB and MiDAC complexes shows the majority of KDM6A/UTX is present in the MLL3 and MLL4 methyltransferase complexes. The catalytic enzymes BAP1 and HDAC1/2 of the PR-DUB and MiDAC complexes were present at a 300-fold lower stoichiometry. This suggests that the main protein activity of KDM6A/UTX is in context of H3K4 methylation. However, this does not automatically imply the function is less important in context of the PR-DUB and MiDAC complexes. Changes in extracellular cues or cell cycle progression could shift the balance in which KDM6A/UTX integrates and performs separate molecular functions. Differentiation of cells is known to cause derepression of H3K27 methyl marks that could shift the balance towards a PR-DUB association. Cell cycle synchronization experiments coupled to quantitative mass spectrometry could be used to determine whether the association of KDM6A/UTX with MiDAC is regulated in a cell cycle-dependent manner.

Experimental procedures

Plasmids

PCR was performed on plasmids containing UTX cDNAs using primers containing attB1 or attB2 sites to obtain Gateway compatible constructs (Table 1) (325). pFLAG-CMV2-UTX and pFLAG-CMV2-UTX-NLS cDNAs were generous gifts from dr. Ramin Shiekhattar (115). PCR products were recombined with pDON201 in a Gateway BP reaction to obtain pENTR vectors according to manufacturer's instructions (Thermo Fischer Scientific, Waltham, MA). All pENTR vectors were confirmed using Sanger sequencing. UTX clones were identical to the Refseq (NM_021140.3). To obtain HeLa FRT compatible expression vectors pENTR vectors containing UTX variants were recombined with Gateway-compatible pCDNA5/FRT/TO/N-GFP in an LR reaction according to manufacturer's instructions (Thermo Fischer Scientific, Waltham, MA).

Cell culture

HeLa cells containing the Flp Recombination Target (FRT) site and Tet repressor were grown and cell lines were generated as described before (345). In short: cells were grown in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with the appropriate selection compounds. Vectors containing GFP-UTX fusion constructs were polyethyleneimine (PEI) co-transfected with pOG44 and cells were selected on selection medium.

Microscopy slide preparation

Cover glasses containing treated cells were washed once with phosphate-buffered saline (PBS). Cells were fixed by incubation with 4% formaldehyde in PBS for 20 minutes at room temperature. Cells were incubated with DAPI solution (2 mg/L 4',6-diamidino-2-phenylindole [Sigma-Aldrich, St. Louis, MO] in 50 mM Tris-HCl pH 7.4, 100 mM NaCl) to stain nuclei and phalloidin-633 (Phalloidin CruzFluor 633 conjugate, Santa Cruz, Dallas, TX) in PBS to stain actin. Cover glasses were mounted onto microscope slides (Thermo Fischer Scientific, Waltham, MA) using Immu-Mount (Invitrogen, Carlsbad, CA) and left to dry overnight at 4°C.

Confocal microscopy

Images were taken using a Leica TCS SP8 X confocal microscope with HC PL APO 63x/1.40 Oil CS2 objective. Samples were scanned using a white light laser using 405 nm for DAPI staining, 488 nm for GFP and 633 nm for phalloidin staining. Microscope gain settings were kept constant within experiments.

Immunoblot Analysis

Cells were seeded in 6-wells plates at a density of 30,000 cells/well. Cells were induced 16 hours prior to harvesting by addition of 1 µg/mL doxycycline to the culture media. Cells were lysed in sample buffer (160 mM Tris-HCl pH 6.8, 4% SDS, 20% glycerol, 0.05% bromophenol blue) and equal amounts of protein were separated using an 8% SDS-PAGE

Table 1. Primers used for cloning

Primer	Sequence
UTX GW Fw	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCAAATCCTGCGGAGTGTCG
UTX GW Rv	GGGGACCACTTTGTACAAGAAAGCTGGGTCTCAAGATGAGGCGGATGGTAATG
UTX NLS GW Rv	GGGGACCACTTTGTACAAGAAAGCTGGGTCTCATACCTTTCTCTTCTTTTTTGGATC

and transferred onto PVDF membrane. The membrane was developed with the appropriate antibodies and ECL and scanned using ChemiDoc Touch Imaging System (Bio-Rad, Hercules, CA).

Antibodies

Antibodies were used from the following sources: GFP (for immunofluorescence, in-house), GFP (for immunoblotting, kind gift from Geert Kops), TBP (20C7 in-house hybridoma (329)), α-Tubulin (DM1A, CP06, Calbiochem).

Nuclear extract preparation and GFP affinity purification

Cytoplasmic and nuclear extract preparation, protein concentration determination, GFP affinity purification and peptide elution were performed as described before (268). In short: Cells were grown and extracts were collected. Protein concentration was determined using Bradford assay. Extracts were bound to the GFP affinity beads on a rotating wheel at 4°C using 1 mg input protein. Beads were washed and peptides were eluted and trypsin digested. Tryptic digests were desalted using Stage-tips (330).

Mass spectrometry

Mass spectrometry analysis was performed as previously described (345). In short: Tryptic peptides were separated using an EASY-nLC 1000 (Thermo Fischer Scientific, Waltham, MA) mounted with a C18 column operating online with an LTQ Orbitrap Velos. Raw data was processed using MaxQuant version 1.5.2.8, Perseus version 1.5.4.0 and R version 3.2.2 (331) with Rstudio version 0.99.486.

Protein Stoichiometry determination

Stoichiometry was determined for listed interactors as described (353). In short: mass spectrometry intensities were corrected for the theoretical number of observable peptides resulting in an iBAQ value. Background was determined from the control immunoprecipitation and subtracted. Next, the values were scaled to a protein complex member as indicated.

Acknowledgements

We thank Proteins at Work (www.netherlandsproteomicscenter.nl/paw) for providing usage and support of their mass spectrometry facility. The authors would like to thank all the members of the Timmers group for their helpful discussions and suggestions.

A

Chapter 5

General discussion

Cellular regulation of molecular processes is vital for proper functioning of the cell. Multiple layers of regulation exist in the cell, which form an interplaying network of control to constrain and direct these processes. Misregulation can lead to serious cellular defects resulting in diseased states, or uncontrolled cell division, which can eventually become fatal for the organism. In this thesis multiple mechanisms of regulation have been studied. Chapter 2 describes development of Chromatin Regulation Ontology SiRNA Screening (CROSS) for identification of histone modifying enzymes and cofactors that regulate histone modification states. In chapter 3 the nuclear importer IPO5 is identified as a *de novo* interacting protein of BMP R-SMADs, regulating the spatiotemporal control of these transcription factors. Chapter 4 describes the transcription dependence and physical interaction of the MLL4 histone methyltransferase complex to SMAD target genes resulting in SMAD-mediated transcriptional activation. In this discussion chapter these results will be put into perspective of the existing knowledge of transcription and chromatin regulation.

Expanding the possibilities of CROSS

The various histone modifications and histone modifying enzyme complexes provide an extensive molecular toolbox for the cell. Whether histone modifications are cause or consequences of transcription is still unknown (354). The histone modifications on various residues of nucleosomes allow for association of transcription factors with chromatin like TFIID and SAGA binding to H3K4me3 and p300 binding to acetylated lysines. *Vice versa* active transcription by RNA pol II promotes the recruitment of additional histone modifying enzyme complexes like Set1, resulting in H3K4 methylation (355).

In Chapter 2 we developed a microscopy-based method to investigate cross-talk between different histone modifications. The CROSS method relies on antibody binding for detection of post-translational modifications (PTMs) on the tails of histones. These histone tails provide proper access of the epitope to the PTM specific antibody. As proof of principle, CROSS was applied to H3K4 and H3K27 methylation and identified known writers, erasers and cofactors for these PTMs. Reduction of H3K4 methylation was observed for the shared core subunits, WDR5, RBBP5, ASH2L and DPY30 (WRAD) of the SET/MLL methyltransferase complex but not with the catalytic enzymes of this complex, supporting their functional redundancy. Conversely, the H3K4 demethylases KDM5A and KDM5C show an increase in H3K4me3. CROSS for H3K27 methylation identified the core subunits of the PRC2 complex as writers and the H3K27 demethylase UTX. Additional information on genes affecting fitness was obtained by quantifying cell growth. A clear subset of genes affected cell fitness in MCF7 but not U2OS and *vice versa*. An example is the reduced MCF7 fitness on BRD2 knockdown. Cell viability assays confirm MCF7 to be more sensitive towards the BRD inhibitor I-BET151.

The globular domain of histones also provides amino acid residues that are present at the surface which can become modified by histone modifying enzymes like H2BK120 ubiquitination (136) and H3K79 methylation (356). Methylation of H3K79 is catalyzed by the mammalian DOT1L (62,357-359). However, the demethylase reversing this H3K79 methylation is still unknown. CROSS was used to pursue the identity of the KDM responsible for H3K79 demethylation, but did not succeed in identifying the catalytic enzyme (data not shown). The siRNA library used for screening with CROSS is composed of 529 writers, erasers, readers and cofactors of known chromatin modifying complexes over seven 96-well plates. Possibly, the siRNA library did not contain the target enzyme for the H3K79

demethylase because the enzyme has not been linked with histone modifying activity before. By screening against a genome-wide library the enzyme may be identified. However, currently a genome wide screen on 96-well format requires a lot of resources to be performed as it requires an approximate 40-fold increase of the amount of library plates, resulting in approximately 280 plates per screen. In order for a genome-wide screen to be technically feasible, CROSS needs to be scaled to 384- or 1536-well plates. Additionally, the H3K79 epitope was difficult to be accessed by our antibodies and harsh conditions were required to allow our antibody to detect the modification site. Exposure of the epitope for the antibody required boiling of the plates, resulting in edge-specific plate effects, *i.e.* wells at the edges and corners of the plate showed significant different fluorescence intensities compared to wells in the center of the plate, resulting in false hit identification. These technical limitations can be expected at more modification sites in the globular domain of histones and pose a potential limit for CROSS. Proper experimental conditions could be determined to access these sites to study globular histone PTMs. CROSS requires specific recognition of antibodies to detect global histone PTM levels. Traditional antibodies obtained from mice or rabbits consist of two heavy and two light chains that form the classical Y-shaped protein. These proteins are relatively large (150 – 160 kDa) and could be sterically hindered in recognition of the epitope. The identification of camelid single chain antibodies allows generation site specific nanobodies that are smaller (12 – 15 kDa) than the conventionally used antibodies (360). These nanobodies are directly compatible with CROSS and potentially make new tools for identifying and detecting these epitopes.

Histone modifying enzymes catalyze the addition or removal of small (*e.g.* acetylation, methylation and phosphorylation) or large (*e.g.* ubiquitination and other ubiquitin-like) chemical modifications to various amino acid residues. However, the enzyme activities of these proteins are not limited to histone proteins (361). For example, ubiquitination is a well-established PTM that is present on many different targets acting in various cellular processes (362,363) and alternate links for histone modifying enzymes to different non-histone target might exist. CROSS is not limited to antibody-mediated staining of histones. Rather, it depends on the availability of proper antibodies recognizing site-specific modifications on any protein. These modifications can be visualized at global levels. Thus, CROSS could serve as a platform for detecting their regulators. The algorithm used for quantification of the antibody signal is currently dependent on the staining and automated identification of the cell nucleus via DAPI staining. This limits CROSS for quantification of signals inside the cell nucleus. By selectively staining the cell cytoplasm and modifying the recognition algorithm, proteins outside the nucleus could be automatically quantified using CROSS.

Development of new techniques allows the study of biological processes from novel perspectives. CROSS is a screening platform that uses siRNA to mediate knockdown in eukaryotic cell lines. The disadvantage of siRNA-mediated knockdown includes off-target effects and possible incomplete gene knockdown. A low knockdown efficiency results in residual protein that can partially rescue the turnover of an associated histone PTM in comparison to a full knockdown of the siRNA-targeted gene. Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)/Cas9 mediated gene knockout allows for targeted knockout of endogenous genes in mammalian cells and is rapidly adopted by the scientific community to study gene knockout effects. CROSS is a modular technique that allows robot-assisted automated staining of fixed cell lines and could be adapted to integrate the CRISPR/Cas9 mediated knockout of genes. A cell line could be transfected with a

CRISPR/Cas9 library to generate these knockout cell lines after which antibody staining and automated detection could be performed. The disadvantage of this approach lies in lethality as a result of the removal of essential genes (364). Gene knockout by CRISPR/Cas9 is binary while siRNA-mediated knockdown can result in low levels of the targeted gene product, which may not compromise cellular survival.

SMAD interactors

TGF-β signaling is dependent on SMAD trimer complex formation and association with other nuclear transcription factors to navigate to their targets. This makes SMAD-mediated TGF-β signaling a very dynamic system that changes its interaction partners depending on time and activation state of this signaling pathway. In addition, members of the SMAD family mediate intercellular signaling via BMP (Bone Morphogenetic Protein) ligands. We applied a quantitative proteomics-based interaction approach in chapter 3 to identify SMAD interacting proteins. Using this approach we identified known interactors like SKI, SKIL and LEMD3, but not all known SMAD interacting proteins were present in our datasets. The TGF-β or BMP type I or II receptors were not identified for any SMAD and this possibly relates to the poor extractability of membrane-bound proteins (365). Absence of transcription factors like MYOD1 or OCT4 might be explained by the lack of expression of these proteins in the used cell type (203). HeLa cells were used for generation of cytoplasmic and nuclear extracts and might not express these interacting transcription factors.

Absence of known SMAD-interacting proteins might be due to the low affinity and transient nature of the interaction. These proteins will be removed during the washing steps in the affinity purification protocol. To increase the amount of identified SMAD interactors we applied a crosslinking step on the extracts after bead binding to covalently bind these low affinity and transiently binding proteins. The used dithiobis(succinimidyl propionate) (DSP) crosslinker contains two amine-reactive N-hydroxysuccinimide (NHS) esters on a spacer arm that can covalently bind primary amines found at N-termini and lysine residues (366). This can result in the covalent attachment of two nearby proteins, which will remain bound during the affinity purification washing steps. In the middle of the spacer a disulfide bond is present that is cleaved upon addition of dithiothreitol (DTT). This allows for separation of the proteins that subsequently can be identified using mass spectrometry. Identification of known SMAD interactors in the volcano plot showed a small enrichment, whereby proteins moved towards the upper right quadrant representing the proteins that interact with the bait (data not shown). Unfortunately these proteins remained in the non-significant background cloud. Additionally, our crosslinking approach resulted in an increased background signal and did not yield additional interactors. Possibly, the absence of DTT in the crosslinked sample allows for intramolecular protein disulfide bonds to be formed. Affinity purification experiments in the absence of the DSP crosslinker contain low levels of DTT that prevents this background binding.

Cytoplasmic and nuclear extracts represent the available protein pool in solution in these two compartments. A nuclear pellet remains after protein extraction containing proteins that are tightly bound to the chromatin. The proteins that are present in the nuclear pellet represent the proteins that are not released from chromatin by increased ionic strength and may play a part of active nuclear processes like transcription, repression and nucleosome remodeling. Presence of DNA might play a role in the association of proteins and certain interactions might only be observed in protein extracts from the nuclear pellet.

These different associations are indicative of different regulation. Therefore, it is worth testing whether the SMADs obtained from the nuclear pellet interact with proteins that are not observed in the nuclear extract like the associated coactivator complexes MLL4 and p300 (209). Additionally, synthetic DNA containing SMAD binding elements might be added to stimulate complex formation forming to prevent loss of interaction during affinity purification washes.

IPO7 and IPO8 were shown to mediate nuclear localization for both BMP and TGF-β R-SMADs (179,180). Our quantitative proteomics approach identified IPO5 as a novel BMP R-SMAD transport protein. Interestingly, forced nuclear localization of R-SMADs by overexpression of IPO7 was more pronounced than overexpression of IPO5. This might be the reason why IPO5 was not identified before as a mediator of SMAD nuclear import. We found that RanBP1 and not RanBP2 acts on SMAD transport. Only knockdown of RanBP2 but not RanBP1 affected nuclear localization of SMAD1 and SMAD3. This corroborates the observation that RanBP1 alone is sufficient to stimulate IPO5, but not importin-β GTPase activity (166). This suggests that SMAD import might be decoupled from SMAD export and this process might be regulated. Localization experiments involving comparison and mutation of the nuclear export sequence might shed light on this molecular control. Alternatively, small molecule inhibitors targeting importins and exportins might be used as a tool to further study SMAD localization control (367).

SMAD transcription regulation

Previously it was shown that SMAD2 interacts with the MLL3/4 subunit PTIP (76,344), but the link to the MLL3/MLL4 methyltransferase complexes was not studied before. The transcriptional effects on SMAD target genes upon knockdown of MLL3/MLL4 subunits using siRNA suggests recruitment of MLL3/MLL4 to these genes resulting in activation of transcription. Co-localization experiments using chromatin immunoprecipitation followed by sequencing (ChIP-seq) could show that both the TGF-β R-SMADs and MLL3/4 and their subunits are present at these target genes. Follow-up of ChIP-seq for H3K4 and H3K27 methylation is then expected to show reduced levels of H3K27 methylation and increased levels of H3K4 methylation at these binding sites due to the recruitment of the MLL3/4 complex. Furthermore, these experiments would indicate whether all TGF-β target genes are occupied by MLL4 or only a subset contains these coactivator complexes and furthermore identify regulatory sites involved in MLL4 binding.

BMP and TGF-β SMAD signaling pathways resemble each other in many different aspects in terms of molecular regulation, yet differences can be identified between the two pathways. For example, both pathways make use of IPO7-mediated protein transport while IPO5 only mediates BMP R-SMADs. SMADs recruit p300/CBP as coactivator for both TGF-β and BMP signaling (209). The association of MLL4 with TGF-β but not BMP R-SMADs represents another difference between signaling pathways. Interestingly, SMAD2/3 binding sites are enriched at active enhancers that are occupied by MLL4 and the MLL4 protein recruits p300 to these sites (347). These observations suggest a synergistic effect between these elements on target gene binding by TGF-β R-SMADs, yet raises the question of whether a similar recruitment mechanism is present in BMP R-SMAD signaling. Careful dissection of the TGF-β and BMP signaling pathways and their interaction partners may allow identification of various druggable cofactors. SB-431542 and LDN-193189 block TGF-β and BMP signaling respectively, however are not used in the clinic (368,369). Alternatively,

LY2157299 is used in the clinic as a general TGF-β signaling drug that downregulates SMAD2 phosphorylation (370). Specific drug combinations could lead to selective inhibition of specific TGF-β or BMP targets despite the generic nonspecific nature of SMAD signaling.

KDM6A as part of multiple protein complexes

Affinity purification of tagged recombinant proteins allows for identification of multiprotein complexes, suggesting that these proteins are folded and incorporated properly without interference of the tag. Some proteins, like WDR5 and KDM6A, are shared between different complexes (76). When these shared proteins are used as bait, a mixture of the subunits of the different complexes is identified. These type of experiments can identify interaction partners of a given bait. However they do not distinguish between an isolated binding event and assembly of a multi subunit complex. Furthermore, affinity purification of full-length proteins cannot identify the specific binding domains by which the proteins interact. To study the binding location of the tagged protein, each domain needs to be individually cloned and analyzed. KDM6A was identified to interact with MLL3 and MLL4, PR-DUB and MiDAC. MLL3/MLL4 and PR-DUB complexes both stimulate a chromatin environment that is transcriptionally permissive. This could be the result of the sum of individual activity of both complexes whereby the MLL3 MLL4 complexes mono-methylates H3K4 and PR-DUB removes H2AK119ub1 and H3K27me3. Alternatively, these protein complexes act synergistically by recruiting each other via the interaction with KDM6A. This is strengthened by the observation that MLL3 is observed when using ASXL2 as bait (143).

In contrast, MiDAC is a deacetylase complex that results in transcriptionally repressive chromatin by removing acetyl groups from histones. Interaction with KDM6A might be favored towards a certain complex during defined cell stages. Cell synchronization experiments could be performed to study this dependency. The site of interaction might overlap with sites required for MLL3/4 or PR-DUB binding, resulting in a mutually exclusive interaction. Furthermore, the biological function of KDM6A in MiDAC is not known and ChIP-seq experiments in synchronized mitotic cells could elucidate genomic targets of this complex. Alternatively, recruitment of KDM6A to the MiDAC complex could alter the activity of the complex instead of regulating a specific subset of targets.

To conclude

In this thesis we investigated the molecular interactions of machinery of cellular control on various levels. Enzymatic activity of proteins influence the PTM state of histones affecting gene transcription while physical binding of interactors result in modulation of activity and spatiotemporal distribution in cells. The combinatorial effect of these effects determines protein state and molecular outcome. A common property for these factors is that their deregulation disrupts cellular homeostasis. This disruption can lead to disease and many of the involved proteins are misregulated in cancer. Specific small molecule inhibitors targeting the proteins involved in this regulation like TGF-β inhibitors or epidrugs are being developed and tested for clinical use (370-373). The combinatorial nature of cellular regulation underscores the potential of combinatorial drug treatment. Therefore, it is required to precisely study and understand these regulation methods. The findings of this thesis contribute to understanding of these molecular control processes and provide directions for future research.

5

Appendices &

References

1. Lagrange, T., Kapanidis, A. N., Tang, H., Reinberg, D., and Ebright, R. H. (1998) New core promoter element in RNA polymerase II-dependent transcription: sequence-specific DNA binding by transcription factor IIB. Genes Dev 12, 34-44

2. Saxonov, S., Berg, P., and Brutlag, D. L. (2006) A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. Proc Natl Acad Sci U S A 103, 1412-1417

3. Burke, T. W., and Kadonaga, J. T. (1996) Drosophila TFIID binds to a conserved downstream basal promoter element that is present in many TATA-box-deficient promoters. Genes Dev 10, 711-724

4. Xi, H., Yu, Y., Fu, Y., Foley, J., Halees, A., and Weng, Z. (2007) Analysis of overrepresented motifs in human core promoters reveals dual regulatory roles of YY1. Genome Res 17, 798-806

5. Lim, C. Y., Santoso, B., Boulay, T., Dong, E., Ohler, U., and Kadonaga, J. T. (2004) The MTE, a new core promoter element for transcription by RNA polymerase II. Genes Dev 18, 1606-1617

6. Buratowski, S., Hahn, S., Guarente, L., and Sharp, P. A. (1989) Five intermediate complexes in transcription initiation by RNA polymerase II. Cell 56, 549-561

7. Buratowski, S., Hahn, S., Sharp, P. A., and Guarente, L. (1988) Function of a yeast TATA element-binding protein in a mammalian transcription system. Nature 334, 37-42

8. Hahn, S., Buratowski, S., Sharp, P. A., and Guarente, L. (1989) Yeast TATA-binding protein TFIID binds to TATA elements with both consensus and nonconsensus DNA sequences. Proc Natl Acad Sci U S A 86, 5718-5722

9. Bieniossek, C., Papai, G., Schaffitzel, C., Garzoni, F., Chaillet, M., Scheer, E., Papadopoulos, P., Tora, L., Schultz, P., and Berger, I. (2013) The architecture of human general transcription factor TFIID core complex. Nature 493, 699-702

10. Louder, R. K., He, Y., Lopez-Blanco, J. R., Fang, J., Chacon, P., and Nogales, E. (2016) Structure of promoter-bound TFIID and model of human pre-initiation complex assembly. Nature 531, 604-609

11. Lu, H., Flores, O., Weinmann, R., and Reinberg, D. (1991) The nonphosphorylated form of RNA polymerase II preferentially associates with the preinitiation complex. Proc Natl Acad Sci U S A 88, 10004-10008

12. Fischer, L., Gerard, M., Chalut, C., Lutz, Y., Humbert, S., Kanno, M., Chambon, P., and Egly, J. M. (1992) Cloning of the 62-kilodalton component of basic transcription factor BTF2. Science 257, 1392-1395

13. Schaeffer, L., Roy, R., Humbert, S., Moncollin, V., Vermeulen, W., Hoeijmakers, J. H., Chambon, P., and Egly, J. M. (1993) DNA repair helicase: a component of BTF2 (TFIIH) basic transcription factor. Science 260, 58-63

14. Wang, W., Carey, M., and Gralla, J. D. (1992) Polymerase II promoter activation: closed complex formation and ATP-driven start site opening. Science 255, 450-453

15. Holstege, F. C., Fiedler, U., and Timmers, H. T. (1997) Three transitions in the RNA polymerase II transcription complex during initiation. EMBO J 16, 7468-7480

16. Watson, J. D., and Crick, F. H. (1953) The structure of DNA. Cold Spring Harb Symp Quant Biol 18, 123-131

17. Fawcett, D. W. (1966) An atlas of fine structure: the cell, its organelles, and inclusions, W. B. Saunders Co., Philadelphia,

18. Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F., and Richmond, T. J. (1997) Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature 389, 251-260

19. Strahl, B. D., and Allis, C. D. (2000) The language of covalent histone modifications. Nature 403, 41-45

20. Kouzarides, T. (2007) Chromatin modifications and their function. Cell 128, 693-705

21. Sterner, D. E., and Berger, S. L. (2000) Acetylation of histones and transcription-related factors. Microbiol Mol Biol Rev 64, 435-459

22. Zhang, Y., and Reinberg, D. (2001) Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. Genes Dev 15, 2343-2360

23. Nowak, S. J., and Corces, V. G. (2004) Phosphorylation of histone H3: a balancing act between chromosome condensation and transcriptional activation. Trends Genet 20, 214-220

24. Dawson, M. A., Bannister, A. J., Gottgens, B., Foster, S. D., Bartke, T., Green, A. R., and Kouzarides, T. (2009) JAK2 phosphorylates histone H3Y41 and excludes HP1alpha from chromatin. Nature 461, 819-822

25. Kurihara, D., Matsunaga, S., Omura, T., Higashiyama, T., and Fukui, K. (2011) Identification and characterization of plant Haspin kinase as a histone H3 threonine kinase. BMC Plant Biol 11, 73

26. Shilatifard, A. (2006) Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression. Annu Rev Biochem 75, 243-269

27. Nathan, D., Ingvarsdottir, K., Sterner, D. E., Bylebyl, G. R., Dokmanovic, M., Dorsey, J. A., Whelan, K. A., Krsmanovic, M., Lane, W. S., Meluh, P. B., Johnson, E. S., and Berger, S. L. (2006) Histone sumoylation is a negative regulator in Saccharomyces cerevisiae and shows dynamic interplay with positive-acting histone modifications. Genes Dev 20, 966-976

28. Li, T., Guan, J., Huang, Z., Hu, X., and Zheng, X. (2014) RNF168-mediated H2A neddylation antagonizes ubiquitylation of H2A and regulates DNA damage repair. J Cell Sci 127, 2238-2248

29. Creyghton, M. P., Cheng, A. W., Welstead, G. G., Kooistra, T., Carey, B. W., Steine, E. J., Hanna, J., Lodato, M. A., Frampton, G. M., Sharp, P. A., Boyer, L. A., Young, R. A., and Jaenisch, R. (2010) Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc Natl Acad Sci U S A 107, 21931-21936

30. Jacobson, R. H., Ladurner, A. G., King, D. S., and Tjian, R. (2000) Structure and function of a human TAFII250 double bromodomain module. Science 288, 1422-1425

31. Tie, F., Banerjee, R., Stratton, C. A., Prasad-Sinha, J., Stepanik, V., Zlobin, A., Diaz, M. O., Scacheri, P. C., and Harte, P. J. (2009) CBP-mediated acetylation of histone H3 lysine 27 antagonizes Drosophila Polycomb silencing. Development 136, 3131-3141

32. Goodman, R. H., and Smolik, S. (2000) CBP/p300 in cell growth, transformation, and development. Genes Dev 14, 1553-1577

33. Lee, K. K., Sardiu, M. E., Swanson, S. K., Gilmore, J. M., Torok, M., Grant, P. A., Florens, L., Workman, J. L., and Washburn, M. P. (2011) Combinatorial depletion analysis to assemble the network architecture of the SAGA and ADA chromatin remodeling complexes. Mol Syst Biol 7, 503

34. Wu, P. Y., Ruhlmann, C., Winston, F., and Schultz, P. (2004) Molecular architecture of the S. cerevisiae SAGA complex. Mol Cell 15, 199-208

35. Henry, K. W., Wyce, A., Lo, W. S., Duggan, L. J., Emre, N. C., Kao, C. F., Pillus, L., Shilatifard, A., Osley, M. A., and Berger, S. L. (2003) Transcriptional activation via sequential histone H2B ubiquitylation and deubiquitylation, mediated by SAGA-associated Ubp8. Genes Dev 17, 2648-2663

36. Grant, P. A., Duggan, L., Cote, J., Roberts, S. M., Brownell, J. E., Candau, R., Ohba, R., Owen-Hughes, T., Allis, C. D., Winston, F., Berger, S. L., and Workman, J. L. (1997) Yeast Gcn5 functions in two multisubunit complexes to acetylate nucleosomal histones: characterization of an Ada complex and the SAGA (Spt/Ada) complex. Genes Dev 11, 1640-1650

37. Balasubramanian, R., Pray-Grant, M. G., Selleck, W., Grant, P. A., and Tan, S. (2002) Role of the Ada2 and Ada3 transcriptional coactivators in histone acetylation. J Biol Chem 277, 7989-7995

38. Lee, T. I., Causton, H. C., Holstege, F. C., Shen, W. C., Hannett, N., Jennings, E. G., Winston, F., Green, M. R., and Young, R. A. (2000) Redundant roles for the TFIID and SAGA complexes in global transcription. Nature 405, 701-704

39. Belotserkovskaya, R., Sterner, D. E., Deng, M., Sayre, M. H., Lieberman, P. M., and Berger, S. L. (2000) Inhibition of TATA-binding protein function by SAGA subunits Spt3 and Spt8 at Gcn4-activated promoters. Mol Cell Biol 20, 634-647

40. Huisinga, K. L., and Pugh, B. F. (2004) A genome-wide housekeeping role for TFIID and a highly regulated stress-related role for SAGA in Saccharomyces cerevisiae. Mol Cell 13, 573-585

41. Krebs, A. R., Karmodiya, K., Lindahl-Allen, M., Struhl, K., and Tora, L. (2011) SAGA and ATAC histone acetyl transferase complexes regulate distinct sets of genes and ATAC defines a class of p300-independent enhancers. Mol Cell 44, 410-423

42. Xue, Y., Wong, J., Moreno, G. T., Young, M. K., Cote, J., and Wang, W. (1998) NURD, a novel complex with both ATP-dependent chromatin-remodeling and histone deacetylase activities. Mol Cell 2, 851-861

43. Laherty, C. D., Yang, W. M., Sun, J. M., Davie, J. R., Seto, E., and Eisenman, R. N. (1997) Histone deacetylases associated with the mSin3 corepressor mediate mad transcriptional repression. Cell 89, 349-356

44. You, A., Tong, J. K., Grozinger, C. M., and Schreiber, S. L. (2001) CoREST is an integral component of the CoREST- human histone deacetylase complex. Proc Natl Acad Sci U S A 98, 1454-1458

45. Humphrey, G. W., Wang, Y., Russanova, V. R., Hirai, T., Qin, J., Nakatani, Y., and Howard, B. H. (2001) Stable histone deacetylase complexes distinguished by the presence of SANT domain proteins CoREST/kiaa0071 and Mta-L1. J Biol Chem 276, 6817-6824

46. Bantscheff, M., Hopf, C., Savitski, M. M., Dittmann, A., Grandi, P., Michon, A. M., Schlegl, J., Abraham, Y.,

Becher, I., Bergamini, G., Boesche, M., Delling, M., Dumpelfeld, B., Eberhard, D., Huthmacher, C., Mathieson, T., Poeckel, D., Reader, V., Strunk, K., Sweetman, G., Kruse, U., Neubauer, G., Ramsden, N. G., and Drewes, G. (2011) Chemoproteomics profiling of HDAC inhibitors reveals selective targeting of HDAC complexes. Nat Biotechnol 29, 255-265

47. Choi, E., Han, C., Park, I., Lee, B., Jin, S., Choi, H., Kim, D. H., Park, Z. Y., Eddy, E. M., and Cho, C. (2008) A novel germ cell-specific protein, SHIP1, forms a complex with chromatin remodeling activity during spermatogenesis. J Biol Chem 283, 35283-35294

48. Hao, Y., Xu, N., Box, A. C., Schaefer, L., Kannan, K., Zhang, Y., Florens, L., Seidel, C., Washburn, M. P., Wiegraebe, W., and Mak, H. Y. (2011) Nuclear cGMP-dependent kinase regulates gene expression via activity-dependent recruitment of a conserved histone deacetylase complex. PLoS Genet 7, e1002065

49. Cantoni, G. L. (1952) The nature of the active methyl donor formed enzymatically from L-methionine and adenosinetriphosphate. Journal of the American Chemical Society 74, 2942-2943

50. Bedford, M. T. (2007) Arginine methylation at a glance. J Cell Sci 120, 4243-4246

51. Molina-Serrano, D., Schiza, V., and Kirmizis, A. (2013) Cross-talk among epigenetic modifications: lessons from histone arginine methylation. Biochem Soc Trans 41, 751-759

52. Kirmizis, A., Santos-Rosa, H., Penkett, C. J., Singer, M. A., Green, R. D., and Kouzarides, T. (2009) Distinct transcriptional outputs associated with mono- and dimethylated histone H3 arginine 2. Nat Struct Mol Biol 16, 449-451

53. Migliori, V., Muller, J., Phalke, S., Low, D., Bezzi, M., Mok, W. C., Sahu, S. K., Gunaratne, J., Capasso, P., Bassi, C., Cecatiello, V., De Marco, A., Blackstock, W., Kuznetsov, V., Amati, B., Mapelli, M., and Guccione, E. (2012) Symmetric dimethylation of H3R2 is a newly identified histone mark that supports euchromatin maintenance. Nat Struct Mol Biol 19, 136-144

54. Yuan, C. C., Matthews, A. G., Jin, Y., Chen, C. F., Chapman, B. A., Ohsumi, T. K., Glass, K. C., Kutateladze, T. G., Borowsky, M. L., Struhl, K., and Oettinger, M. A. (2012) Histone H3R2 symmetric dimethylation and histone H3K4 trimethylation are tightly correlated in eukaryotic genomes. Cell Rep 1, 83-90

55. Guccione, E., Bassi, C., Casadio, F., Martinato, F., Cesaroni, M., Schuchlautz, H., Luscher, B., and Amati, B. (2007) Methylation of histone H3R2 by PRMT6 and H3K4 by an MLL complex are mutually exclusive. Nature 449, 933-937

56. Hyllus, D., Stein, C., Schnabel, K., Schiltz, E., Imhof, A., Dou, Y., Hsieh, J., and Bauer, U. M. (2007) PRMT6-mediated methylation of R2 in histone H3 antagonizes H3 K4 trimethylation. Genes Dev 21, 3369-3380

57. Kirmizis, A., Santos-Rosa, H., Penkett, C. J., Singer, M. A., Vermeulen, M., Mann, M., Bahler, J., Green, R. D., and Kouzarides, T. (2007) Arginine methylation at histone H3R2 controls deposition of H3K4 trimethylation. Nature 449, 928-932

58. Cuthbert, G. L., Daujat, S., Snowden, A. W., Erdjument-Bromage, H., Hagiwara, T., Yamada, M., Schneider, R., Gregory, P. D., Tempst, P., Bannister, A. J., and Kouzarides, T. (2004) Histone deimination antagonizes arginine methylation. Cell 118, 545-553

59. Wang, Y., Wysocka, J., Sayegh, J., Lee, Y. H., Perlin, J. R., Leonelli, L., Sonbuchner, L. S., McDonald, C. H., Cook, R. G., Dou, Y., Roeder, R. G., Clarke, S., Stallcup, M. R., Allis, C. D., and Coonrod, S. A. (2004) Human PAD4 regulates histone arginine methylation levels via demethylimination. Science 306, 279-283

60. Qian, C., and Zhou, M. M. (2006) SET domain protein lysine methyltransferases: Structure, specificity and catalysis. Cell Mol Life Sci 63, 2755-2763

61. Min, J., Feng, Q., Li, Z., Zhang, Y., and Xu, R. M. (2003) Structure of the catalytic domain of human DOT1L, a non-SET domain nucleosomal histone methyltransferase. Cell 112, 711-723

62. Feng, Q., Wang, H., Ng, H. H., Erdjument-Bromage, H., Tempst, P., Struhl, K., and Zhang, Y. (2002) Methylation of H3-lysine 79 is mediated by a new family of HMTases without a SET domain. Curr Biol 12, 1052-1058

63. Wagner, E. J., and Carpenter, P. B. (2012) Understanding the language of Lys36 methylation at histone H3. Nat Rev Mol Cell Biol 13, 115-126

64. Ng, H. H., Ciccone, D. N., Morshead, K. B., Oettinger, M. A., and Struhl, K. (2003) Lysine-79 of histone H3 is hypomethylated at silenced loci in yeast and mammalian cells: a potential mechanism for position-effect variegation. Proc Natl Acad Sci U S A 100, 1820-1825

65. Wang, Z., Zang, C., Rosenfeld, J. A., Schones, D. E., Barski, A., Cuddapah, S., Cui, K., Roh, T. Y., Peng, W., Zhang, M. Q., and Zhao, K. (2008) Combinatorial patterns of histone acetylations and methylations in the human

genome. Nat Genet 40, 897-903

66. Huyen, Y., Zgheib, O., Ditullio, R. A., Jr., Gorgoulis, V. G., Zacharatos, P., Petty, T. J., Sheston, E. A., Mellert, H. S., Stavridi, E. S., and Halazonetis, T. D. (2004) Methylated lysine 79 of histone H3 targets 53BP1 to DNA double-strand breaks. Nature 432, 406-411

67. Vermeulen, M., Eberl, H. C., Matarese, F., Marks, H., Denissov, S., Butter, F., Lee, K. K., Olsen, J. V., Hyman, A. A., Stunnenberg, H. G., and Mann, M. (2010) Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. Cell 142, 967-980

68. Vermeulen, M., Mulder, K. W., Denissov, S., Pijnappel, W. W., van Schaik, F. M., Varier, R. A., Baltissen, M. P., Stunnenberg, H. G., Mann, M., and Timmers, H. T. (2007) Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. Cell 131, 58-69

69. Yun, M., Wu, J., Workman, J. L., and Li, B. (2011) Readers of histone modifications. Cell Res 21, 564-578

70. Huang, Y., Fang, J., Bedford, M. T., Zhang, Y., and Xu, R. M. (2006) Recognition of histone H3 lysine-4 methylation by the double tudor domain of JMJD2A. Science 312, 748-751

71. Varier, R. A., Outchkourov, N. S., de Graaf, P., van Schaik, F. M., Ensing, H. J., Wang, F., Higgins, J. M., Kops, G. J., and Timmers, H. T. (2010) A phospho/methyl switch at histone H3 regulates TFIID association with mitotic chromosomes. EMBO J 29, 3967-3978

72. Wu, M., Wang, P. F., Lee, J. S., Martin-Brown, S., Florens, L., Washburn, M., and Shilatifard, A. (2008) Molecular regulation of H3K4 trimethylation by Wdr82, a component of human Set1/COMPASS. Mol Cell Biol 28, 7337-7344

73. Wang, P., Lin, C., Smith, E. R., Guo, H., Sanderson, B. W., Wu, M., Gogol, M., Alexander, T., Seidel, C., Wiedemann, L. M., Ge, K., Krumlauf, R., and Shilatifard, A. (2009) Global analysis of H3K4 methylation defines MLL family member targets and points to a role for MLL1-mediated H3K4 methylation in the regulation of transcriptional initiation by RNA polymerase II. Mol Cell Biol 29, 6074-6085

74. Herz, H. M., Mohan, M., Garruss, A. S., Liang, K., Takahashi, Y. H., Mickey, K., Voets, O., Verrijzer, C. P., and Shilatifard, A. (2012) Enhancer-associated H3K4 monomethylation by Trithorax-related, the Drosophila homolog of mammalian Mll3/Mll4. Genes Dev 26, 2604-2620

75. Hu, D., Gao, X., Morgan, M. A., Herz, H. M., Smith, E. R., and Shilatifard, A. (2013) The MLL3/MLL4 branches of the COMPASS family function as major histone H3K4 monomethylases at enhancers. Mol Cell Biol 33, 4745-4754

76. van Nuland, R., Smits, A. H., Pallaki, P., Jansen, P. W., Vermeulen, M., and Timmers, H. T. (2013) Quantitative dissection and stoichiometry determination of the human SET1/MLL histone methyltransferase complexes. Mol Cell Biol 33, 2067-2077

77. Cho, Y. W., Hong, T., Hong, S., Guo, H., Yu, H., Kim, D., Guszczynski, T., Dressler, G. R., Copeland, T. D., Kalkum, M., and Ge, K. (2007) PTIP associates with MLL3- and MLL4-containing histone H3 lysine 4 methyltransferase complex. J Biol Chem 282, 20395-20406

78. Gu, Y., Nakamura, T., Alder, H., Prasad, R., Canaani, O., Cimino, G., Croce, C. M., and Canaani, E. (1992) The t(4;11) chromosome translocation of human acute leukemias fuses the ALL-1 gene, related to Drosophila trithorax, to the AF-4 gene. Cell 71, 701-708

79. Meyer, C., Kowarz, E., Hofmann, J., Renneville, A., Zuna, J., Trka, J., Ben Abdelali, R., Macintyre, E., De Braekeleer, E., De Braekeleer, M., Delabesse, E., de Oliveira, M. P., Cave, H., Clappier, E., van Dongen, J. J., Balgobind, B. V., van den Heuvel-Eibrink, M. M., Beverloo, H. B., Panzer-Grumayer, R., Teigler-Schlegel, A., Harbott, J., Kjeldsen, E., Schnittger, S., Koehl, U., Gruhn, B., Heidenreich, O., Chan, L. C., Yip, S. F., Krzywinski, M., Eckert, C., Moricke, A., Schrappe, M., Alonso, C. N., Schafer, B. W., Krauter, J., Lee, D. A., Zur Stadt, U., Te Kronnie, G., Sutton, R., Izraeli, S., Trakhtenbrot, L., Lo Nigro, L., Tsaur, G., Fechina, L., Szczepanski, T., Strehl, S., Ilencikova, D., Molkentin, M., Burmeister, T., Dingermann, T., Klingebiel, T., and Marschalek, R. (2009) New insights to the MLL recombinome of acute leukemias. Leukemia 23, 1490-1499

80. van der Linden, M. H., Valsecchi, M. G., De Lorenzo, P., Moricke, A., Janka, G., Leblanc, T. M., Felice, M., Biondi, A., Campbell, M., Hann, I., Rubnitz, J. E., Stary, J., Szczepanski, T., Vora, A., Ferster, A., Hovi, L., Silverman, L. B., and Pieters, R. (2009) Outcome of congenital acute lymphoblastic leukemia treated on the Interfant-99 protocol. Blood 114, 3764-3768

81. Tomizawa, D., Koh, K., Sato, T., Kinukawa, N., Morimoto, A., Isoyama, K., Kosaka, Y., Oda, T., Oda, M., Hayashi, Y., Eguchi, M., Horibe, K., Nakahata, T., Mizutani, S., and Ishii, E. (2007) Outcome of risk-based therapy for infant acute lymphoblastic leukemia with or without an MLL gene rearrangement, with emphasis on late effects: a final report of two consecutive studies, MLL96 and MLL98, of the Japan Infant Leukemia Study Group.

Leukemia 21, 2258-2263

82. Dordelmann, M., Reiter, A., Borkhardt, A., Ludwig, W. D., Gotz, N., Viehmann, S., Gadner, H., Riehm, H., and Schrappe, M. (1999) Prednisone response is the strongest predictor of treatment outcome in infant acute lymphoblastic leukemia. Blood 94, 1209-1217

83. Ferster, A., Benoit, Y., Francotte, N., Dresse, M. F., Uyttebroeck, A., Plouvier, E., Thyss, A., Lutz, P., Marguerite, G., Behar, C., Mazingue, F., Boutard, P., Millot, F., Rialland, X., Mechinaud, F., Norton, L., Robert, A., Otten, J., Vilmer, E., Philippe, N., Waterkeyn, C., and Suciu, S. (2000) Treatment outcome in infant acute lymphoblastic leukemia. Children Leukemia Cooperative Group--EORTC. European Organization for Research and Treatment of Cancer. Blood 95, 2729-2731

84. Reaman, G. H., Sposto, R., Sensel, M. G., Lange, B. J., Feusner, J. H., Heerema, N. A., Leonard, M., Holmes, E. J., Sather, H. N., Pendergrass, T. W., Johnstone, H. S., O'Brien, R. T., Steinherz, P. G., Zeltzer, P. M., Gaynon, P. S., Trigg, M. E., and Uckun, F. M. (1999) Treatment outcome and prognostic factors for infants with acute lymphoblastic leukemia treated on two consecutive trials of the Children's Cancer Group. J Clin Oncol 17, 445-455

85. Mohan, M., Herz, H. M., Takahashi, Y. H., Lin, C., Lai, K. C., Zhang, Y., Washburn, M. P., Florens, L., and Shilatifard, A. (2010) Linking H3K79 trimethylation to Wnt signaling through a novel Dot1-containing complex (DotCom). Genes Dev 24, 574-589

86. He, N., Liu, M., Hsu, J., Xue, Y., Chou, S., Burlingame, A., Krogan, N. J., Alber, T., and Zhou, Q. (2010) HIV-1 Tat and host AFF4 recruit two transcription elongation factors into a bifunctional complex for coordinated activation of HIV-1 transcription. Mol Cell 38, 428-438

87. Lin, C., Smith, E. R., Takahashi, H., Lai, K. C., Martin-Brown, S., Florens, L., Washburn, M. P., Conaway, J. W., Conaway, R. C., and Shilatifard, A. (2010) AFF4, a component of the ELL/P-TEFb elongation complex and a shared subunit of MLL chimeras, can link transcription elongation to leukemia. Mol Cell 37, 429-437

88. Sobhian, B., Laguette, N., Yatim, A., Nakamura, M., Levy, Y., Kiernan, R., and Benkirane, M. (2010) HIV-1 Tat assembles a multifunctional transcription elongation complex and stably associates with the 7SK snRNP. Mol Cell 38, 439-451

89. Varier, R. A., and Timmers, H. T. (2011) Histone lysine methylation and demethylation pathways in cancer. Biochim Biophys Acta 1815, 75-89

90. Knutson, B. A., Smith, M. L., Walker-Kopp, N., and Xu, X. (2016) Super elongation complex contains a TFIIF-related subcomplex. Transcription 7, 133-140

91. Saksouk, N., Simboeck, E., and Dejardin, J. (2015) Constitutive heterochromatin formation and transcription in mammals. Epigenetics Chromatin 8, 3

92. Tachibana, M., Sugimoto, K., Fukushima, T., and Shinkai, Y. (2001) Set domain-containing protein, G9a, is a novel lysine-preferring mammalian histone methyltransferase with hyperactivity and specific selectivity to lysines 9 and 27 of histone H3. J Biol Chem 276, 25309-25317

93. Tachibana, M., Sugimoto, K., Nozaki, M., Ueda, J., Ohta, T., Ohki, M., Fukuda, M., Takeda, N., Niida, H., Kato, H., and Shinkai, Y. (2002) G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. Genes Dev 16, 1779-1791

94. Tachibana, M., Ueda, J., Fukuda, M., Takeda, N., Ohta, T., Iwanari, H., Sakihama, T., Kodama, T., Hamakubo, T., and Shinkai, Y. (2005) Histone methyltransferases G9a and GLP form heteromeric complexes and are both crucial for methylation of euchromatin at H3-K9. Genes Dev 19, 815-826

95. Chin, H. G., Esteve, P. O., Pradhan, M., Benner, J., Patnaik, D., Carey, M. F., and Pradhan, S. (2007) Automethylation of G9a and its implication in wider substrate specificity and HP1 binding. Nucleic Acids Res 35, 7313-7323

96. Schultz, D. C., Ayyanathan, K., Negorev, D., Maul, G. G., and Rauscher, F. J., 3rd. (2002) SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. Genes Dev 16, 919-932

97. Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B. D., Sun, Z. W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C. P., Allis, C. D., and Jenuwein, T. (2000) Regulation of chromatin structure by site-specific histone H3 methyltransferases. Nature 406, 593-599

98. Nakayama, J., Rice, J. C., Strahl, B. D., Allis, C. D., and Grewal, S. I. (2001) Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. Science 292, 110-113

99. Bannister, A. J., Zegerman, P., Partridge, J. F., Miska, E. A., Thomas, J. O., Allshire, R. C., and Kouzarides, T. (2001) Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. Nature 410, 120-

124

100. Mishima, Y., Watanabe, M., Kawakami, T., Jayasinghe, C. D., Otani, J., Kikugawa, Y., Shirakawa, M., Kimura, H., Nishimura, O., Aimoto, S., Tajima, S., and Suetake, I. (2013) Hinge and chromoshadow of HP1alpha participate in recognition of K9 methylated histone H3 in nucleosomes. J Mol Biol 425, 54-70

101. Lachner, M., O'Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T. (2001) Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. Nature 410, 116-120

102. Canzio, D., Chang, E. Y., Shankar, S., Kuchenbecker, K. M., Simon, M. D., Madhani, H. D., Narlikar, G. J., and Al-Sady, B. (2011) Chromodomain-mediated oligomerization of HP1 suggests a nucleosome-bridging mechanism for heterochromatin assembly. Mol Cell 41, 67-81

103. Epsztejn-Litman, S., Feldman, N., Abu-Remaileh, M., Shufaro, Y., Gerson, A., Ueda, J., Deplus, R., Fuks, F., Shinkai, Y., Cedar, H., and Bergman, Y. (2008) De novo DNA methylation promoted by G9a prevents reprogramming of embryonically silenced genes. Nat Struct Mol Biol 15, 1176-1183

104. Lehnertz, B., Ueda, Y., Derijck, A. A., Braunschweig, U., Perez-Burgos, L., Kubicek, S., Chen, T., Li, E., Jenuwein, T., and Peters, A. H. (2003) Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. Curr Biol 13, 1192-1200

105. Hawkins, R. D., Hon, G. C., Lee, L. K., Ngo, Q., Lister, R., Pelizzola, M., Edsall, L. E., Kuan, S., Luu, Y., Klugman, S., Antosiewicz-Bourget, J., Ye, Z., Espinoza, C., Agarwahl, S., Shen, L., Ruotti, V., Wang, W., Stewart, R., Thomson, J. A., Ecker, J. R., and Ren, B. (2010) Distinct epigenomic landscapes of pluripotent and lineage-committed human cells. Cell Stem Cell 6, 479-491

106. Trojer, P., and Reinberg, D. (2007) Facultative heterochromatin: is there a distinctive molecular signature? Mol Cell 28, 1-13

107. Nekrasov, M., Wild, B., and Muller, J. (2005) Nucleosome binding and histone methyltransferase activity of Drosophila PRC2. EMBO Rep 6, 348-353

108. Ketel, C. S., Andersen, E. F., Vargas, M. L., Suh, J., Strome, S., and Simon, J. A. (2005) Subunit contributions to histone methyltransferase activities of fly and worm polycomb group complexes. Mol Cell Biol 25, 6857-6868

109. Hadorn, E. (1968) Transdetermination in cells. Sci Am 219, 110-114 passim

110. Shao, Z., Raible, F., Mollaaghababa, R., Guyon, J. R., Wu, C. T., Bender, W., and Kingston, R. E. (1999) Stabilization of chromatin structure by PRC1, a Polycomb complex. Cell 98, 37-46

111. Ma, R. G., Zhang, Y., Sun, T. T., and Cheng, B. (2014) Epigenetic regulation by polycomb group complexes: focus on roles of CBX proteins. J Zhejiang Univ Sci B 15, 412-428

112. Gao, Z., Zhang, J., Bonasio, R., Strino, F., Sawai, A., Parisi, F., Kluger, Y., and Reinberg, D. (2012) PCGF homologs, CBX proteins, and RYBP define functionally distinct PRC1 family complexes. Mol Cell 45, 344-356

113. Margueron, R., and Reinberg, D. (2011) The Polycomb complex PRC2 and its mark in life. Nature 469, 343-349

114. Hauri, S., Comoglio, F., Seimiya, M., Gerstung, M., Glatter, T., Hansen, K., Aebersold, R., Paro, R., Gstaiger, M., and Beisel, C. (2016) A High-Density Map for Navigating the Human Polycomb Complexome. Cell Rep 17, 583-595

115. Lee, M. G., Villa, R., Trojer, P., Norman, J., Yan, K. P., Reinberg, D., Di Croce, L., and Shiekhattar, R. (2007) Demethylation of H3K27 regulates polycomb recruitment and H2A ubiquitination. Science 318, 447-450

116. Ferrari, K. J., Scelfo, A., Jammula, S., Cuomo, A., Barozzi, I., Stutzer, A., Fischle, W., Bonaldi, T., and Pasini, D. (2014) Polycomb-dependent H3K27me1 and H3K27me2 regulate active transcription and enhancer fidelity. Mol Cell 53, 49-62

117. Barski, A., Cuddapah, S., Cui, K., Roh, T. Y., Schones, D. E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007) High-resolution profiling of histone methylations in the human genome. Cell 129, 823-837

118. Cui, K., Zang, C., Roh, T. Y., Schones, D. E., Childs, R. W., Peng, W., and Zhao, K. (2009) Chromatin signatures in multipotent human hematopoietic stem cells indicate the fate of bivalent genes during differentiation. Cell Stem Cell 4, 80-93

119. Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J. R., Cole, P. A., Casero, R. A., and Shi, Y. (2004) Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. Cell 119, 941-953

120. Tsukada, Y., Fang, J., Erdjument-Bromage, H., Warren, M. E., Borchers, C. H., Tempst, P., and Zhang, Y. (2006) Histone demethylation by a family of JmjC domain-containing proteins. Nature 439, 811-816

121. Metzger, E., Wissmann, M., Yin, N., Muller, J. M., Schneider, R., Peters, A. H., Gunther, T., Buettner, R.,

and Schule, R. (2005) LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. Nature 437, 436-439

122. Fang, R., Barbera, A. J., Xu, Y., Rutenberg, M., Leonor, T., Bi, Q., Lan, F., Mei, P., Yuan, G. C., Lian, C., Peng, J., Cheng, D., Sui, G., Kaiser, U. B., Shi, Y., and Shi, Y. G. (2010) Human LSD2/KDM1b/AOF1 regulates gene transcription by modulating intragenic H3K4me2 methylation. Mol Cell 39, 222-233

123. Ciccone, D. N., Su, H., Hevi, S., Gay, F., Lei, H., Bajko, J., Xu, G., Li, E., and Chen, T. (2009) KDM1B is a histone H3K4 demethylase required to establish maternal genomic imprints. Nature 461, 415-418

124. Klose, R. J., Kallin, E. M., and Zhang, Y. (2006) JmjC-domain-containing proteins and histone demethylation. Nat Rev Genet 7, 715-727

125. Outchkourov, N. S., Muino, J. M., Kaufmann, K., van Ijcken, W. F., Groot Koerkamp, M. J., van Leenen, D., de Graaf, P., Holstege, F. C., Grosveld, F. G., and Timmers, H. T. (2013) Balancing of histone H3K4 methylation states by the Kdm5c/SMCX histone demethylase modulates promoter and enhancer function. Cell Rep 3, 1071-1079

126. Walport, L. J., Hopkinson, R. J., Vollmar, M., Madden, S. K., Gileadi, C., Oppermann, U., Schofield, C. J., and Johansson, C. (2014) Human UTY(KDM6C) is a male-specific N-methyl lysyl demethylase. J Biol Chem 289, 18302-18313

127. Woelk, T., Sigismund, S., Penengo, L., and Polo, S. (2007) The ubiquitination code: a signalling problem. Cell Div 2, 11

128. Walter, D., Matter, A., and Fahrenkrog, B. (2010) Bre1p-mediated histone H2B ubiquitylation regulates apoptosis in Saccharomyces cerevisiae. J Cell Sci 123, 1931-1939

129. Hwang, W. W., and Madhani, H. D. (2009) Nonredundant requirement for multiple histone modifications for the early anaphase release of the mitotic exit regulator Cdc14 from nucleolar chromatin. PLoS Genet 5, e1000588

130. Fuchs, G., Shema, E., Vesterman, R., Kotler, E., Wolchinsky, Z., Wilder, S., Golomb, L., Pribluda, A., Zhang, F., Haj-Yahya, M., Feldmesser, E., Brik, A., Yu, X., Hanna, J., Aberdam, D., Domany, E., and Oren, M. (2012) RNF20 and USP44 regulate stem cell differentiation by modulating H2B monoubiquitylation. Mol Cell 46, 662-673

131. Wright, D. E., Wang, C. Y., and Kao, C. F. (2011) Flickin' the ubiquitin switch: the role of H2B ubiquitylation in development. Epigenetics 6, 1165-1175

132. Shema, E., Tirosh, I., Aylon, Y., Huang, J., Ye, C., Moskovits, N., Raver-Shapira, N., Minsky, N., Pirngruber, J., Tarcic, G., Hublarova, P., Moyal, L., Gana-Weisz, M., Shiloh, Y., Yarden, Y., Johnsen, S. A., Vojtesek, B., Berger, S. L., and Oren, M. (2008) The histone H2B-specific ubiquitin ligase RNF20/hBRE1 acts as a putative tumor suppressor through selective regulation of gene expression. Genes Dev 22, 2664-2676

133. Stock, J. K., Giadrossi, S., Casanova, M., Brookes, E., Vidal, M., Koseki, H., Brockdorff, N., Fisher, A. G., and Pombo, A. (2007) Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells. Nat Cell Biol 9, 1428-1435

134. Mattiroli, F., Vissers, J. H., van Dijk, W. J., Ikpa, P., Citterio, E., Vermeulen, W., Marteijn, J. A., and Sixma, T. K. (2012) RNF168 ubiquitinates K13-15 on H2A/H2AX to drive DNA damage signaling. Cell 150, 1182-1195

135. Gatti, M., Pinato, S., Maspero, E., Soffientini, P., Polo, S., and Penengo, L. (2012) A novel ubiquitin mark at the N-terminal tail of histone H2As targeted by RNF168 ubiquitin ligase. Cell Cycle 11, 2538-2544

136. Zhu, B., Zheng, Y., Pham, A. D., Mandal, S. S., Erdjument-Bromage, H., Tempst, P., and Reinberg, D. (2005) Monoubiquitination of human histone H2B: the factors involved and their roles in HOX gene regulation. Mol Cell 20, 601-611

137. Thorne, A. W., Sautiere, P., Briand, G., and Crane-Robinson, C. (1987) The structure of ubiquitinated histone H2B. EMBO J 6, 1005-1010

138. Sun, Z. W., and Allis, C. D. (2002) Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. Nature 418, 104-108

139. Wood, A., Krogan, N. J., Dover, J., Schneider, J., Heidt, J., Boateng, M. A., Dean, K., Golshani, A., Zhang, Y., Greenblatt, J. F., Johnston, M., and Shilatifard, A. (2003) Bre1, an E3 ubiquitin ligase required for recruitment and substrate selection of Rad6 at a promoter. Mol Cell 11, 267-274

140. Bonnet, J., Devys, D., and Tora, L. (2014) Histone H2B ubiquitination: signaling not scrapping. Drug Discov Today Technol 12, e19-27

141. Clague, M. J., Barsukov, I., Coulson, J. M., Liu, H., Rigden, D. J., and Urbe, S. (2013) Deubiquitylases from

genes to organism. Physiol Rev 93, 1289-1315

142. Nijman, S. M., Luna-Vargas, M. P., Velds, A., Brummelkamp, T. R., Dirac, A. M., Sixma, T. K., and Bernards, R. (2005) A genomic and functional inventory of deubiquitinating enzymes. Cell 123, 773-786

143. Baymaz, H. I., Fournier, A., Laget, S., Ji, Z., Jansen, P. W., Smits, A. H., Ferry, L., Mensinga, A., Poser, I., Sharrocks, A., Defossez, P. A., and Vermeulen, M. (2014) MBD5 and MBD6 interact with the human PR-DUB complex through their methyl-CpG-binding domain. Proteomics 14, 2179-2189

144. Park, U. H., Kang, M. R., Kim, E. J., Kwon, Y. S., Hur, W., Yoon, S. K., Song, B. J., Park, J. H., Hwang, J. T., Jeong, J. C., and Um, S. J. (2016) ASXL2 promotes proliferation of breast cancer cells by linking ERalpha to histone methylation. Oncogene 35, 3742-3752

145. Daniel, J. A., Torok, M. S., Sun, Z. W., Schieltz, D., Allis, C. D., Yates, J. R., 3rd, and Grant, P. A. (2004) Deubiquitination of histone H2B by a yeast acetyltransferase complex regulates transcription. J Biol Chem 279, 1867-1871

146. Morgan, M. T., Haj-Yahya, M., Ringel, A. E., Bandi, P., Brik, A., and Wolberger, C. (2016) Structural basis for histone H2B deubiquitination by the SAGA DUB module. Science 351, 725-728

147. Joo, H. Y., Jones, A., Yang, C., Zhai, L., Smith, A. D. t., Zhang, Z., Chandrasekharan, M. B., Sun, Z. W., Renfrow, M. B., Wang, Y., Chang, C., and Wang, H. (2011) Regulation of histone H2A and H2B deubiquitination and Xenopus development by USP12 and USP46. J Biol Chem 286, 7190-7201

148. Nicassio, F., Corrado, N., Vissers, J. H., Areces, L. B., Bergink, S., Marteijn, J. A., Geverts, B., Houtsmuller, A. B., Vermeulen, W., Di Fiore, P. P., and Citterio, E. (2007) Human USP3 is a chromatin modifier required for S phase progression and genome stability. Curr Biol 17, 1972-1977

149. van der Knaap, J. A., Kumar, B. R., Moshkin, Y. M., Langenberg, K., Krijgsveld, J., Heck, A. J., Karch, F., and Verrijzer, C. P. (2005) GMP synthetase stimulates histone H2B deubiquitylation by the epigenetic silencer USP7. Mol Cell 17, 695-707

150. Zhang, X. Y., Varthi, M., Sykes, S. M., Phillips, C., Warzecha, C., Zhu, W., Wyce, A., Thorne, A. W., Berger, S. L., and McMahon, S. B. (2008) The putative cancer stem cell marker USP22 is a subunit of the human SAGA complex required for activated transcription and cell-cycle progression. Mol Cell 29, 102-111

151. Zhang, Z., Jones, A., Joo, H. Y., Zhou, D., Cao, Y., Chen, S., Erdjument-Bromage, H., Renfrow, M., He, H., Tempst, P., Townes, T. M., Giles, K. E., Ma, L., and Wang, H. (2013) USP49 deubiquitinates histone H2B and regulates cotranscriptional pre-mRNA splicing. Genes Dev 27, 1581-1595

152. Drivas, G. T., Shih, A., Coutavas, E., Rush, M. G., and D'Eustachio, P. (1990) Characterization of four novel ras-like genes expressed in a human teratocarcinoma cell line. Mol Cell Biol 10, 1793-1798

153. Scheffzek, K., Klebe, C., Fritz-Wolf, K., Kabsch, W., and Wittinghofer, A. (1995) Crystal structure of the nuclear Ras-related protein Ran in its GDP-bound form. Nature 374, 378-381

154. Vetter, I. R., Nowak, C., Nishimoto, T., Kuhlmann, J., and Wittinghofer, A. (1999) Structure of a Ran-binding domain complexed with Ran bound to a GTP analogue: implications for nuclear transport. Nature 398, 39-46

155. Yaseen, N. R., and Blobel, G. (1997) Cloning and characterization of human karyopherin beta3. Proc Natl Acad Sci U S A 94, 4451-4456

156. Mosammaparast, N., and Pemberton, L. F. (2004) Karyopherins: from nuclear-transport mediators to nuclear-function regulators. Trends Cell Biol 14, 547-556

157. Melchior, F., Paschal, B., Evans, J., and Gerace, L. (1993) Inhibition of nuclear protein import by nonhydrolyzable analogues of GTP and identification of the small GTPase Ran/TC4 as an essential transport factor. J Cell Biol 123, 1649-1659

158. Moore, M. S., and Blobel, G. (1993) The GTP-binding protein Ran/TC4 is required for protein import into the nucleus. Nature 365, 661-663

159. Chen, X., and Xu, L. (2010) Specific nucleoporin requirement for Smad nuclear translocation. Mol Cell Biol 30, 4022-4034

160. Ribbeck, K., Lipowsky, G., Kent, H. M., Stewart, M., and Gorlich, D. (1998) NTF2 mediates nuclear import of Ran. EMBO J 17, 6587-6598

161. Smith, A., Brownawell, A., and Macara, I. G. (1998) Nuclear import of Ran is mediated by the transport factor NTF2. Curr Biol 8, 1403-1406

162. Bischoff, F. R., and Ponstingl, H. (1991) Catalysis of guanine nucleotide exchange on Ran by the mitotic regulator RCC1. Nature 354, 80-82

163. Klebe, C., Bischoff, F. R., Ponstingl, H., and Wittinghofer, A. (1995) Interaction of the nuclear GTP-binding protein Ran with its regulatory proteins RCC1 and RanGAP1. Biochemistry 34, 639-647

164. Bischoff, F. R., and Ponstingl, H. (1995) Catalysis of guanine nucleotide exchange of Ran by RCC1 and stimulation of hydrolysis of Ran-bound GTP by Ran-GAP1. Methods Enzymol 257, 135-144

165. Kutay, U., Bischoff, F. R., Kostka, S., Kraft, R., and Gorlich, D. (1997) Export of importin alpha from the nucleus is mediated by a specific nuclear transport factor. Cell 90, 1061-1071

166. Deane, R., Schafer, W., Zimmermann, H. P., Mueller, L., Gorlich, D., Prehn, S., Ponstingl, H., and Bischoff, F. R. (1997) Ran-binding protein 5 (RanBP5) is related to the nuclear transport factor importin-beta but interacts differently with RanBP1. Mol Cell Biol 17, 5087-5096

167. Huminiecki, L., Goldovsky, L., Freilich, S., Moustakas, A., Ouzounis, C., and Heldin, C. H. (2009) Emergence, development and diversification of the TGF-beta signalling pathway within the animal kingdom. BMC Evol Biol 9, 28

168. Assoian, R. K., Komoriya, A., Meyers, C. A., Miller, D. M., and Sporn, M. B. (1983) Transforming growth factor-beta in human platelets. Identification of a major storage site, purification, and characterization. J Biol Chem 258, 7155-7160

169. Herpin, A., Lelong, C., and Favrel, P. (2004) Transforming growth factor-beta-related proteins: an ancestral and widespread superfamily of cytokines in metazoans. Dev Comp Immunol 28, 461-485

170. Savage, C., Das, P., Finelli, A. L., Townsend, S. R., Sun, C. Y., Baird, S. E., and Padgett, R. W. (1996) Caenorhabditis elegans genes sma-2, sma-3, and sma-4 define a conserved family of transforming growth factor beta pathway components. Proc Natl Acad Sci U S A 93, 790-794

171. Sekelsky, J. J., Newfeld, S. J., Raftery, L. A., Chartoff, E. H., and Gelbart, W. M. (1995) Genetic characterization and cloning of mothers against dpp, a gene required for decapentaplegic function in Drosophila melanogaster. Genetics 139, 1347-1358

172. Miyazawa, K., Shinozaki, M., Hara, T., Furuya, T., and Miyazono, K. (2002) Two major Smad pathways in TGF-beta superfamily signalling. Genes Cells 7, 1191-1204

173. He, K., Yan, X., Li, N., Dang, S., Xu, L., Zhao, B., Li, Z., Lv, Z., Fang, X., Zhang, Y., and Chen, Y. G. (2015) Internalization of the TGF-beta type I receptor into caveolin-1 and EEA1 double-positive early endosomes. Cell Res 25, 738-752

174. Itoh, F., Divecha, N., Brocks, L., Oomen, L., Janssen, H., Calafat, J., Itoh, S., and Dijke Pt, P. (2002) The FYVE domain in Smad anchor for receptor activation (SARA) is sufficient for localization of SARA in early endosomes and regulates TGF-beta/Smad signalling. Genes Cells 7, 321-331

175. Shi, W., Chang, C., Nie, S., Xie, S., Wan, M., and Cao, X. (2007) Endofin acts as a Smad anchor for receptor activation in BMP signaling. J Cell Sci 120, 1216-1224

176. Tsukazaki, T., Chiang, T. A., Davison, A. F., Attisano, L., and Wrana, J. L. (1998) SARA, a FYVE domain protein that recruits Smad2 to the TGFbeta receptor. Cell 95, 779-791

177. Qin, B. Y., Chacko, B. M., Lam, S. S., de Caestecker, M. P., Correia, J. J., and Lin, K. (2001) Structural basis of Smad1 activation by receptor kinase phosphorylation. Mol Cell 8, 1303-1312

178. Chacko, B. M., Qin, B. Y., Tiwari, A., Shi, G., Lam, S., Hayward, L. J., De Caestecker, M., and Lin, K. (2004) Structural basis of heteromeric smad protein assembly in TGF-beta signaling. Mol Cell 15, 813-823

179. Xu, L., Yao, X., Chen, X., Lu, P., Zhang, B., and Ip, Y. T. (2007) Msk is required for nuclear import of TGF-{beta}/BMP-activated Smads. J Cell Biol 178, 981-994

180. Yao, X., Chen, X., Cottonham, C., and Xu, L. (2008) Preferential utilization of Imp7/8 in nuclear import of Smads. J Biol Chem 283, 22867-22874

181. Xiao, Z., Liu, X., and Lodish, H. F. (2000) Importin beta mediates nuclear translocation of Smad 3. J Biol Chem 275, 23425-23428

182. Xiao, Z., Liu, X., Henis, Y. I., and Lodish, H. F. (2000) A distinct nuclear localization signal in the N terminus of Smad 3 determines its ligand-induced nuclear translocation. Proc Natl Acad Sci U S A 97, 7853-7858

183. Xiao, Z., Latek, R., and Lodish, H. F. (2003) An extended bipartite nuclear localization signal in Smad4 is required for its nuclear import and transcriptional activity. Oncogene 22, 1057-1069

184. Watanabe, M., Masuyama, N., Fukuda, M., and Nishida, E. (2000) Regulation of intracellular dynamics of Smad4 by its leucine-rich nuclear export signal. EMBO Rep 1, 176-182

185. Shi, Y., Wang, Y. F., Jayaraman, L., Yang, H., Massague, J., and Pavletich, N. P. (1998) Crystal structure of a

Smad MH1 domain bound to DNA: insights on DNA binding in TGF-beta signaling. Cell 94, 585-594

186. Baburajendran, N., Jauch, R., Tan, C. Y., Narasimhan, K., and Kolatkar, P. R. (2011) Structural basis for the cooperative DNA recognition by Smad4 MH1 dimers. Nucleic Acids Res 39, 8213-8222

187. Baburajendran, N., Palasingam, P., Narasimhan, K., Sun, W., Prabhakar, S., Jauch, R., and Kolatkar, P. R. (2010) Structure of Smad1 MH1/DNA complex reveals distinctive rearrangements of BMP and TGF-beta effectors. Nucleic Acids Res 38, 3477-3488

188. Zawel, L., Dai, J. L., Buckhaults, P., Zhou, S., Kinzler, K. W., Vogelstein, B., and Kern, S. E. (1998) Human Smad3 and Smad4 are sequence-specific transcription activators. Mol Cell 1, 611-617

189. Yagi, K., Goto, D., Hamamoto, T., Takenoshita, S., Kato, M., and Miyazono, K. (1999) Alternatively spliced variant of Smad2 lacking exon 3. Comparison with wild-type Smad2 and Smad3. J Biol Chem 274, 703-709

190. Gao, S., Steffen, J., and Laughon, A. (2005) Dpp-responsive silencers are bound by a trimeric Mad-Medea complex. J Biol Chem 280, 36158-36164

191. Pyrowolakis, G., Hartmann, B., Muller, B., Basler, K., and Affolter, M. (2004) A simple molecular complex mediates widespread BMP-induced repression during Drosophila development. Dev Cell 7, 229-240

192. Weiss, A., Charbonnier, E., Ellertsdottir, E., Tsirigos, A., Wolf, C., Schuh, R., Pyrowolakis, G., and Affolter, M. (2010) A conserved activation element in BMP signaling during Drosophila development. Nat Struct Mol Biol 17, 69-76

193. Blitz, I. L., and Cho, K. W. (2009) Finding partners: how BMPs select their targets. Dev Dyn 238, 1321-1331

194. Korchynskyi, O., and ten Dijke, P. (2002) Identification and functional characterization of distinct critically important bone morphogenetic protein-specific response elements in the Id1 promoter. J Biol Chem 277, 4883-4891

195. Morikawa, M., Koinuma, D., Tsutsumi, S., Vasilaki, E., Kanki, Y., Heldin, C. H., Aburatani, H., and Miyazono, K. (2011) ChIP-seq reveals cell type-specific binding patterns of BMP-specific Smads and a novel binding motif. Nucleic Acids Res 39, 8712-8727

196. Nakahiro, T., Kurooka, H., Mori, K., Sano, K., and Yokota, Y. (2010) Identification of BMP-responsive elements in the mouse Id2 gene. Biochem Biophys Res Commun 399, 416-421

197. Zhang, Y., Feng, X. H., and Derynck, R. (1998) Smad3 and Smad4 cooperate with c-Jun/c-Fos to mediate TGF-beta-induced transcription. Nature 394, 909-913

198. Koinuma, D., Tsutsumi, S., Kamimura, N., Taniguchi, H., Miyazawa, K., Sunamura, M., Imamura, T., Miyazono, K., and Aburatani, H. (2009) Chromatin immunoprecipitation on microarray analysis of Smad2/3 binding sites reveals roles of ETS1 and TFAP2A in transforming growth factor beta signaling. Mol Cell Biol 29, 172-186

199. Gomis, R. R., Alarcon, C., Nadal, C., Van Poznak, C., and Massague, J. (2006) C/EBPbeta at the core of the TGFbeta cytostatic response and its evasion in metastatic breast cancer cells. Cancer Cell 10, 203-214

200. Silvestri, C., Narimatsu, M., von Both, I., Liu, Y., Tan, N. B., Izzi, L., McCaffery, P., Wrana, J. L., and Attisano, L. (2008) Genome-wide identification of Smad/Foxh1 targets reveals a role for Foxh1 in retinoic acid regulation and forebrain development. Dev Cell 14, 411-423

201. Attisano, L., Silvestri, C., Izzi, L., and Labbe, E. (2001) The transcriptional role of Smads and FAST (FoxH1) in TGFbeta and activin signalling. Mol Cell Endocrinol 180, 3-11

202. Gomis, R. R., Alarcon, C., He, W., Wang, Q., Seoane, J., Lash, A., and Massague, J. (2006) A FoxO-Smad synexpression group in human keratinocytes. Proc Natl Acad Sci U S A 103, 12747-12752

203. Mullen, A. C., Orlando, D. A., Newman, J. J., Loven, J., Kumar, R. M., Bilodeau, S., Reddy, J., Guenther, M. G., DeKoter, R. P., and Young, R. A. (2011) Master transcription factors determine cell-type-specific responses to TGF-beta signaling. Cell 147, 565-576

204. Ikushima, H., and Miyazono, K. (2010) TGFbeta signalling: a complex web in cancer progression. Nat Rev Cancer 10, 415-424

205. de Jong, D. S., Vaes, B. L., Dechering, K. J., Feijen, A., Hendriks, J. M., Wehrens, R., Mummery, C. L., van Zoelen, E. J., Olijve, W., and Steegenga, W. T. (2004) Identification of novel regulators associated with early-phase osteoblast differentiation. J Bone Miner Res 19, 947-958

206. Levy, L., and Hill, C. S. (2005) Smad4 dependency defines two classes of transforming growth factor {beta} (TGF-{beta}) target genes and distinguishes TGF-{beta}-induced epithelial-mesenchymal transition from its antiproliferative and migratory responses. Mol Cell Biol 25, 8108-8125

207. Locklin, R. M., Riggs, B. L., Hicok, K. C., Horton, H. F., Byrne, M. C., and Khosla, S. (2001) Assessment of gene regulation by bone morphogenetic protein 2 in human marrow stromal cells using gene array technology. J Bone Miner Res 16, 2192-2204

208. Prashar, P., Yadav, P. S., Samarjeet, F., and Bandyopadhyay, A. (2014) Microarray meta-analysis identifies evolutionarily conserved BMP signaling targets in developing long bones. Dev Biol 389, 192-207

209. Nishihara, A., Hanai, J. I., Okamoto, N., Yanagisawa, J., Kato, S., Miyazono, K., and Kawabata, M. (1998) Role of p300, a transcriptional coactivator, in signalling of TGF-beta. Genes Cells 3, 613-623

210. Kang, Y., Chen, C. R., and Massague, J. (2003) A self-enabling TGFbeta response coupled to stress signaling: Smad engages stress response factor ATF3 for Id1 repression in epithelial cells. Mol Cell 11, 915-926

211. Germain, S., Howell, M., Esslemont, G. M., and Hill, C. S. (2000) Homeodomain and winged-helix transcription factors recruit activated Smads to distinct promoter elements via a common Smad interaction motif. Genes Dev 14, 435-451

212. Kim, S. W., Yoon, S. J., Chuong, E., Oyolu, C., Wills, A. E., Gupta, R., and Baker, J. (2011) Chromatin and transcriptional signatures for Nodal signaling during endoderm formation in hESCs. Dev Biol 357, 492-504

213. Kang, J. S., Alliston, T., Delston, R., and Derynck, R. (2005) Repression of Runx2 function by TGF-beta through recruitment of class II histone deacetylases by Smad3. EMBO J 24, 2543-2555

214. Beyer, T. A., Weiss, A., Khomchuk, Y., Huang, K., Ogunjimi, A. A., Varelas, X., and Wrana, J. L. (2013) Switch enhancers interpret TGF-beta and Hippo signaling to control cell fate in human embryonic stem cells. Cell Rep 5, 1611-1624

215. Hayashi, H., Abdollah, S., Qiu, Y., Cai, J., Xu, Y. Y., Grinnell, B. W., Richardson, M. A., Topper, J. N., Gimbrone, M. A., Jr., Wrana, J. L., and Falb, D. (1997) The MAD-related protein Smad7 associates with the TGFbeta receptor and functions as an antagonist of TGFbeta signaling. Cell 89, 1165-1173

216. Imamura, T., Takase, M., Nishihara, A., Oeda, E., Hanai, J., Kawabata, M., and Miyazono, K. (1997) Smad6 inhibits signalling by the TGF-beta superfamily. Nature 389, 622-626

217. Nakao, A., Afrakhte, M., Moren, A., Nakayama, T., Christian, J. L., Heuchel, R., Itoh, S., Kawabata, M., Heldin, N. E., Heldin, C. H., and ten Dijke, P. (1997) Identification of Smad7, a TGFbeta-inducible antagonist of TGF-beta signalling. Nature 389, 631-635

218. Wang, Q., Wei, X., Zhu, T., Zhang, M., Shen, R., Xing, L., O'Keefe, R. J., and Chen, D. (2007) Bone morphogenetic protein 2 activates Smad6 gene transcription through bone-specific transcription factor Runx2. J Biol Chem 282, 10742-10748

219. Hanyu, A., Ishidou, Y., Ebisawa, T., Shimanuki, T., Imamura, T., and Miyazono, K. (2001) The N domain of Smad7 is essential for specific inhibition of transforming growth factor-beta signaling. J Cell Biol 155, 1017-1027

220. Souchelnytskyi, S., Nakayama, T., Nakao, A., Moren, A., Heldin, C. H., Christian, J. L., and ten Dijke, P. (1998) Physical and functional interaction of murine and Xenopus Smad7 with bone morphogenetic protein receptors and transforming growth factor-beta receptors. J Biol Chem 273, 25364-25370

221. Hata, A., Lagna, G., Massague, J., and Hemmati-Brivanlou, A. (1998) Smad6 inhibits BMP/Smad1 signaling by specifically competing with the Smad4 tumor suppressor. Genes Dev 12, 186-197

222. Ebisawa, T., Fukuchi, M., Murakami, G., Chiba, T., Tanaka, K., Imamura, T., and Miyazono, K. (2001) Smurf1 interacts with transforming growth factor-beta type I receptor through Smad7 and induces receptor degradation. J Biol Chem 276, 12477-12480

223. Kavsak, P., Rasmussen, R. K., Causing, C. G., Bonni, S., Zhu, H., Thomsen, G. H., and Wrana, J. L. (2000) Smad7 binds to Smurf2 to form an E3 ubiquitin ligase that targets the TGF beta receptor for degradation. Mol Cell 6, 1365-1375

224. Murakami, G., Watabe, T., Takaoka, K., Miyazono, K., and Imamura, T. (2003) Cooperative inhibition of bone morphogenetic protein signaling by Smurf1 and inhibitory Smads. Mol Biol Cell 14, 2809-2817

225. Stroschein, S. L., Wang, W., Zhou, S., Zhou, Q., and Luo, K. (1999) Negative feedback regulation of TGF-beta signaling by the SnoN oncoprotein. Science 286, 771-774

226. Sun, Y., Liu, X., Eaton, E. N., Lane, W. S., Lodish, H. F., and Weinberg, R. A. (1999) Interaction of the Ski oncoprotein with Smad3 regulates TGF-beta signaling. Mol Cell 4, 499-509

227. Lin, F., Morrison, J. M., Wu, W., and Worman, H. J. (2005) MAN1, an integral protein of the inner nuclear membrane, binds Smad2 and Smad3 and antagonizes transforming growth factor-beta signaling. Hum Mol Genet 14, 437-445

228. Pan, D., Estevez-Salmeron, L. D., Stroschein, S. L., Zhu, X., He, J., Zhou, S., and Luo, K. (2005) The integral

inner nuclear membrane protein MAN1 physically interacts with the R-Smad proteins to repress signaling by the transforming growth factor-{beta} superfamily of cytokines. J Biol Chem 280, 15992-16001

229. Luo, K., Stroschein, S. L., Wang, W., Chen, D., Martens, E., Zhou, S., and Zhou, Q. (1999) The Ski oncoprotein interacts with the Smad proteins to repress TGFbeta signaling. Genes Dev 13, 2196-2206

230. Luo, K. (2004) Ski and SnoN: negative regulators of TGF-beta signaling. Curr Opin Genet Dev 14, 65-70

231. Ueki, N., and Hayman, M. J. (2003) Signal-dependent N-CoR requirement for repression by the Ski oncoprotein. J Biol Chem 278, 24858-24864

232. Bonni, S., Wang, H. R., Causing, C. G., Kavsak, P., Stroschein, S. L., Luo, K., and Wrana, J. L. (2001) TGF-beta induces assembly of a Smad2-Smurf2 ubiquitin ligase complex that targets SnoN for degradation. Nat Cell Biol 3, 587-595

233. Sun, Y., Liu, X., Ng-Eaton, E., Lodish, H. F., and Weinberg, R. A. (1999) SnoN and Ski protooncoproteins are rapidly degraded in response to transforming growth factor beta signaling. Proc Natl Acad Sci U S A 96, 12442-12447

234. Bitzer, M., von Gersdorff, G., Liang, D., Dominguez-Rosales, A., Beg, A. A., Rojkind, M., and Bottinger, E. P. (2000) A mechanism of suppression of TGF-beta/SMAD signaling by NF-kappa B/RelA. Genes Dev 14, 187-197

235. Ulloa, L., Doody, J., and Massague, J. (1999) Inhibition of transforming growth factor-beta/SMAD signalling by the interferon-gamma/STAT pathway. Nature 397, 710-713

236. Mulder, K. M., and Morris, S. L. (1992) Activation of p21ras by transforming growth factor beta in epithelial cells. J Biol Chem 267, 5029-5031

237. Yan, Z., Winawer, S., and Friedman, E. (1994) Two different signal transduction pathways can be activated by transforming growth factor beta 1 in epithelial cells. J Biol Chem 269, 13231-13237

238. Thiery, J. P. (2003) Epithelial-mesenchymal transitions in development and pathologies. Curr Opin Cell Biol 15, 740-746

239. Engel, M. E., McDonnell, M. A., Law, B. K., and Moses, H. L. (1999) Interdependent SMAD and JNK signaling in transforming growth factor-beta-mediated transcription. J Biol Chem 274, 37413-37420

240. Frey, R. S., and Mulder, K. M. (1997) Involvement of extracellular signal-regulated kinase 2 and stress-activated protein kinase/Jun N-terminal kinase activation by transforming growth factor beta in the negative growth control of breast cancer cells. Cancer Res 57, 628-633

241. Hanafusa, H., Ninomiya-Tsuji, J., Masuyama, N., Nishita, M., Fujisawa, J., Shibuya, H., Matsumoto, K., and Nishida, E. (1999) Involvement of the p38 mitogen-activated protein kinase pathway in transforming growth factor-beta-induced gene expression. J Biol Chem 274, 27161-27167

242. Edlund, S., Bu, S., Schuster, N., Aspenstrom, P., Heuchel, R., Heldin, N. E., ten Dijke, P., Heldin, C. H., and Landstrom, M. (2003) Transforming growth factor-beta1 (TGF-beta)-induced apoptosis of prostate cancer cells involves Smad7-dependent activation of p38 by TGF-beta-activated kinase 1 and mitogen-activated protein kinase kinase 3. Mol Biol Cell 14, 529-544

243. Kimura, N., Matsuo, R., Shibuya, H., Nakashima, K., and Taga, T. (2000) BMP2-induced apoptosis is mediated by activation of the TAK1-p38 kinase pathway that is negatively regulated by Smad6. J Biol Chem 275, 17647-17652

244. Jaffe, A. B., and Hall, A. (2005) Rho GTPases: biochemistry and biology. Annu Rev Cell Dev Biol 21, 247-269

245. Bhowmick, N. A., Ghiassi, M., Bakin, A., Aakre, M., Lundquist, C. A., Engel, M. E., Arteaga, C. L., and Moses, H. L. (2001) Transforming growth factor-beta1 mediates epithelial to mesenchymal transdifferentiation through a RhoA-dependent mechanism. Mol Biol Cell 12, 27-36

246. Edlund, S., Landstrom, M., Heldin, C. H., and Aspenstrom, P. (2002) Transforming growth factor-beta-induced mobilization of actin cytoskeleton requires signaling by small GTPases Cdc42 and RhoA. Mol Biol Cell 13, 902-914

247. Bakin, A. V., Tomlinson, A. K., Bhowmick, N. A., Moses, H. L., and Arteaga, C. L. (2000) Phosphatidylinositol 3-kinase function is required for transforming growth factor beta-mediated epithelial to mesenchymal transition and cell migration. J Biol Chem 275, 36803-36810

248. Wilkes, M. C., Mitchell, H., Penheiter, S. G., Dore, J. J., Suzuki, K., Edens, M., Sharma, D. K., Pagano, R. E., and Leof, E. B. (2005) Transforming growth factor-beta activation of phosphatidylinositol 3-kinase is independent of Smad2 and Smad3 and regulates fibroblast responses via p21-activated kinase-2. Cancer Res 65, 10431-10440

249. Yi, J. Y., Shin, I., and Arteaga, C. L. (2005) Type I transforming growth factor beta receptor binds to and activates phosphatidylinositol 3-kinase. J Biol Chem 280, 10870-10876

250. Massague, J. (1998) TGF-beta signal transduction. Annu Rev Biochem 67, 753-791

251. Hahn, S. A., Schutte, M., Hoque, A. T., Moskaluk, C. A., da Costa, L. T., Rozenblum, E., Weinstein, C. L., Fischer, A., Yeo, C. J., Hruban, R. H., and Kern, S. E. (1996) DPC4, a candidate tumor suppressor gene at human chromosome 18q21.1. Science 271, 350-353

252. Kandoth, C., McLellan, M. D., Vandin, F., Ye, K., Niu, B., Lu, C., Xie, M., Zhang, Q., McMichael, J. F., Wyczalkowski, M. A., Leiserson, M. D., Miller, C. A., Welch, J. S., Walter, M. J., Wendl, M. C., Ley, T. J., Wilson, R. K., Raphael, B. J., and Ding, L. (2013) Mutational landscape and significance across 12 major cancer types. Nature 502, 333-339

253. Lawrence, M. S., Stojanov, P., Mermel, C. H., Robinson, J. T., Garraway, L. A., Golub, T. R., Meyerson, M., Gabriel, S. B., Lander, E. S., and Getz, G. (2014) Discovery and saturation analysis of cancer genes across 21 tumour types. Nature 505, 495-501

254. Ojesina, A. I., Lichtenstein, L., Freeman, S. S., Pedamallu, C. S., Imaz-Rosshandler, I., Pugh, T. J., Cherniack, A. D., Ambrogio, L., Cibulskis, K., Bertelsen, B., Romero-Cordoba, S., Trevino, V., Vazquez-Santillan, K., Guadarrama, A. S., Wright, A. A., Rosenberg, M. W., Duke, F., Kaplan, B., Wang, R., Nickerson, E., Walline, H. M., Lawrence, M. S., Stewart, C., Carter, S. L., McKenna, A., Rodriguez-Sanchez, I. P., Espinosa-Castilla, M., Woie, K., Bjorge, L., Wik, E., Halle, M. K., Hoivik, E. A., Krakstad, C., Gabino, N. B., Gomez-Macias, G. S., Valdez-Chapa, L. D., Garza-Rodriguez, M. L., Maytorena, G., Vazquez, J., Rodea, C., Cravioto, A., Cortes, M. L., Greulich, H., Crum, C. P., Neuberg, D. S., Hidalgo-Miranda, A., Escareno, C. R., Akslen, L. A., Carey, T. E., Vintermyr, O. K., Gabriel, S. B., Barrera-Saldana, H. A., Melendez-Zajgla, J., Getz, G., Salvesen, H. B., and Meyerson, M. (2014) Landscape of genomic alterations in cervical carcinomas. Nature 506, 371-375

255. Levy, L., and Hill, C. S. (2006) Alterations in components of the TGF-beta superfamily signaling pathways in human cancer. Cytokine Growth Factor Rev 17, 41-58

256. Massague, J. (2008) TGFbeta in Cancer. Cell 134, 215-230

257. Derynck, R., Akhurst, R. J., and Balmain, A. (2001) TGF-beta signaling in tumor suppression and cancer progression. Nat Genet 29, 117-129

258. Wikstrom, P., Stattin, P., Franck-Lissbrant, I., Damber, J. E., and Bergh, A. (1998) Transforming growth factor beta1 is associated with angiogenesis, metastasis, and poor clinical outcome in prostate cancer. Prostate 37, 19-29

259. Ito, N., Kawata, S., Tamura, S., Shirai, Y., Kiso, S., Tsushima, H., and Matsuzawa, Y. (1995) Positive correlation of plasma transforming growth factor-beta 1 levels with tumor vascularity in hepatocellular carcinoma. Cancer Lett 89, 45-48

260. Janda, E., Lehmann, K., Killisch, I., Jechlinger, M., Herzig, M., Downward, J., Beug, H., and Grunert, S. (2002) Ras and TGF[beta] cooperatively regulate epithelial cell plasticity and metastasis: dissection of Ras signaling pathways. J Cell Biol 156, 299-313

261. David, C. J., Huang, Y. H., Chen, M., Su, J., Zou, Y., Bardeesy, N., Iacobuzio-Donahue, C. A., and Massague, J. (2016) TGF-beta Tumor Suppression through a Lethal EMT. Cell 164, 1015-1030

262. Shimomura, O., Johnson, F. H., and Saiga, Y. (1962) Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, Aequorea. J Cell Comp Physiol 59, 223-239

263. Ormo, M., Cubitt, A. B., Kallio, K., Gross, L. A., Tsien, R. Y., and Remington, S. J. (1996) Crystal structure of the Aequorea victoria green fluorescent protein. Science 273, 1392-1395

264. Shaner, N. C., Steinbach, P. A., and Tsien, R. Y. (2005) A guide to choosing fluorescent proteins. Nat Methods 2, 905-909

265. Heim, R., Cubitt, A. B., and Tsien, R. Y. (1995) Improved green fluorescence. Nature 373, 663-664

266. McRae, S. R., Brown, C. L., and Bushell, G. R. (2005) Rapid purification of EGFP, EYFP, and ECFP with high yield and purity. Protein Expr Purif 41, 121-127

267. Kirchhofer, A., Helma, J., Schmidthals, K., Frauer, C., Cui, S., Karcher, A., Pellis, M., Muyldermans, S., Casas-Delucchi, C. S., Cardoso, M. C., Leonhardt, H., Hopfner, K. P., and Rothbauer, U. (2010) Modulation of protein properties in living cells using nanobodies. Nat Struct Mol Biol 17, 133-138

268. Baymaz, H. I., Spruijt, C. G., and Vermeulen, M. (2014) Identifying nuclear protein-protein interactions using GFP affinity purification and SILAC-based quantitative mass spectrometry. Methods Mol Biol 1188, 207-226

269. Rothbauer, U., Zolghadr, K., Tillib, S., Nowak, D., Schermelleh, L., Gahl, A., Backmann, N., Conrath, K., Muyldermans, S., Cardoso, M. C., and Leonhardt, H. (2006) Targeting and tracing antigens in live cells with

fluorescent nanobodies. Nat Methods 3, 887-889

270. Rothbauer, U., Zolghadr, K., Muyldermans, S., Schepers, A., Cardoso, M. C., and Leonhardt, H. (2008) A versatile nanotrap for biochemical and functional studies with fluorescent fusion proteins. Mol Cell Proteomics 7, 282-289

271. Schmidthals, K., Helma, J., Zolghadr, K., Rothbauer, U., and Leonhardt, H. (2010) Novel antibody derivatives for proteome and high-content analysis. Anal Bioanal Chem 397, 3203-3208

272. Inman, G. J., Nicolas, F. J., Callahan, J. F., Harling, J. D., Gaster, L. M., Reith, A. D., Laping, N. J., and Hill, C. S. (2002) SB-431542 is a potent and specific inhibitor of transforming growth factor-beta superfamily type I activin receptor-like kinase (ALK) receptors ALK4, ALK5, and ALK7. Mol Pharmacol 62, 65-74

273. Cuny, G. D., Yu, P. B., Laha, J. K., Xing, X., Liu, J. F., Lai, C. S., Deng, D. Y., Sachidanandan, C., Bloch, K. D., and Peterson, R. T. (2008) Structure-activity relationship study of bone morphogenetic protein (BMP) signaling inhibitors. Bioorg Med Chem Lett 18, 4388-4392

274. Boergermann, J. H., Kopf, J., Yu, P. B., and Knaus, P. (2010) Dorsomorphin and LDN-193189 inhibit BMP-mediated Smad, p38 and Akt signalling in C2C12 cells. Int J Biochem Cell Biol 42, 1802-1807

275. Huang, H., Sabari, B. R., Garcia, B. A., Allis, C. D., and Zhao, Y. (2014) SnapShot: histone modifications. Cell 159, 458-458 e451

276. Rodriguez-Paredes, M., and Esteller, M. (2011) Cancer epigenetics reaches mainstream oncology. Nat Med 17, 330-339

277. Clarke, P. R., and Zhang, C. (2008) Spatial and temporal coordination of mitosis by Ran GTPase. Nat Rev Mol Cell Biol 9, 464-477

278. Berger, S. L. (2007) The complex language of chromatin regulation during transcription. Nature 447, 407-412

279. Chi, P., Allis, C. D., and Wang, G. G. (2010) Covalent histone modifications--miswritten, misinterpreted and mis-erased in human cancers. Nat Rev Cancer 10, 457-469

280. Vogelstein, B., Papadopoulos, N., Velculescu, V. E., Zhou, S., Diaz, L. A., Jr., and Kinzler, K. W. (2013) Cancer genome landscapes. Science 339, 1546-1558

281. Dawson, M. A., and Kouzarides, T. (2012) Cancer epigenetics: from mechanism to therapy. Cell 150, 12-27

282. Ruthenburg, A. J., Allis, C. D., and Wysocka, J. (2007) Methylation of lysine 4 on histone H3: intricacy of writing and reading a single epigenetic mark. Mol Cell 25, 15-30

283. van Nuland, R., van Schaik, F. M., Simonis, M., van Heesch, S., Cuppen, E., Boelens, R., Timmers, H. M., and van Ingen, H. (2013) Nucleosomal DNA binding drives the recognition of H3K36-methylated nucleosomes by the PSIP1-PWWP domain. Epigenetics Chromatin 6, 12

284. Lee, J. S., Smith, E., and Shilatifard, A. (2010) The language of histone crosstalk. Cell 142, 682-685

285. Schneider, J., Dover, J., Johnston, M., and Shilatifard, A. (2004) Global proteomic analysis of S. cerevisiae (GPS) to identify proteins required for histone modifications. Methods Enzymol 377, 227-234

286. Lee, J. S., Shukla, A., Schneider, J., Swanson, S. K., Washburn, M. P., Florens, L., Bhaumik, S. R., and Shilatifard, A. (2007) Histone crosstalk between H2B monoubiquitination and H3 methylation mediated by COMPASS. Cell 131, 1084-1096

287. Mulder, K. W., Wang, X., Escriu, C., Ito, Y., Schwarz, R. F., Gillis, J., Sirokmany, G., Donati, G., Uribe-Lewis, S., Pavlidis, P., Murrell, A., Markowetz, F., and Watt, F. M. (2012) Diverse epigenetic strategies interact to control epidermal differentiation. Nat Cell Biol 14, 753-763

288. Seidler, E., and Wohlrab, F. (1973) [Suitability of monotetrazolium salt MTT derivatives for histochemical dehydrogenase demonstration]. Acta Histochem 46, 202-208

289. Shilatifard, A. (2012) The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. Annu Rev Biochem 81, 65-95

290. R Foundation for Statistical Computing. (2012) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria

291. Christensen, J., Agger, K., Cloos, P. A., Pasini, D., Rose, S., Sennels, L., Rappsilber, J., Hansen, K. H., Salcini, A. E., and Helin, K. (2007) RBP2 belongs to a family of demethylases, specific for tri-and dimethylated lysine 4 on histone 3. Cell 128, 1063-1076

292. Wu, S. Y., and Chiang, C. M. (2007) The double bromodomain-containing chromatin adaptor Brd4 and transcriptional regulation. J Biol Chem 282, 13141-13145

293. Zuber, J., Shi, J., Wang, E., Rappaport, A. R., Herrmann, H., Sison, E. A., Magoon, D., Qi, J., Blatt, K., Wunderlich, M., Taylor, M. J., Johns, C., Chicas, A., Mulloy, J. C., Kogan, S. C., Brown, P., Valent, P., Bradner, J. E., Lowe, S. W., and Vakoc, C. R. (2011) RNAi screen identifies Brd4 as a therapeutic target in acute myeloid leukaemia. Nature 478, 524-528

294. Filippakopoulos, P., Qi, J., Picaud, S., Shen, Y., Smith, W. B., Fedorov, O., Morse, E. M., Keates, T., Hickman, T. T., Felletar, I., Philpott, M., Munro, S., McKeown, M. R., Wang, Y., Christie, A. L., West, N., Cameron, M. J., Schwartz, B., Heightman, T. D., La Thangue, N., French, C. A., Wiest, O., Kung, A. L., Knapp, S., and Bradner, J. E. (2010) Selective inhibition of BET bromodomains. Nature 468, 1067-1073

295. Nicodeme, E., Jeffrey, K. L., Schaefer, U., Beinke, S., Dewell, S., Chung, C. W., Chandwani, R., Marazzi, I., Wilson, P., Coste, H., White, J., Kirilovsky, J., Rice, C. M., Lora, J. M., Prinjha, R. K., Lee, K., and Tarakhovsky, A. (2010) Suppression of inflammation by a synthetic histone mimic. Nature 468, 1119-1123

296. Jacob, Y., and Michaels, S. D. (2009) H3K27me1 is E(z) in animals, but not in plants. Epigenetics 4, 366-369

297. Xu, K., Wu, Z. J., Groner, A. C., He, H. H., Cai, C., Lis, R. T., Wu, X., Stack, E. C., Loda, M., Liu, T., Xu, H., Cato, L., Thornton, J. E., Gregory, R. I., Morrissey, C., Vessella, R. L., Montironi, R., Magi-Galluzzi, C., Kantoff, P. W., Balk, S. P., Liu, X. S., and Brown, M. (2012) EZH2 oncogenic activity in castration-resistant prostate cancer cells is Polycomb-independent. Science 338, 1465-1469

298. Lewis, P. W., Muller, M. M., Koletsky, M. S., Cordero, F., Lin, S., Banaszynski, L. A., Garcia, B. A., Muir, T. W., Becher, O. J., and Allis, C. D. (2013) Inhibition of PRC2 activity by a gain-of-function H3 mutation found in pediatric glioblastoma. Science 340, 857-861

299. Knutson, S. K., Wigle, T. J., Warholic, N. M., Sneeringer, C. J., Allain, C. J., Klaus, C. R., Sacks, J. D., Raimondi, A., Majer, C. R., Song, J., Scott, M. P., Jin, L., Smith, J. J., Olhava, E. J., Chesworth, R., Moyer, M. P., Richon, V. M., Copeland, R. A., Keilhack, H., Pollock, R. M., and Kuntz, K. W. (2012) A selective inhibitor of EZH2 blocks H3K27 methylation and kills mutant lymphoma cells. Nat Chem Biol 8, 890-896

300. McCabe, M. T., Ott, H. M., Ganji, G., Korenchuk, S., Thompson, C., Van Aller, G. S., Liu, Y., Graves, A. P., Della Pietra, A., 3rd, Diaz, E., LaFrance, L. V., Mellinger, M., Duquenne, C., Tian, X., Kruger, R. G., McHugh, C. F., Brandt, M., Miller, W. H., Dhanak, D., Verma, S. K., Tummino, P. J., and Creasy, C. L. (2012) EZH2 inhibition as a therapeutic strategy for lymphoma with EZH2-activating mutations. Nature 492, 108-112

301. Smyth, G. K. (2005) Limma: Linear Models for Microarray Data, Bioinformatics and Computational Biology Solutions Using R and Bioconductor. Springer: New York

302. Tan, M., Luo, H., Lee, S., Jin, F., Yang, J. S., Montellier, E., Buchou, T., Cheng, Z., Rousseaux, S., Rajagopal, N., Lu, Z., Ye, Z., Zhu, Q., Wysocka, J., Ye, Y., Khochbin, S., Ren, B., and Zhao, Y. (2011) Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. Cell 146, 1016-1028

303. Shyh-Chang, N., Locasale, J. W., Lyssiotis, C. A., Zheng, Y., Teo, R. Y., Ratanasirintrawoot, S., Zhang, J., Onder, T., Unternaehrer, J. J., Zhu, H., Asara, J. M., Daley, G. Q., and Cantley, L. C. (2013) Influence of threonine metabolism on S-adenosylmethionine and histone methylation. Science 339, 222-226

304. Cervera, A. M., Bayley, J. P., Devilee, P., and McCreath, K. J. (2009) Inhibition of succinate dehydrogenase dysregulates histone modification in mammalian cells. Mol Cancer 8, 89

305. Letouze, E., Martinelli, C., Loriot, C., Burnichon, N., Abermil, N., Ottolenghi, C., Janin, M., Menara, M., Nguyen, A. T., Benit, P., Buffet, A., Marcaillou, C., Bertherat, J., Amar, L., Rustin, P., De Reynies, A., Gimenez-Roqueplo, A. P., and Favier, J. (2013) SDH mutations establish a hypermethylator phenotype in paraganglioma. Cancer Cell 23, 739-752

306. Koster, M. J., Snel, B., and Timmers, H. T. (2015) Genesis of chromatin and transcription dynamics in the origin of species. Cell 161, 724-736

307. Vaquerizas, J. M., Kummerfeld, S. K., Teichmann, S. A., and Luscombe, N. M. (2009) A census of human transcription factors: function, expression and evolution. Nat Rev Genet 10, 252-263

308. Lechleider, R. J., de Caestecker, M. P., Dehejia, A., Polymeropoulos, M. H., and Roberts, A. B. (1996) Serine phosphorylation, chromosomal localization, and transforming growth factor-beta signal transduction by human bsp-1. J Biol Chem 271, 17617-17620

309. Yingling, J. M., Das, P., Savage, C., Zhang, M., Padgett, R. W., and Wang, X. F. (1996) Mammalian dwarfins are phosphorylated in response to transforming growth factor beta and are implicated in control of cell growth. Proc Natl Acad Sci U S A 93, 8940-8944

310. Zhang, Y., Feng, X., We, R., and Derynck, R. (1996) Receptor-associated Mad homologues synergize as effectors of the TGF-beta response. Nature 383, 168-172

311. Xiao, Z., Watson, N., Rodriguez, C., and Lodish, H. F. (2001) Nucleocytoplasmic shuttling of Smad1

conferred by its nuclear localization and nuclear export signals. J Biol Chem 276, 39404-39410

312. Tsubakihara, Y., Hikita, A., Yamamoto, S., Matsushita, S., Matsushita, N., Oshima, Y., Miyazawa, K., and Imamura, T. (2015) Arkadia enhances BMP signalling through ubiquitylation and degradation of Smad6. J Biochem 158, 61-71

313. Lebrun, J. J., Takabe, K., Chen, Y., and Vale, W. (1999) Roles of pathway-specific and inhibitory Smads in activin receptor signaling. Mol Endocrinol 13, 15-23

314. Hjelmeland, M. D., Hjelmeland, A. B., Sathornsumetee, S., Reese, E. D., Herbstreith, M. H., Laping, N. J., Friedman, H. S., Bigner, D. D., Wang, X. F., and Rich, J. N. (2004) SB-431542, a small molecule transforming growth factor-beta-receptor antagonist, inhibits human glioma cell line proliferation and motility. Mol Cancer Ther 3, 737-745

315. Levy, L., Howell, M., Das, D., Harkin, S., Episkopou, V., and Hill, C. S. (2007) Arkadia activates Smad3/Smad4-dependent transcription by triggering signal-induced SnoN degradation. Mol Cell Biol 27, 6068-6083

316. Takahata, M., Inoue, Y., Tsuda, H., Imoto, I., Koinuma, D., Hayashi, M., Ichikura, T., Yamori, T., Nagasaki, K., Yoshida, M., Matsuoka, M., Morishita, K., Yuki, K., Hanyu, A., Miyazawa, K., Inazawa, J., Miyazono, K., and Imamura, T. (2009) SKI and MEL1 cooperate to inhibit transforming growth factor-beta signal in gastric cancer cells. J Biol Chem 284, 3334-3344

317. Varelas, X., Sakuma, R., Samavarchi-Tehrani, P., Peerani, R., Rao, B. M., Dembowy, J., Yaffe, M. B., Zandstra, P. W., and Wrana, J. L. (2008) TAZ controls Smad nucleocytoplasmic shuttling and regulates human embryonic stem-cell self-renewal. Nat Cell Biol 10, 837-848

318. Lee, J. E., Wang, C., Xu, S., Cho, Y. W., Wang, L., Feng, X., Baldridge, A., Sartorelli, V., Zhuang, L., Peng, W., and Ge, K. (2013) H3K4 mono- and di-methyltransferase MLL4 is required for enhancer activation during cell differentiation. Elife 2, e01503

319. Guo, X., and Wang, X. F. (2013) A P(E)RM(I)T for BMP signaling. Mol Cell 51, 1-2

320. Singh, A. M., Reynolds, D., Cliff, T., Ohtsuka, S., Mattheyses, A. L., Sun, Y., Menendez, L., Kulik, M., and Dalton, S. (2012) Signaling network crosstalk in human pluripotent cells: a Smad2/3-regulated switch that controls the balance between self-renewal and differentiation. Cell Stem Cell 10, 312-326

321. Alarcon, C., Zaromytidou, A. I., Xi, Q., Gao, S., Yu, J., Fujisawa, S., Barlas, A., Miller, A. N., Manova-Todorova, K., Macias, M. J., Sapkota, G., Pan, D., and Massague, J. (2009) Nuclear CDKs drive Smad transcriptional activation and turnover in BMP and TGF-beta pathways. Cell 139, 757-769

322. Gao, S., Alarcon, C., Sapkota, G., Rahman, S., Chen, P. Y., Goerner, N., Macias, M. J., Erdjument-Bromage, H., Tempst, P., and Massague, J. (2009) Ubiquitin ligase Nedd4L targets activated Smad2/3 to limit TGF-beta signaling. Mol Cell 36, 457-468

323. Isogaya, K., Koinuma, D., Tsutsumi, S., Saito, R. A., Miyazawa, K., Aburatani, H., and Miyazono, K. (2014) A Smad3 and TTF-1/NKX2-1 complex regulates Smad4-independent gene expression. Cell Res 24, 994-1008

324. Katagiri, T., and Watabe, T. (2016) Bone Morphogenetic Proteins. Cold Spring Harb Perspect Biol 8

325. Hartley, J. L., Temple, G. F., and Brasch, M. A. (2000) DNA cloning using in vitro site-specific recombination. Genome Res 10, 1788-1795

326. Nakao, A., Imamura, T., Souchelnytskyi, S., Kawabata, M., Ishisaki, A., Oeda, E., Tamaki, K., Hanai, J., Heldin, C. H., Miyazono, K., and ten Dijke, P. (1997) TGF-beta receptor-mediated signalling through Smad2, Smad3 and Smad4. EMBO J 16, 5353-5362

327. Chao, H. W., Lai, Y. T., Lu, Y. L., Lin, C. L., Mai, W., and Huang, Y. S. (2012) NMDAR signaling facilitates the IPO5-mediated nuclear import of CPEB3. Nucleic Acids Res 40, 8484-8498

328. Saijou, E., Itoh, T., Kim, K. W., Iemura, S., Natsume, T., and Miyajima, A. (2007) Nucleocytoplasmic shuttling of the zinc finger protein EZI Is mediated by importin-7-dependent nuclear import and CRM1-independent export mechanisms. J Biol Chem 282, 32327-32337

329. Pereira, L. A., van der Knaap, J. A., van den Boom, V., van den Heuvel, F. A., and Timmers, H. T. (2001) TAF(II)170 interacts with the concave surface of TATA-binding protein to inhibit its DNA binding activity. Mol Cell Biol 21, 7523-7534

330. Rappsilber, J., Ishihama, Y., and Mann, M. (2003) Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. Anal Chem 75, 663-670

331. R Core Team. (2015) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria

332. Vizcaino, J. A., Csordas, A., del-Toro, N., Dianes, J. A., Griss, J., Lavidas, I., Mayer, G., Perez-Riverol, Y.,

Reisinger, F., Ternent, T., Xu, Q. W., Wang, R., and Hermjakob, H. (2016) 2016 update of the PRIDE database and its related tools. Nucleic Acids Res 44, D447-456

333. Kawarada, Y., Inoue, Y., Kawasaki, F., Fukuura, K., Sato, K., Tanaka, T., Itoh, Y., and Hayashi, H. (2016) TGF-beta induces p53/Smads complex formation in the PAI-1 promoter to activate transcription. Sci Rep 6, 35483

334. Janknecht, R., Wells, N. J., and Hunter, T. (1998) TGF-beta-stimulated cooperation of smad proteins with the coactivators CBP/p300. Genes Dev 12, 2114-2119

335. Pearson, K. L., Hunter, T., and Janknecht, R. (1999) Activation of Smad1-mediated transcription by p300/CBP. Biochim Biophys Acta 1489, 354-364

336. Pouponnot, C., Jayaraman, L., and Massague, J. (1998) Physical and functional interaction of SMADs and p300/CBP. J Biol Chem 273, 22865-22868

337. Spedale, G., Timmers, H. T., and Pijnappel, W. W. (2012) ATAC-king the complexity of SAGA during evolution. Genes Dev 26, 527-541

338. Dhalluin, C., Carlson, J. E., Zeng, L., He, C., Aggarwal, A. K., and Zhou, M. M. (1999) Structure and ligand of a histone acetyltransferase bromodomain. Nature 399, 491-496

339. Lange, M., Kaynak, B., Forster, U. B., Tonjes, M., Fischer, J. J., Grimm, C., Schlesinger, J., Just, S., Dunkel, I., Krueger, T., Mebus, S., Lehrach, H., Lurz, R., Gobom, J., Rottbauer, W., Abdelilah-Seyfried, S., and Sperling, S. (2008) Regulation of muscle development by DPF3, a novel histone acetylation and methylation reader of the BAF chromatin remodeling complex. Genes Dev 22, 2370-2384

340. Sanchez, R., and Zhou, M. M. (2009) The role of human bromodomains in chromatin biology and gene transcription. Curr Opin Drug Discov Devel 12, 659-665

341. Patel, S. R., Kim, D., Levitan, I., and Dressler, G. R. (2007) The BRCT-domain containing protein PTIP links PAX2 to a histone H3, lysine 4 methyltransferase complex. Dev Cell 13, 580-592

342. Starnes, L. M., Su, D., Pikkupeura, L. M., Weinert, B. T., Santos, M. A., Mund, A., Soria, R., Cho, Y. W., Pozdnyakova, I., Kubec Hojfeldt, M., Vala, A., Yang, W., Lopez-Mendez, B., Lee, J. E., Peng, W., Yuan, J., Ge, K., Montoya, G., Nussenzweig, A., Choudhary, C., and Daniel, J. A. (2016) A PTIP-PA1 subcomplex promotes transcription for IgH class switching independently from the associated MLL3/MLL4 methyltransferase complex. Genes Dev 30, 149-163

343. Gong, Z., Cho, Y. W., Kim, J. E., Ge, K., and Chen, J. (2009) Accumulation of Pax2 transactivation domain interaction protein (PTIP) at sites of DNA breaks via RNF8-dependent pathway is required for cell survival after DNA damage. J Biol Chem 284, 7284-7293

344. Shimizu, K., Bourillot, P. Y., Nielsen, S. J., Zorn, A. M., and Gurdon, J. B. (2001) Swift is a novel BRCT domain coactivator of Smad2 in transforming growth factor beta signaling. Mol Cell Biol 21, 3901-3912

345. Baas, R., Sijm, A., van Teeffelen, H. A., van Es, R., Vos, H. R., and Marc Timmers, H. T. (2016) Quantitative Proteomics of the SMAD (Suppressor of Mothers against Decapentaplegic) Transcription Factor Family Identifies Importin 5 as a Bone Morphogenic Protein Receptor SMAD-specific Importin. J Biol Chem 291, 24121-24132

346. Baas, R., Lelieveld, D., van Teeffelen, H., Lijnzaad, P., Castelijns, B., van Schaik, F. M., Vermeulen, M., Egan, D. A., Timmers, H. T., and de Graaf, P. (2014) A novel microscopy-based high-throughput screening method to identify proteins that regulate global histone modification levels. J Biomol Screen 19, 287-296

347. Wang, C., Lee, J. E., Lai, B., Macfarlan, T. S., Xu, S., Zhuang, L., Liu, C., Peng, W., and Ge, K. (2016) Enhancer priming by H3K4 methyltransferase MLL4 controls cell fate transition. Proc Natl Acad Sci U S A 113, 11871-11876

348. Akhtar-Zaidi, B., Cowper-Sal-lari, R., Corradin, O., Saiakhova, A., Bartels, C. F., Balasubramanian, D., Myeroff, L., Lutterbaugh, J., Jarrar, A., Kalady, M. F., Willis, J., Moore, J. H., Tesar, P. J., Laframboise, T., Markowitz, S., Lupien, M., and Scacheri, P. C. (2012) Epigenomic enhancer profiling defines a signature of colon cancer. Science 336, 736-739

349. Kirmizis, A., Bartley, S. M., Kuzmichev, A., Margueron, R., Reinberg, D., Green, R., and Farnham, P. J. (2004) Silencing of human polycomb target genes is associated with methylation of histone H3 Lys 27. Genes Dev 18, 1592-1605

350. Fang, J., Chen, T., Chadwick, B., Li, E., and Zhang, Y. (2004) Ring1b-mediated H2A ubiquitination associates with inactive X chromosomes and is involved in initiation of X inactivation. J Biol Chem 279, 52812-52815

351. Scheuermann, J. C., de Ayala Alonso, A. G., Oktaba, K., Ly-Hartig, N., McGinty, R. K., Fraterman, S., Wilm, M., Muir, T. W., and Muller, J. (2010) Histone H2A deubiquitinase activity of the Polycomb repressive complex PR-DUB. Nature 465, 243-247

352. Yang, W. M., Inouye, C., Zeng, Y., Bearss, D., and Seto, E. (1996) Transcriptional repression by YY1 is

mediated by interaction with a mammalian homolog of the yeast global regulator RPD3. Proc Natl Acad Sci U S A 93, 12845-12850

353. Smits, A. H., Jansen, P. W., Poser, I., Hyman, A. A., and Vermeulen, M. (2013) Stoichiometry of chromatin-associated protein complexes revealed by label-free quantitative mass spectrometry-based proteomics. Nucleic Acids Res 41, e28

354. Henikoff, S., and Shilatifard, A. (2011) Histone modification: cause or cog? Trends Genet 27, 389-396

355. Ng, H. H., Robert, F., Young, R. A., and Struhl, K. (2003) Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. Mol Cell 11, 709-719

356. Nguyen, A. T., and Zhang, Y. (2011) The diverse functions of Dot1 and H3K79 methylation. Genes Dev 25, 1345-1358

357. Lacoste, N., Utley, R. T., Hunter, J. M., Poirier, G. G., and Cote, J. (2002) Disruptor of telomeric silencing-1 is a chromatin-specific histone H3 methyltransferase. J Biol Chem 277, 30421-30424

358. Ng, H. H., Feng, Q., Wang, H., Erdjument-Bromage, H., Tempst, P., Zhang, Y., and Struhl, K. (2002) Lysine methylation within the globular domain of histone H3 by Dot1 is important for telomeric silencing and Sir protein association. Genes Dev 16, 1518-1527

359. van Leeuwen, F., Gafken, P. R., and Gottschling, D. E. (2002) Dot1p modulates silencing in yeast by methylation of the nucleosome core. Cell 109, 745-756

360. Harmsen, M. M., and De Haard, H. J. (2007) Properties, production, and applications of camelid single-domain antibody fragments. Appl Microbiol Biotechnol 77, 13-22

361. Zhang, X., Wen, H., and Shi, X. (2012) Lysine methylation: beyond histones. Acta Biochim Biophys Sin (Shanghai) 44, 14-27

362. Qi, J., and Ronai, Z. A. (2015) Dysregulation of ubiquitin ligases in cancer. Drug Resist Updat 23, 1-11

363. Teixeira, L. K., and Reed, S. I. (2013) Ubiquitin ligases and cell cycle control. Annu Rev Biochem 82, 387-414

364. Blomen, V. A., Majek, P., Jae, L. T., Bigenzahn, J. W., Nieuwenhuis, J., Staring, J., Sacco, R., van Diemen, F. R., Olk, N., Stukalov, A., Marceau, C., Janssen, H., Carette, J. E., Bennett, K. L., Colinge, J., Superti-Furga, G., and Brummelkamp, T. R. (2015) Gene essentiality and synthetic lethality in haploid human cells. Science 350, 1092-1096

365. Seddon, A. M., Curnow, P., and Booth, P. J. (2004) Membrane proteins, lipids and detergents: not just a soap opera. Biochim Biophys Acta 1666, 105-117

366. Mattson, G., Conklin, E., Desai, S., Nielander, G., Savage, M. D., and Morgensen, S. (1993) A practical approach to crosslinking. Mol Biol Rep 17, 167-183

367. Mahipal, A., and Malafa, M. (2016) Importins and exportins as therapeutic targets in cancer. Pharmacol Ther 164, 135-143

368. Laping, N. J., Grygielko, E., Mathur, A., Butter, S., Bomberger, J., Tweed, C., Martin, W., Fornwald, J., Lehr, R., Harling, J., Gaster, L., Callahan, J. F., and Olson, B. A. (2002) Inhibition of transforming growth factor (TGF)-beta1-induced extracellular matrix with a novel inhibitor of the TGF-beta type I receptor kinase activity: SB-431542. Mol Pharmacol 62, 58-64

369. Sanvitale, C. E., Kerr, G., Chaikuad, A., Ramel, M. C., Mohedas, A. H., Reichert, S., Wang, Y., Triffitt, J. T., Cuny, G. D., Yu, P. B., Hill, C. S., and Bullock, A. N. (2013) A new class of small molecule inhibitor of BMP signaling. PLoS One 8, e62721

370. Herbertz, S., Sawyer, J. S., Stauber, A. J., Gueorguieva, I., Driscoll, K. E., Estrem, S. T., Cleverly, A. L., Desaiah, D., Guba, S. C., Benhadji, K. A., Slapak, C. A., and Lahn, M. M. (2015) Clinical development of galunisertib (LY2157299 monohydrate), a small molecule inhibitor of transforming growth factor-beta signaling pathway. Drug Des Devel Ther 9, 4479-4499

371. Neuzillet, C., Tijeras-Raballand, A., Cohen, R., Cros, J., Faivre, S., Raymond, E., and de Gramont, A. (2015) Targeting the TGFbeta pathway for cancer therapy. Pharmacol Ther 147, 22-31

372. Nervi, C., De Marinis, E., and Codacci-Pisanelli, G. (2015) Epigenetic treatment of solid tumours: a review of clinical trials. Clin Epigenetics 7, 127

373. de Lera, A. R., and Ganesan, A. (2016) Epigenetic polypharmacology: from combination therapy to multitargeted drugs. Clin Epigenetics 8, 105

374. Website: http://www.mun.ca/biology/desmid/brian/BIOL2060/BIOL2060-18/18_21.jpg.

List of abbreviations	
ac	acetylation
ACVR	Activin receptor
ALK	Activin receptor-like kinase
ALL	acute lymphoblastic leukemia
AMH	Anti-Müllerian Hormone
AML	acute myelogenous leukemia
ASXL	additional sex comb like
BAP1	BRCA1 associated protein 1
BMP	Bone Morphogenetic Protein
bp	base pair
CBP	CREB-binding protein
CBX	Chromobox homolog
COMPASS	complex of proteins associated with Set1
Co-SMAD	co-activator SMAD
CROSS	Chromatin Regulation Ontology SiRNA Screening
DOT1L	disruptor of telomeric silencing 1 like
DotCom	DOT1L H3K79 methyltransferase complex
dox	doxycycline
DSP	dithiobis(succinimidyl propionate)
DTT	dithiothreitol
DUB	deubiquitinase
EMT	epithelial to mesenchymal transition
EZH	enhancer of zeste homologue
FAD	flavin adenine dinucleotide
GAP	GTPase-activating protein
GDF	Growth Differentiation Factor
GEF	guanine exchange factors
GFP	green fluorescent protein
GPS	global proteomic analysis in *Saccharomyces cerevisiae*
GSTF	gene-specific transcription factor
GTF	general transcription factor
GWAS	genome-wide association studies
HAT	histone acetyltransferase
HDAC	histone deacetylases
hESC	human embryonic stem cell
HP1	heterochromatin protein 1
ID	inhibitor of differentiation
IFN-γ	interferon-γ
IPO	importin
I-SMAD	inhibitory SMAD

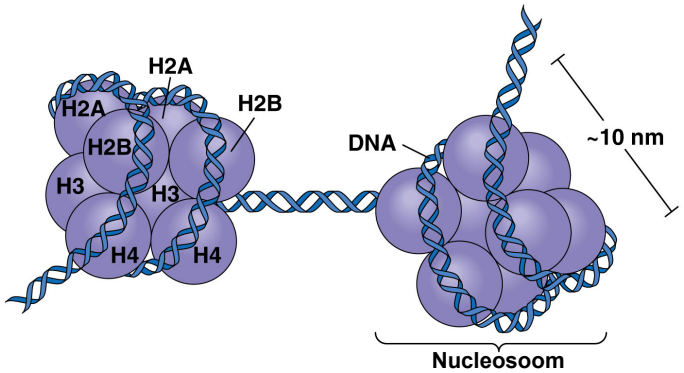
JAMM	jab1/mov34/mpn
JmjC	Jumonji
JMJD	Jumonji domain-containing protein
KAT	lysine acetyltransferase
KDM	lysine demethylase
KMT	methyltransferase
MAD	Mothers against decapentaplegic
MAPK	mitogen-activated protein kinases
me	methylation
MH	MAD homology
MiDAC	mitotic deacetylase
MJD	Machado-Joseph domain-containing protein
MLL	Mixed Lineage Leukemia
NES	nuclear export signal
NF-κB	nuclear factor κB
NLS	nuclear localization signal
NTF	nuclear transport factor
NuRD	Nucleosome Remodeling Deacetylase
OUT	otubain domain-containing protein
PAI-I	plasminogen activator inhibitor-1
Pc	Polycomb
PcG	Polycomb group
PCGF	Polycomb group RING finger protein
ph	phosphorylation
Ph	Polyhomeotic
PHD	plant homeodomain
PIC	pre-initiation complex
PRC	polycomb repressive complex
PR-DUB	polycomb repressive deubiquitinase complex
PRMT	protein arginine methyl transferases
Psc	Posterior sex combs
PtdIns(3)P	phosphatidylinositol 3-phosphate
PTM	post-translational modification
RanBP	Ran binding protein
RanGDP	GDP bound Ran
RanGTP	GTP bound Ran
RCC	regulator of chromosome condensation
RNA pol II	RNA polymerase II
R-SMAD	receptor activated SMAD
RT	room temperature
SAD	SMAD activating domain

SAGA	Spt-Ada-Gcn5 acetyltransferase
SAM	S-adenosyl-L-methionine
SARA	SMAD anchor for receptor activation
SBE	SMAD binding elements
SEC	super elongation complex
SET	Su (var), Enhancer of Zeste, and Trithorax
SMA	Small
SMAD	SMA and MAD related
TAF	TBP associated factor
TBP	TATA-binding protein
TFIID	transcription factor II D
TGFBR	TGF-β receptor
TGF-β	Transforming Growth Factor-β
Trr	Trithorax-related
Trx	Trithorax
TSO	TEAD/SMAD/OCT4
ub	ubiquitination
UCH	Ubiquitin C-terminal hydrolase
USP	Ubiquitin-specific protease
UTX	Ubiquitously-transcribed TPR protein on the X chromosome
WRAD	WRD5, RBBP5, ASH2L and DPY30

Nederlandse samenvatting voor niet-ingewijden

In dit proefschrift worden verschillende niveaus van controle mechanismen van de cel onderzocht en gebruiken we een nieuwe techniek om meer van deze mechanismen te kunnen identificeren. Begrip van transcriptie en chromatine regulatie mechanismen is onmisbaar en brengt ons een stap dichterbij het genereren van medicijnen voor het behandelen van cellen als deze mechanismen ontspoord zijn in ziektes zoals kanker.

In het menselijke lichaam zijn biljoenen cellen aanwezig, die elk een eigen functie vervullen. Een cel in de lever filtert bijvoorbeeld schadelijke stoffen uit het bloed, terwijl een neuron in de hersenen elektrische signalen door moet geven. Ook is het belangrijk dat een stamcel kan blijven delen, terwijl een gespecialiseerde cel zich daar juist niet mee bezig houdt. Aan de basis van al deze cellen ligt een bouwplan, het DNA, waarin de genetische informatie voor elke cel in het lichaam ligt opgeslagen. Op dit DNA ligt de informatie om functionele eenheden, eiwitten, te genereren. De verschillende cellen verspreid over de organen in het lichaam voeren elk een andere taak uit en hebben dus ook niet allemaal dezelfde stukjes van dit DNA nodig. Toch is alle informatie voor alle cellen in het DNA aanwezig. Als je het DNA van één cel in een rechte lijn achter elkaar zou leggen zou dit wel twee meter lang zijn. De volledige lengte DNA moet in de kern van de cel worden opgeborgen die zes duizendste van een millimeter groot is. Dit vereist een zeer efficiënte manier van opslag. Om het DNA zo goed op te kunnen slaan in de celkern ligt het opgerold om bolletjes van eiwitten (Figuur 1). Deze eiwitten noemen we histonen. Elk bolletje waar DNA omheen ligt bestaat uit twee kopieën van vier verschillende histonen: H2A, H2B, H3 en H4. Verder ligt er nog één histon eiwit bovenop het DNA die het DNA bij elkaar houdt: histon H1. Het geheel van de histon eiwitten samen met het omgewonden DNA vormt het nucleosoom. Alle nucleosomen en geassocieerde eiwitten samen noemen we chromatine. De histonen in het nucleosoom hebben staarten, die uit het nucleosoom steken. Deze staarten kunnen worden aangepast door bepaalde type eiwitten die een kleine chemische modificatie toevoegen (schrijven). Een andere klasse van eiwitten kan juist deze modificaties herkennen (lezen) of weghalen (wissen). Vaak zijn dit soort schrijvende, lezende en wissende eiwitten afhankelijk van andere eiwitten (cofactoren) om hun functie goed te kunnen uitvoeren en zitten ze samen in grotere eiwitcomplexen. Er zijn veel van dit soort eiwitcomplexen en het soort eiwitcomplex dat deze modificaties leest is afhankelijk van het soort en de plaats van



Figuur 1. Versimpelde weergave van twee nucleosomen. Vier paar histonen zijn omwikkeld door het DNA. Histon H1 en de staarten van de histonen zijn niet getekend. Figuur aangepast van (374).

de modificatie. Sommige complexen zorgen ervoor dat de nucleosomen dicht op elkaar worden gepakt waardoor het DNA niet meer te bereiken is. Andere complexen zorgen juist voor ruimte tussen deze nucleosomen waardoor het DNA wel af te lezen is voor de cel. Voorbeelden van zulke complexen die zorgen voor de toegankelijkheid van het DNA zijn de "Mixed Lineage Leukemia 3 en 4 (MLL3 en MLL4)" complexen. Op deze manier heeft de cel controle over de beschikbare stukjes DNA en is alleen het DNA beschikbaar wat van toepassing is voor het type cel. Dit is van belang voor het gezond functioneren van de cel.

In het menselijk lichaam is het essentieel dat een cel communiceert met nabij gelegen cellen. Dit kan een cel op een paar manieren doen. Als een cel moet communiceren met een buurman gebeurt dit via direct cel-cel contact. Maar als een cel een stuk verderop ligt, bijvoorbeeld in een ander orgaan, is dit geen optie. Cellen kunnen er dan voor kiezen om kleine moleculen in de bloedbaan te brengen die meeliften in de bloedsomloop. Voorbeelden van zulke moleculen zijn: "Transforming Growth Factor beta (TGF-β)" en "Bone Morphogenetic Protein (BMP)". Een stukje verderop kunnen deze moleculen dan worden opgevangen door andere cellen die een sensor (receptor) bevatten voor deze moleculen. In de ontvangende cel wordt dan een signaal door eiwitten doorgegeven, die uiteindelijk in de kern van de cel terecht moet komen. In het geval van TGF-β en BMP wordt dit gedaan door de signaaltransductie eiwitten "SMA and MAD related (SMADs)". De kern van een cel is een compartiment, dat grotendeels afgesloten is voor de rest van de cel. Kleine poriën zijn aanwezig aan de buitenkant van de kern, die toestaan dat eiwitten de kern binnen kunnen komen. Om de kern in te kunnen komen moet een signaaleiwit echter wel worden gebracht door een speciaal transporteiwit. Als het signaaleiwit vervolgens de kern bereikt kan het binden aan herkenningspunten in het DNA. Deze signaaleiwitten noemen we transcriptie factoren. Dit heeft dan tot gevolg dat eiwitcomplexen worden gerekruteerd, die dat stukje DNA vrij maken om te kunnen worden overgeschreven in boodschapper mRNA. Ook het fysiek afscheiden van de signaaleiwitten en het rekruteren van deze eiwitcomplexen zijn mechanismen van de cel om controle over de functie uit te kunnen oefenen.

De controle van al deze moleculaire processen zorgt voor balans in de cel. Deze balans kan soms worden verstoord, bijvoorbeeld door een signaal van een nabijgelegen cel. Bij voldoende controle kan de cel die het signaal ontvangt hierop reageren om vervolgens weer in balans te komen. Als de controle op deze processen echter verstoord is, kan dit leiden tot grote onbalans in de cel. Deze onbalans kan ervoor zorgen dat een cel ziek wordt en dood gaat. Een andere uitkomst is dat de cel opeens toegang heeft tot informatie van andere processen dan de bedoeling is. Als een cel opeens toegang krijgt tot de informatie om te kunnen gaan delen kan een cel dit ongecontroleerd gaan doen en zo tot een kankergezwel uitgroeien. Het is daarom van groot belang dat de controlesystemen op orde zijn. Om controle op deze systemen te kunnen uitvoeren kunnen we gebruik maken van medicijnen. Deze medicijnen oefenen invloed uit op de controlemechanismen door deze bijvoorbeeld te blokkeren of juist te activeren. Voordat we een medicijn kunnen ontwikkelen moeten we weten hoe deze controlesystemen werken.

In dit proefschrift worden verschillende controlemechanismen van de cel onderzocht. In hoofdstuk 2 beschrijven we een nieuwe methode die we "Chromatin Regulation Ontology SiRNA Screening" (CROSS) noemen. In deze techniek beschrijven we hoe we op zoek gaan naar schrijvers, wissers en cofactoren van histon modificaties. Dit doen we door systematisch genen één voor één uit te schakelen in menselijke cellijnen. Vervolgens meten we met een

volledig geautomatiseerde microscoop hoeveel van deze modificatie aanwezig is in deze cellen.

In hoofdstuk 3 gaan we op zoek naar eiwitten die binden aan een klasse signaaleiwitten met behulp van een massa spectrometer. Hiermee wordt de massa van een eiwit gedetecteerd en weet je welke eiwitten aanwezig zijn in het experiment. Eiwitten die aan elkaar binden trekken elkaar sterker aan dan wanneer ze geen interactie hebben. Door een eiwit van interesse te zuiveren uit een cel trek je deze interactiepartners mee. Door aan de SMAD signaaleiwitten te trekken hebben we IPO5 als transporteiwit voor de BMP subgroep van deze SMADs gevonden. Alleen de SMADs, die door BMP maar niet door TGF-β worden geactiveerd, worden door IPO5 naar de kern gebracht. Het onderscheid hierin wordt gemaakt door een volgorde aminozuren in deze SMADs.

In hoofdstuk 4 wordt vervolgens beschreven wat er gebeurt nadat de TGF-β geactiveerde SMADs naar de kern zijn gebracht en ze landen op het DNA. Met behulp van eiwit expressie proeven wordt aangetoond dat de SMADs een interactie aangaan met eiwitten van de MLL3 en MLL4 complexen. Door vervolgens te kijken naar de producten die ontstaan nadat DNA wordt afgelezen, wordt aangetoond dat deze complexen nodig zijn om het DNA af te lezen op deze plekken en blijkt dat het MLL4 complex hierbij cruciaal is.

Vervolgens onderzoeken we in addendum 1 de interactiepartners van UTX. Dit eiwit is aanwezig in het MLL3/4 complex maar ook in het "polycomb repressive deubiquitinase (PR-DUB)" complex. Door UTX te zuiveren wordt een derde complex gevonden waarin UTX zich bevindt. Dit complex is het "mitotic deacetylase (MiDAC)" complex.

In de discussie wordt het proefschrift afgesloten door deze bevindingen in biologische context te plaatsen.

Beste iedereen,

Aan het begin van mijn PhD traject leek dit boekje iets dat nog ver weg voor me lag. Maar de tijd vliegt, vooral als je een PhD doet, en voordat je het weet ben je het afsluitende woord aan het schrijven. Tijdens mijn PhD heb ik zowel op het lab als daarbuiten enorm veel kennis en ervaring opgedaan. Dit alles zou onmogelijk zijn geweest zonder mijn zeer gewaardeerde collega's, vrienden en familie en ik heb dan ook veel aan iedereen te danken.

Ik wil graag traditioneel beginnen met **Marc**. In 2012 ben ik bij jou begonnen in de groep en dankzij jou heb ik me kunnen ontwikkelen op wetenschappelijk vlak. Ik heb een grote waardering voor jouw enthousiasme voor de wetenschap die je graag op andere mensen over wil brengen. We waren het niet altijd met elkaar eens, maar uiteindelijk kwamen we er samen toch altijd weer uit. Je staat nu op het punt om met het lab te verhuizen naar de universiteit van Freiburg in Duitsland, een flinke onderneming maar tegelijkertijd ook een hele eer. En je hebt alvast geoefend met het verhuizen door twee jaar geleden naar het Hubrecht te gaan. Ik wens je heel veel succes en plezier in Freiburg en ik hoop dat er nog vele mooie wetenschappelijke jaren mogen volgen.

Petra, tijdens mijn PhD heb jij me veel begeleid. In het begin van mijn PhD traject was dit intensief toen we samen werkten aan het CROSS project (een paper waar ik nog steeds trots op ben). Later was dit iets minder frequent, toen je noodgedwongen het lab bij urologie op bent gaan zetten. We kwamen echter snel weer bij elkaar toen we bij toeval naar dezelfde verdieping in hetzelfde gebouw verhuisden. Ik kon altijd bij je terecht als ik advies nodig had, en daar heb ik veel aan gehad. Ik wens je het allerbeste toe en hoop dat er nog vele publicaties mogen volgen.

Michiel, jouw input tijdens de werkbespreking en later de supergroup meeting op het gebied van massaspectrometrie was voor mij van grote waarde. Tijdens mijn PhD ben je gegroeid van groepsleider naar professor en sta je aan het hoofd van een succesvolle groep in Nijmegen. Ik verwacht nog veel van je onderzoek te mogen horen en lezen.

SimonaC, my paranimf. Hello, it's me. When you arrived at the lab, things directly worked out between us. Your positive attitude always brightened up the mood of the group and I am certain you'll keep doing that in the future. It is very brave of you, that you'll be moving with the lab to Freiburg. Luckily you have **Derek** and **Aria** to support you. Keep going strong.

SimonA, you are an energetic person who gets her priorities straight. You are very creative and I enjoyed your humorous drawings on the whiteboard. All of these skills are invaluable for completing your PhD. Keep handing out those red and white spring bracelets and best of luck finishing. **Maria**, je bracht altijd sfeer in de groep met je scherpe opmerkingen en discussie. Na je reis naar Azië kwam je weer terug om verder te gaan als postdoc. Je kwam er snel achter dat dit toch niets voor je was en hebt de academie gewisseld voor een baan waarvan jij wist dat het bij je past. Ik wens je heel veel succes op deze nieuwe ingeslagen weg. **Silke**, je bent voortvarend begonnen aan je PhD traject bij Marc en kon het goed vinden in de groep. Het was een harde schok toen bleek dat het lab naar Duitsland ging verhuizen terwijl jij juist bewust had gekozen voor een plek in Nederland. Gelukkig heb je een nieuwe plek weten te vinden in de groep van **Jop Kind** en ik hoop dat je het hier net zoveel naar je zin zult hebben met goede collega's zoals **Corina** en **Franka**. Je enthousiasme is een goede

drive voor het halen van je PhD, zet hem op.

Hetty, tijdens mijn PhD ben jij onmisbaar voor mij geweest. Of je nou bezig was met het uitvoeren van mijn oneindig aantal IPs en immunoblots of je algemene taken voor het lab en de afdeling, je zette je altijd voor 100% in. Je hebt zelfs een belangrijke rol gespeeld in de verhuizing van het lab van het Stratenum naar het Hubrecht. Nu zet je datzelfde enthousiasme in voor je hobby fotografie. Je toonde altijd interesse op persoonlijk vlak en als je iets interessants voor mij had gezien in de krant legde je dat op mijn bureau, dat ga ik zeker missen. Blijf altijd positief en enthousiast, dat zijn twee mooie eigenschappen van je. **Richard**, jouw kennis van biochemie is ontzettend groot en tijdens mijn PhD heb je die aangevuld met proeven omtrent massaspectrometrie. Je bent een grote aanwinst voor het lab, maar nu het lab naar Duitsland vertrekt is de toekomst onzeker. Ik weet zeker dat je goed terecht zult komen met de kennis en kwaliteiten die jij bezit, onafhankelijk van je keuze. Heel veel succes deze keuze te maken.

Koen, ik heb respect voor hoe jij het werk als arts combineert met het werk in het lab. Ondanks dat je er niet altijd was wist je precies wat er in de groep speelde. Nu Marc naar Duitsland vertrekt hoop ik dat je het werk in het lab kan blijven voortzetten. **Elfi**, ook jij hebt gemerkt dat het werk in het lab heel anders is dan je werk als arts. Toch is het je gelukt om je in hoog tempo aan te passen. Ik vond het leuk te mogen rijden in je Audi tijdens de retraite in Schoorl. Nu de tijd in het lab voorbij is ben je weer aan de slag als arts. Ik wens je veel succes in het tussendoor afronden van je PhD en zie graag het boekje tegemoet. **Rick**, je vertelde me dat af en toe een schouderklopje nodig is in de wetenschap. Die positieve instelling heeft mij geholpen en probeer ik over te dragen op anderen. **Hugo**, je was altijd erg gedreven en het is je gelukt een eigen NMR groep te starten. **Andrée**, Je hebt gekozen voor een carrière buiten de academie, ik hoop dat je het daar naar je zin hebt. **Pim**, kort nadat ik ben begonnen ben jij je eigen groep gestart en we hebben elkaar dus maar kort meegemaakt. Heel veel succes met je groep in Rotterdam.

During my PhD I had the opportunity to teach a set of amazing students. **Bas**, je was mijn allereerste student die ik heb mogen begeleiden en daar heb ik enorm veel van geleerd. Als ik het JBS tijdschrift in mijn kast zie staan denk ik af en toe nog terug aan de avond in Florin toen ik je een fysieke kopie mocht geven. Heel veel succes met je PhD in het **Creyghton** lab! **Ayestha**, ik heb genoten van onze kleine wandelingen waarin we dingen als toekomst en carrière hebben besproken. Na je stage bij mij ben je aan de slag gegaan als PhD bij **Peter Verrijzer** in Rotterdam en ik heb er alle vertrouwen in dat je daar op je plek bent. Ik zie graag jouw boekje tegemoet! **Michele**, the project on TAF1 turned out to be challenging and somewhat plagued with bad luck. Nevertheless, I think you are a fine researcher and I hope that you, together with **Lise**, will find a place to work that suits you. During your internship you introduced me into the world of bouldering which I still very much enjoy. **Sjoerd**, ik heb je leren kennen toen je samen met **Anne-Wil** de stage van de MMK-cursus bij mij kwam lopen. Dat vond je zo leuk dat je maar gelijk bent blijven plakken. En dat was niet met het minste resultaat, ons manuscript over de SMAD target genen ligt bij de editor van Transcription. Veel succes bij het afronden van je studie.

Ik vond het altijd leuk om even langs te lopen bij het secretariaat, ook al was het om even te zwaaien. **Betty**, **Mirjam**, **Cristina** en **Marian**, hoe zou ik mijn PhD zijn doorgekomen zonder jullie? Altijd kon ik bij jullie terecht als er een meeting moest worden gepland en

jullie stonden altijd voor me klaar en keken voor me uit. **Betty**, je plotselinge vertrek bij de groep was een grote verrassing, gelukkig kwam je terecht op een nieuwe afdeling waar je het erg naar je zin hebt. Ik vond het leuk dat we het contact bleven houden en om het kopje thee met je te drinken op de dag dat ik mijn proefschrift af leverde bij de leescommissie. **Mirjam**, je had het altijd heel goed door als het even wat minder goed ging en het huppelen door de gang wist me dan altijd op te vrolijken. Ook toen we verhuisd waren naar het Hubrecht wist je de groep te ondersteunen en was je betrokken bij de groep.

De samenwerking met de Vermeulen groep was altijd erg gezellig en constructief, ook na de verhuizing naar Nijmegen. **Nelleke**, heel veel succes met het uitpluizen van NuRD. Je bent gedreven en ik hoop dat het je gaat lukken om een eigen groep te gaan starten. **Arne**, altijd heel vrolijk en handig met R. Je hebt me veel geleerd en ik maak er nog steeds gebruik van. Blijf vooral bezig met de bioinformatica. **Irem**, best of luck with your postdoc in Germany, I hope the place suits you and allows you to pursue what you want to do. **Susan**, thank you for teaching me how to perform GFP-affinity pulldowns. This technique is at the basis of my SMAD proteomics chapter and I still use the tricks you taught me. **Radhika**, I wish you the best establishing a new life in London. **Anneloes**, hoewel je het soms druk had, vond je het werk in het lab erg leuk. Na je vertrek uit het Vermeulen lab kwam ik je nog geregeld tegen, eerst bij celbiologie en later in het Hubrecht. Ik hoop dat je plezier uit de wetenschap blijft halen. **Pascal**, je bent altijd de drijvende kracht achter het Vermeulen lab geweest, ik hoop dat je dat nog lang mag doen. Furthermore, I'd like to thank and wish the best for **Lisa**, **Rik**, **Raghu**, **Xiaofei**, **Cristina**, **Matthew**, **Ino** and **Marijke**.

After the moving of the lab, I enjoyed the interaction with the members of the **Coffer lab**. **Cornelieke**, **Koen**, **Magdalena**, **Jorg**, **Sandra**, **Luca**, **Desiree**, **Enric**, **Catalina**, **Ana**, **Lucas**, **Bas**, **Janneke**, **Cindy** and **Guy**, thank you all for the nice time at the Hubrecht.

Frank, tijdens de afwezigheid van Marc in het begin heb je de begeleiding tijdelijk op je genomen. Verder wilde ik je bedanken dat je het hoofd van mijn leescommissie wilde zijn. Leden van de leescommissie: **Frank Holstege**, **Geert Kops**, **Edwin Cuppen**, **Paul Coffer** en **Peter ten Dijke**, bedankt voor het lezen en beoordelen van mijn proefschrift.

Klaas, tijdens de supergroup meeting heb ik jou en je lab zien ontwikkelen. Ik wens jou, **Jan**, **Georgina**, **Sabine**, **Jessie**, **Roderick** en **Mark** heel veel succes.

Harmjan, je was altijd nieuwsgierig en kritisch en dacht altijd mee als het over massaspectrometrie ging, ik heb veel van je geleerd. **Robert**, je hebt een oneindige hoeveelheid samples voor me geladen op de massaspectrometer. Totdat je vond dat ik het maar zelf moest doen, dank je wel dat je me dit hebt geleerd.

Vittoria, we met when the Saskia van Mil group moved to the Stratenum. Who would have thought that talking to a random person during the introductory drinks could result in a very good friendship? Thank you for sharing the nice moments during our PhDs, I really enjoyed those. I hope that you will be able to find a nice job back in Italy.

Philip, dank je wel voor de bioinformatica ondersteuning en de adviezen als ik even niet uit mijn R analyse kwam.

Wienand, waar een brandweercursus al dan niet goed voor is. Onze discussies over R-scripts, statistiek en bioinformatica waren zeer nuttig voor me. **David** and **Daphne**, thank you for allowing me to work with the liquid handling system, it was really cool to work with.

Huub, ik vond het altijd heel leuk om samen met jou onderwijs te geven. **Joost**, heel veel succes met het invullen van het onderwijs.

Livio, dank je wel voor je hulp als de microscoop het weer een keer af liet weten.

Marcel, tijdens mijn bachelor stage heb je me laten zien wat een PhD inhoudt en mede hierdoor ben ik met mijn PhD gestart. Ik vond het leuk je onverwachts weer tegen te komen bij de cursus Python die je toen gaf en kon het niet laten een klein beetje extra m'n best te doen.

Marjoleine en **Cheuk**, bedankt voor de ondersteuning bij al die bestellingen. Ik vond het leuk om af en toe bij jullie langs te lopen. Samen met **Marcel** zorgen jullie ervoor dat alle labs door kunnen blijven gaan. **Leny** en **Abdel**, dank jullie wel voor het schoonhouden van de afdeling, ook als de borrels weer eens een keer net te lang door zijn gegaan.

Heren van de ICT, **Wim**, **Marc**, **Erik**, **Dennis**, altijd kon ik even naar beneden lopen om hulp te vragen bij een probleem en jullie stonden altijd voor me klaar. **Marc**, enorm bedankt voor alle hulp die je me hebt geboden, zonder jou zou mijn massaspectrometriefiguur hopeloos verloren zijn gegaan.

Eirinn, **Dorien** and **Banafsheh**, organizing the PhD retreat together was an amazing experience!

Joost, ik heb er enorme bewondering voor hoe het je gelukt is een complete afdeling te verhuizen. Regenerative Medicine is met recht een mooie afdeling geworden.

Furthermore I'd like to thank all the members of the labs of **Frank Holstege**, **Saskia van Mil**, **Eric Kalkhoven**, **Hans Bos**, **Boudewijn Burgering**, **Tobias Dansen**, **Suzanne Lens**, **Johan de Rooij**, **Hugo Snippert**, **Geert Kops** and **David Egan**.

Dear new colleagues of **B8**, I'm looking forward to lots of new scientific breakthroughs.

Jan, **Rosina**, **Martijn** en **Clothaire**, ik heb genoten van onze Pathfinder avonturen en ik kijk uit naar de vele vervolgen die ongetwijfeld gaan komen en uiteraard kijk ik uit naar de dag dat **Leander** mee gaat doen. **Koen**, ook al zien we elkaar niet even regelmatig, het is altijd zo dat als we weer eens afspreken het net is alsof we elkaar gisteren nog hebben gezien. **Kevin**, we zagen elkaar standaard tijdens hardloopevenementen, blijf sportief bezig, ik hoop dat er nog vele mogen volgen.

Biologen **Jan**, **Nils**, **Kevin**, **Rick**, **Remon**, **Hugo**, **Bas** en **Nienke**, elke tweede paasdag komen we bij elkaar voor het traditionele potje klaverjassen. En ook al is **Hugo** er niet zo goed in, deze dag is altijd geslaagd! Ik hoop dat we nog veel boompjes mogen leggen.

Annette en **Valerie**, het is altijd gezellig als jullie langskomen en volgens mij vonden jullie het niet vervelend als ik weer eens een baksel uitprobeerde. **Valerie**, ik hoop nog vaak samen te gaan klimmen.

Mike en **Ditha**, samen met **Suus** was ik altijd bij jullie welkom om eventjes weg uit Nederland te zijn. En nog voordat we dan vertrokken werd alweer gevraagd of we dan binnenkort weer langs kwamen. We hebben mooie wandelingen gemaakt en de vakantie naar de Alpen was daar zeker de mooiste van, danke schön!

Piet, **Joke** en **Tom**, jullie hebben mij, en ik heb jullie, vlak voor mijn eindexamen op de middelbare school leren kennen. Jullie hebben me direct opgenomen en jullie hebben altijd als 2e familie voor me gevoeld. Bedankt voor jullie steun en interesse. **Joke**, je adviseerde me altijd dat ik goed op mezelf moest passen en dat is een advies dat ik meeneem. **Piet**, Ik heb genoten van de uitstapjes naar de TT in Assen samen met **Maarten** op de hoofdtribune en het gastvrije verblijf bij **Koert** en **Els**. **Tom**, binnenkort ga je op jezelf wonen, dat is een hele grote en belangrijke stap. Ik kom graag bij je langs en dan mag je me alles vertellen over de vertragingen van de treinen.

Opa Boere, u volgde Suus en mij altijd met veel interesse en was heel trots op ons. Helaas bent u er niet meer bij.

Lisa, je bent altijd bezig met organiseren en al ruim een jaar geleden vroeg je aan mij of je mijn paranimf mocht zijn. Daar had je namelijk over gelezen en dat leek je heel erg leuk om te doen. Omdat ik daar op dat moment nog niet uitgebreid over had nagedacht overviel je me daar behoorlijk mee. Nadat ik had besloten ben je ijverig aan de slag gegaan om het feest te organiseren en je hebt me veel werk uit handen genomen. Enorm bedankt. **Bas**, ik vond het altijd gezellig als je er was, die schoonfamilie is zo gek nog niet toch?

Oma, u vond het altijd leuk als ik in de weekenden in Vlaardingen was en ik vond het altijd heel fijn om u weer een knuffel te mogen geven. Door de jaren heen heeft u me altijd gesteund en vertroeteld. U was apetrots op me en dat motiveerde mij nog meer.

Papa en **mama**, elf jaar geleden nam ik op een boot in Schotland de beslissing om biologie te gaan studeren. Het plan om dieren in het veld te bestuderen is wel wat anders uitgepakt, maar jullie bleven me altijd steunen. En ook al was het soms moeilijk om uit te leggen wat ik precies allemaal deed op het lab, jullie bleven er altijd naar vragen. Ik vond het dan ook erg leuk om jullie het lab te laten zien. Dank jullie wel dat ik de mogelijkheid heb gekregen om te kunnen doen wat ik wilde doen. **Papa**, het is nu dr. Ouwe Beer. Dank je wel voor al die keren dat je me een lift hebt gegeven naar het station. **Mama**, ook jij vond het altijd fijn als ik tijdens de weekenden naar huis kwam. En je zorgde er ook voor dat ik nooit zonder eten weer terug naar Utrecht ging. Dank je wel voor al die goede zorgen.

Lieve **Suus**, de afgelopen jaren heb jij van dichtbij meegemaakt hoe ik mijn PhD traject heb doorlopen. Wat er ook gebeurde, je stond altijd voor me klaar. Ik heb veel steun aan je gehad op de momenten dat ik dat nodig had en dat terwijl je zelf ook bezig bent met je gz-opleiding. Dank je wel dat jij er voor me bent.

Roy

Curriculum vitae

Roy Baas werd geboren op 8 augustus 1988 in Vlaardingen. In 2006 haalde hij zijn vwo-diploma met het profiel 'natuur en gezondheid' aan Het College Vos in Vlaardingen. Datzelfde jaar begon hij zijn bacheloropleiding biologie aan de Universiteit Leiden. Tijdens zijn bacheloropleiding heeft hij een onderzoeksstage gedaan in de Plant Cell Physiology groep onder begeleiding van dr. ing. Marcel van Verk. Tijdens deze stage heeft hij gewerkt aan AtWRKY eiwit-eiwit en eiwit-promotor interacties. In het derde jaar van zijn bacheloropleiding heeft hij meegedaan met het Honours Class Program 2009: From Solar to Fuel with Bio. Tijdens deze Honours Class heeft hij een onderzoeksvoorstel geschreven die hij mocht presenteren tijdens de This Week's Discoveries aan de Universiteit Leiden. Nadat hij in 2009 zijn bachelor heeft afgerond is hij zijn masteropleiding Life Sciences – Molecular Biosciences begonnen aan de Universiteit van Amsterdam. Tijdens zijn masteropleiding heeft hij een onderzoeksstage gedaan bij de Mass Spectrometry groep onder begeleiding van dr. Luitzen de Jong. Tijdens deze stage heeft hij gewerkt aan het cross-linken van eiwitten en de detectie hiervan met behulp van massaspectrometrie. Als tweede stage heeft hij gewerkt in de Molecular Cytology groep onder begeleiding van dr. ir. Joachim Goedhart. Tijdens deze stage heeft hij confocale microscopie gebruikt voor het bestuderen van Gq signaaltransductie in HeLa cellen. Na het afronden van zijn masteropleiding is Roy in 2012 begonnen als onderzoeker in opleiding in de onderzoeksgroep van prof. dr. Marc Timmers op de afdeling Molecular Cancer Research en Regenerative Medicine van het Universitair Medisch Centrum Utrecht. Tijdens zijn promotietraject is hij lid geweest van de Cancer Genomics & Developmental Biology PhD retraite commissie. De resultaten van zijn onderzoek staan beschreven in dit proefschrift. Tijdens het afronden van zijn PhD is Roy begonnen als postdoc in het Nederlands Kanker Instituut in de groep van prof. dr. Titia Sixma.

List of publications

Baas, R., Sijm, A., van Teeffelen, H. A., van Es, R., Vos, H. R., and Marc Timmers, H. T. (2016) Quantitative Proteomics of the SMAD (Suppressor of Mothers against Decapentaplegic) Transcription Factor Family Identifies Importin 5 as a Bone Morphogenic Protein Receptor SMAD-specific Importin. *J Biol Chem* 291, 24121-24132 doi: 10.1074/jbc.M116.748582

Baas, R., Lelieveld, D., van Teeffelen, H., Lijnzaad, P., Castelijns, B., van Schaik, F. M., Vermeulen, M., Egan, D. A., Timmers, H. T., and de Graaf, P. (2014) A novel microscopy-based high-throughput screening method to identify proteins that regulate global histone modification levels. *J Biomol Screen* 19, 287-296 doi: 10.1177/1087057113515024

Schram, A. W., **Baas, R.**, Jansen, P. W., Riss, A., Tora, L., Vermeulen, M., and Timmers, H. T. (2013) A dual role for SAGA-associated factor 29 (SGF29) in ER stress survival by coordination of both histone H3 acetylation and histone H3 lysine-4 trimethylation. *PLoS One* 8, e70035 doi: 10.1371/journal.pone.0070035