Institut für Immunologie

Direktor: Prof. Dr. Hans-Jürgen Thiesen

Transcription Factor Properties of ZNF746 and ZNF777 with Focus on their Domains of Unknown Function 3669 (DUF3669)

Dissertation
zur
Erlangung des akademischen Grades
doctor rerum humanarum: “Dr. rer. hum.”
der Universitätsmedizin Rostock

vorgelegt von
M.Sc. Mohannad Al Chiblak,
geboren am 08.01.1982,
in Daraa, Syrien
Rostock, 06.10.20

https://doi.org/10.18453/rosdok_id00002800

Gutachter

Prof. Dr. med. Markus Tiedge, Universitätsmedizin Rostock,
Institut für Medizinische Biochemie und Molekularbiologie.

Prof. Dr. med. Hans-Jürgen Thiesen, Universitätsmedizin Rostock
Institut für Immunologie.

Prof. Dr. med. Christian Sina, Universitätsklinikum schleswig-holstein
Institut für Ernährungsmedizin.

Datum der Einreichung: 27.01.2020
Datum der Verteidigung: 06.10.2020

Table of contents

1- Summary

Krüppel-associated box-containing zinc finger proteins (KRAB-ZNFs) represent a large family of transcription regulators in tetrapods. Typically, KRAB-ZNF proteins consist of an amino-terminal protein interaction domain, designated KRAB, and a carboxy-terminal tandem array of DNA-binding C2H2 zinc finger motifs as initially determined in KOX1/ZNF10. In the canonical model of function, zinc finger motifs specifically bind to genomic target sites while the KRAB domain interacts with TRIM28 that in turn recruits chromatin-modifying protein complexes including enzymes such as histone methyltransferase SETDB1 and histone deacetylase HDAC1. The KRAB/TRIM28-mediated interaction initiates a local spreading of repressive chromatin structures promoting epigenetic silencing. A subset of KRAB-ZNFs has been described to contain N-terminally an additional conserved protein domain designated "domain of unknown function 3669" (DUF3669). The human genome encodes six human KRAB-ZNF proteins with DUF3669 at their N-termini. One of them is ZNF746/PARIS, which was involved in Parkinson's disease, and ZNF777, which is known as one of the oldest evolutionary family members of the KRAB-ZNF. The aim of this study was to evaluate transcriptional repression functions of DUF3669-KRAB amino-terminal domain arrays in ZNF746 and ZNF777 in comparison to the canonical KRAB/TRIM28 pathway. Further, the aim of this study was to answer the question as to whether DUF3669 can mediate protein oligomerization.

The ZNF746 major isoform (ZNF746a) is characterized by a 15-aa truncated KRAB-A subdomain. The existence of a second isoform (ZNF746b) containing a non-truncated KRAB domain was proven by RT-PCR. Heterologous reporter assays showed that the most efficient downregulation of the reporter gene was achieved by the entire amino-terminal array of a ZNF746 isoform corresponding to DUF3669, KRAB domain variant, and further sequence encoded by exons 5 and 6. Yet, these repression factors were still moderate as compared to the potent KRAB domain of ZNF10. Interestingly, the isolated DUF3669 domain of ZNF746 and ZNF777 showed intrinsic weak transcriptional repression activities. Most importantly, the DUF3669 region was sufficient to confer direct physical interaction among the members of the subfamily in oligomers, which might explain the contribution of DUF3669 to the repression activity. The ability of DUF3669 to modulate the repression activities in DUF3669-KRAB segments of ZNF746 and ZNF777 was found to be context-dependent since the heterologous transfer of DUF3669 to the potent KRAB domain of ZNF10 did not augment the transcriptional repression activity of the latter. Using mutants, the contribution of the

protein segment encoded by exons 5 and 6 to repression was ascribed to a potential SUMOylation event at K189 (in ZNF746a). Neither ZNF746a nor ZNF777 protein segments stably associated with TRIM28 within cells. Conversely, ZNF746b segments that contain the non-truncated KRAB-A subdomain were found in protein complexes containing TRIM28. Importantly, all ZNF746b segments showed substantially higher repression activities as compared to their counterparts of ZNF746a and their repression activities significantly dropped in HAP1 knockout cell models depleted of TRIM28 or SETDB1. The gained repression activities in the protein segments with non-truncated KRAB-A occur mainly through the canonical mechanisms known for the KRAB domain. The fact that the amino-terminal segments of ZNF746 could maintain some repression activities in TRIM28 knockout cells suggests the existence of TRIM28-independent regulatory mechanisms. The moderate repression activity of the full KRAB domain of ZNF746b compared to the ZNF10 KRAB domain turned out to be predominantly due to the absence of a glutamic acid residue that is highly conserved in the canonical KRAB-A consensus within the so-called “MLE”-motif. In agreement with the canonical model of KRAB function, the replacement of glycine residue with the glutamic acid residue in ZNF746b-KRAB concomitantly increased complex formation with TRIM28 as well. Apparently, N-terminal domain configurations including DUF3669 and controlled by SUMOylation events appear to determine protein complex formation of DUF3669-containing KRAB-ZNF proteins and to mediate transcriptional repression. The observed oligomerization property of DUF3669 is proposed to facilitate the forming of regulatory protein assemblies with other DUF3669 containing KRAB ZNF proteins. In addition, occurring SUMOylation might confer an additional docking surface to stabilize complex formation with TRIM28 or directly attract other chromatin modifiers such as SETDB1 or HDAC. Additional research activities are necessary to confirm whether endogenous genomic targets might be regulated in a comparable fashion as initially shown by the reporter assays. Moreover, the DUF3669 oligomerization function determined in this study underlines that KRAB-ZNF proteins acquired highly sophisticated regulatory properties by employing additional protein domains. The conservation of SUMO-receptor lysines in other DUF3669-containing zinc finger proteins suggests common intrinsic regulatory mechanisms for functions in distinct cellular processes from cell cycle regulation to cellular differentiation. This comprehensive study presents an entry to the phylogeny of DUF3669 containing KRAB-ZNF proteins and might be supportive in answering how KRAB mediated functionalities have been evolved from sea urchin to human.

The major part of this study was published in:

Al Chiblak, M., Steinbeck, F., Thiesen, H. J., & Lorenz, P. (2019). DUF3669, a “domain of unknown function” within ZNF746 and ZNF777, oligomerizes and contributes to transcriptional repression. *BMC Molecular and Cell Biology*, *20*(1), 60.

2- Introduction

2.1- Regulation of gene expression

Many mechanisms orchestrate gene expression to maintain the basic biological and metabolic function of an organism. Eukaryotic genes are transcribed under the control of promoter, enhancer, and silencer elements. While promoters reside near the 5' end of the gene, enhancers and silencers are located down- or upstream from their target genes (Maston et al., 2006; Chatterjee and Ahituv, 2017) and in intronic sequences as well (Deng et al., 2017; Finkbeiner, 2001). Sequence-specific transcription factors can bind to promoters, enhancers (activators), or silencers (suppressors) and modulate the recruitment of RNA polymerase II to the promoter and thereby regulate the rate of gene transcription (Martinez, 2002). Regulation of gene expression occurs either by direct targeting of the basal transcriptional machinery or by modulating the DNA-packaging structure (chromatin). It is thought that activators can recruit the transcription machinery to the promoter or enhance conformational changes in the complex. The suppressors may prevent the transcription by competing with transcription machinery components in binding to the promoter (Chatterjee and Ahituv, 2017) and/or through chromatin marks (Dillon et al., 2005).

In eukaryotes, DNA is packaged into chromosomes that are made up of repetitive units called "nucleosomes". Each nucleosome is made up of an octamer core of histones (two dimers of H3: H4, and two dimers of H2A: H2B) (Campos and Reinberg, 2009). Histone amino-terminal tails are highly conserved and heavily prone to post-transitional modifications (PTMs) that can lead to reversible changes in nucleosome assembly and DNA accessibility (Bannister and Kouzarides, 2011). For example, histone lysine acetylation is usually associated with transcriptional activation, presumably by weakening the electrostatic interactions between lysine residues and the phosphate backbone of DNA (Allfrey et al., 1964; Hebbes et al., 1988). Histone arginine methylation serves as either a transcriptional activating mark or transcriptional repressive mark dependent on the targeted arginine residue (Di Lorenzo and Bedford, 2011). Histone lysine methylation is catalyzed predominantly by SETdomain–containing methyltransferase family which targets multiple lysine residues such as; H1K26, H3K4, H3K9, H3K27, H3K36, and H4K20 resulting in transcriptional activation or inhibition depending on targeted residue (Dillon et al., 2005). H3K9me forms a high-affinity mark for the recruitment of heterochromatin protein 1 (HP1), thus, facilitating the deposition of HP1

proteins and making the target promoter inaccessible to the transcription machinery (Schultz et al., 2001, 2002).

2.2- C2H2 zinc finger transcription factors

The C2H2 “Krüppel”-class zinc finger (ZNF) family represents the largest group of transcription factors in human (Vaquerizas et al., 2009). Already back in 1990 with the description of the C2H2 KOX gene family, the expression of about 70 C2H2 zinc finger genes was estimated to occur in human T cells (Thiesen, 1990). Most ZNF proteins are transcriptional regulatory proteins that are involved in important cellular processes, such as proliferation, development, differentiation, speciation, and cancer pathology. ZNF proteins often contain tandem repeats of about 30 amino acids long finger-like motifs connected by short polypeptide linkers (Thiesen, 1990; Wolfe et al., 2000).

The zinc finger motif is folded in a short ßß-hairpin-alpha helix compact domain. Zinc cation stabilizes this structure through direct interaction with two Cys at one end of the ß-sheet and two His in the C-terminal part of the alpha helix (Thiesen, 1990; Thiesen and Bach, 1990, 1991a).

The amino acids of the α-helix are the major determinants for DNA binding and its specificity (Thiesen and Bach, 1991b; Thiesen and Schröder, 1991) supporting the notion of a zinc finger recognition code (Thiesen, 1990). Each zinc finger can bind three to four nucleotides within the major groove of DNA; the amino acids in the positions +6, +3, and -1 in the alpha-helix bind first, second, and third nucleotides at 5’ end, respectively. The amino acid +2 binds the fourth nucleotide in the complementary strand (Urnov et al., 2010; Gutierrez-Guerrero et al., 2018).

The diversity of amino acid sequences of zinc finger, the number of zinc finger motifs, spacing of zinc finger motifs, and higher-order structure increase the ability of zinc finger domain to bind a wide range of ligands that are not only DNA sequences, but also DNA-RNA hybrids, lipids, and proteins (Eom et al., 2016; Fedotova et al., 2017). Currently, designer zinc finger is used in genome editing as engineered zinc finger nucleases (Urnov et al., 2010; Gutierrez-Guerrero et al., 2018)

2.3- Krüppel associated box (KRAB)–containing zinc finger protein family

2.3.1- Discovery

The KRAB domain was initially discovered in December 1988 as heptad repeat of leucines in resemblance to the protein interaction domain found in leucine zipper structures as mentioned in the Annual Report 1988 of the Basel Institute of Immunology and published by Thiesen in 1990 and by Thiesen and Meyer in 1993. Sequences encoding the Krüppel-associated box (KRAB) domain are present in more than 400 of at least 742 human ZNF genes (Huntley et al., 2006). The earliest description in the literature of what was later coined KRAB domain was as N-terminal "heptad repeat of leucines" after analyzing KOX1/ZNF10 in 30 non-overlapping cDNAs of cKox1 to cKox30 isolated from human T-cells (Thiesen, 1990). Shortly thereafter, the name "KRAB domain" was coined referring to a highly conserved 75aa domain located at N-termini of numerous ZNF proteins from various species (Bellefroid et al., 1991; Thiesen et al., 1991). Importantly, several subsequent studies revealed the transcriptional silencing capacity of KRAB domains when they were tethered to DNA via a DNA binding domain in heterologous reporter assays (Margolin et al., 1994; Pengue et al., 1994; Witzgall et al., 1994; Deuschle et al., 1995; Pengue and Lania, 1996).

2.3.2- Genomic organization and protein domain configuration

Typically, the KRAB domain is made up of canonical subdomain A with or without an additional C-terminal subdomain. The most prevalent class of the KRAB-ZNF proteins harbors N-terminal KRAB-A and KRAB-B subdomains and C-terminal ZNF motifs (Fig. 1; (Huntley et al., 2006)). It was reported that KRAB-A confers repression activity while the KRAB-B enhances the KRAB-A mediated transcriptional repression through an unknown mechanism (Vissing et al., 1995). In addition to KRAB-A and KRAB-B subdomains, other subdomains have been characterized; KRAB-b which is a highly divergent KRAB-B (Mark et al., 1999), KRAB-C (Looman et al., 2004) and KRAB-BL (about10 amino acids longer, extending towards the C-terminal direction) (Huntley et al., 2006). Corresponding the evolution of KRAB-A, the KRAB-B coding-exon diverged into KRAB- b in some KRAB ZNF genes and might have been completely lost in others (Looman et al., 2002). While the KRAB subdomains are encoded by individual exons separated by introns with variable sizes, zinc finger domains containing multiple zinc fingers are encoded by a single exon (Fig. 1 (Bellefroid et al., 1991, 1993)). Most of the human KRAB-ZNF genes are concentrated in 25

major genomic clusters, the largest cluster resides on human chromosome 19 (Rousseau-Merck et al., 2002; Huntley et al., 2006). The multi-exon configuration of KRAB-ZNFs proteins led to the acquisition of new protein domains; such as SCAN and DUF3669 domains These accessory domains are found in the amino termini of several highly evolutionarily conserved KRAB-ZNFs (Ecco et al., 2017).

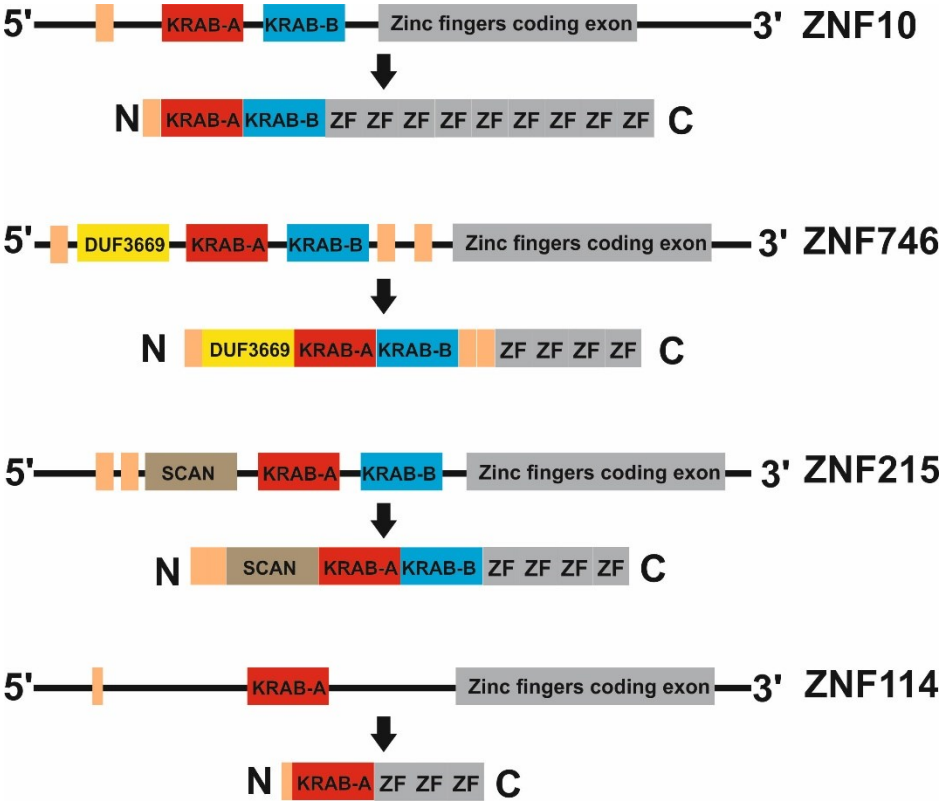

Fig. 1: Examples of gene models and primary protein structures of some KRAB-ZNF proteins represent different subclasses of KRAB-ZNF protein family; *KRAB AB-ZFs (e.g. ZNF10), DUF3669-KRAB AB-ZFs (e.g. ZNF746), SCAN-KRAB AB-ZFs (e.g. ZNF215), and KRAB A-ZFs (eg. ZNF114).*

2.3.3- Molecular mechanisms of canonical KRAB-mediated transcriptional repression

It is widely believed that the interaction between the KRAB domain and TRIM28 is obligatory and direct to function as a transcriptional repressor (Deuschle et al., 1995; Lupo et al., 2013). TRIM28 has initially been visualized as silencing –mediated protein (SMP1) and as a binding partner of the KRAB domain of Kox1 by (Deuschle et al., 1995). This interaction occurs directly through the N-terminal RBCC motif of TRIM28 (Fig. 2) (Friedman et al., 1996; Peng et al., 2000a; Peng et al., 2000b). Recent data showed that TRIM28 forms antiparallel dimers via the coiled-coil domain which primers higher-order assembly of TRIM28 dimers. KRAB domain interacts with TRIM28 in a stoichiometry of KRAB 1:2

RBCC on the dyad of the RBCC dimer that is required for transcriptional repression. The higher oligomerization status of TRIM28 did not affect its transcriptional silencing potential (Stoll et al., 2019).

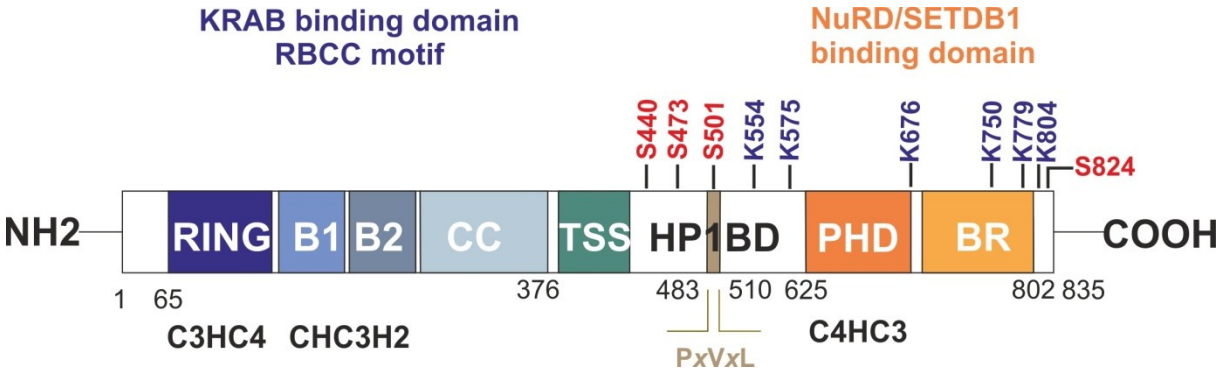

***Fig. 2**: **TRIM28 from Homo sapiens; primary structure and post-translational modifications**. The amino-terminal RBCC motif (RING, B1box, B2box, and coiled-coil) is responsible for the interaction with the KRAB domain. The amino-terminal PHD-BR is responsible for recruiting SETDB1 and NuRD complex to the target genomic locus. HP1-BD is responsible for binding heterochromatin protein 1 (HP1) through its binding site (PxVxL). TRIM28 is SUMOylated at the indicated lysines (K) and phosphorylated at the indicated serines (S), adapted from (Czerwińska et al., 2017).*

The C-terminal Plant Homeodomain (PHD) finger and bromodomain (referred collectively as PHD-BD) are the responsible parts for gene silencing. Isolated PHD-BD served as a repressor when it was tethered to DNA using a heterologous DNA-binding domain (H. Peng et al., 2000). PHD-BD binds CHD3/Mi-2alpha subunit of NuRD/HDAC complex (Schultz et al., 2001) and histone methyltransferase SETDB1 enzyme (Schultz et al., 2002). In addition, a central domain in TRIM28 called HP1BD (HP1 protein binding domain) containing a core PxVxL motif binds directly to the chromoshadow domain of the HP1 protein (Lechner et al., 2000).

SETDB1 (SET domain bifurcated 1) contains SET domain interrupted by an evolutionarily conserved 347-amino-acid insertion to form two functionally distinct subdomains (Harte et al., 1999; Schultz et al., 2002). SETDB1 specifically methylates the lysine 9 in the N-terminal tail of histone 3 (H3K9me), while inactive toward histone H2A, H2B and H4 (Schultz et al., 2002; Yang et al., 2002) to form a high-affinity mark for the recruitment of HP1 protein which facilitates the deposition of HP1 proteins (Schultz et al., 2002, 2001). ATF7IP (also known as MCAF1or hAM) is a SETDB1 interacting protein in humans and promotes the conversion of H3K9me2 mediated by SETDB1 to H3K9me3 *in vivo* (Wang et al., 2013).

It has been reported that the PHD domain in TRIM28 serves as intramolecular E3 SUMO ligase SUMOylating its cognate bromodomain. This SUMOylation event turns to be essential for recruiting SETDB1 and CHD3/NuRD complex, see Fig. 3. The recruitment of HP1-α to

TRIM28 was shown not to be affected by TRIM28 SUMOylation status. The SUMO Interacting Motifs (SIM) of SETDB1 and CHD3 have been identified; 122-IIEI-125 and 1995-VICI-1998 respectively (Fig. 3) (Ivanov et al., 2007; Mascle et al., 2007).

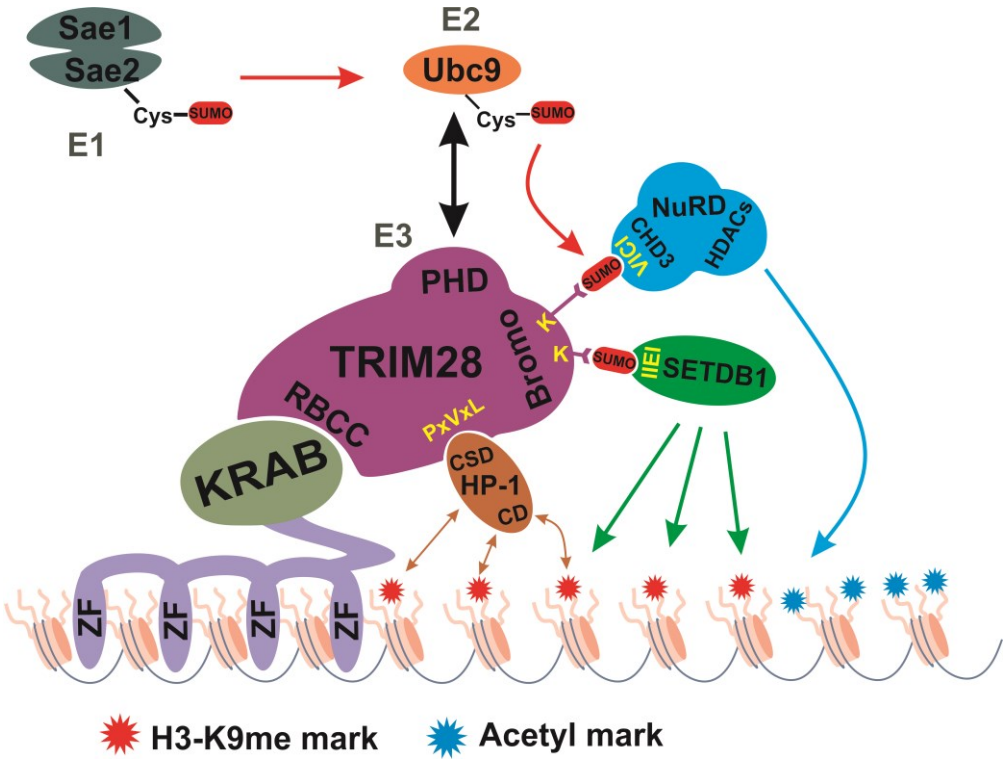

Fig. 3: Model for KRAB-TRIM28 mediated transcriptional repression*. The KRAB-ZNF protein binds specifically to a chromatin locus via its zinc finger (ZF) motifs. The KRAB domain interacts with RBCC motif of the nuclear multimodular hub protein TRIM28. TRIM28 recruits chromatin-modifying complexes via its carboxyl repressive part. Auto-SUMOylation is apparently essential for TRIM28-mediated transcriptional repression. The SUMOylated sites are docking interfaces for CHD3, a component of NuRD complexes with its associated activities like histone deacetylases (HDACs), and the histone methyltransferase SETDB. The HDACs remove histone acetyl marks while SETDB1 methylates H3K9 residues. Heterochromatin protein HP1 recognizes H3K9me marks placed by SETDB1 and triggers heterochromatization at targeted locus making it non-permissive for transcription; adapted from (Ivanov et al., 2007; Mascle et al., 2007) with modifications.*

Heterochromatin protein-1 is made up of two distinct evolutionary conserved domains; amino-terminal chromodomain (CD) and carboxy-terminal chromoshadow domain (CSD), in addition to the non-conserved central region that possesses DNA binding properties (hinge region). HP1 binds H3K9me3 marks with high affinity (Bannister et al., 2001).

It has been reported as well that CSD forms a strong homodimer and interacts with TRIM28 in the stoichiometry of 2:1 (Lechner et al., 2000). HP1 plays an essential role in the KRAB/TRIM28 repression model. CSDs form a strong dimer, while CDs form a weak tetramer. Two CDs recruit two H3K9me3 marks on the surface of the nucleosome and the other two CDs are still unoccupied to form bridges by recruiting additional H3K9me3 marks

on the adjacent nucleosome. In this way, HP1 proteins spread along the chromatin fibers and these spreads are directed by H3K9me3 marks (Canzio et al., 2011).

A gene-trapping method suggested that KRAB-TRIM28 transcriptional silencing machinery can repress genes that are located far away from their primary docking sites (Groner et al., 2010). How KRAB-ZNFs select their genomic targets along this distance for transcriptional repression and how a silenced locus can be converted back to a chromatin configuration that is again permissive for transcription, these topics are still ill-defined. Meylan and coauthors demonstrated that genes most susceptible to KRAB/TRIM28-mediated transcriptional repression are located in genomic regions of high gene activity. Furthermore, the pre-existence of H4K20me3 and H3K9me3 marks at the promoter and the surrounding region are assumed to make the promoters more susceptible to KRAB/TRIM28-mediated transcriptional repression (Meylan et al., 2011). In addition, KRAB/TRIM28 machinery seems to auto-regulate the KRAB-ZNF genes as well. As such, the KRAB-TRIM28 binding sites are located predominantly near 3' end and within the transcribed region of KRAB-ZNF genes. This binding mediates the spread of heterochromatin along the chromatin body (Groner et al., 2010; O'Geen et al., 2007).

2.3.4- The biological function of KRAB-ZNF proteins

It was recently proposed that tandem zinc finger genes originally evolved to repress the transcription of transposable elements (TE). It is like an ''arms race'' against these fossilized genetic invaders. In several vertebrate genomes, a striking correlation between the number of TE elements and ZNF genes has been observed. Additionally, it was hypothesized that there was a selective pressure to preserve KRAB-ZNFs-TEs pairs after the expansion of these TEs for millions of years. (Imbeault et al., 2017; Thomas and Schneider, 2011). The majority of KRAB-ZNF-binding sites of the human genome are located within TEs, and one-third of KRAB-ZNFs bind other genomic targets such as promoters or DNA repeats (Imbeault et al., 2017). Over the course of evolution, the number of zinc finger motifs per protein increased. These added motifs could be functional or non-functional motifs (Looman et al., 2002).

Due to their numerical abundance, the KRAB-ZNF proteins have the potential to form a complex regulatory network that controls the transcription of a wide category of genes and hereby diverse biological processes (Fig. 4). Although the identification of the endogenous targets of all KRAB-ZNFs is far from being clearly assessed, a lot of evidences have been accumulated by now, indicating the importance of KRAB-ZNF proteins and KRAB/TRIM28 in several physiological processes (Ecco et al., 2017; Lupo et al., 2013). For example,

KRAB/TRIM28 controls the retroelements during early embryogenesis. Upon DNA specific binding of KRAB-ZNF/TRIM28 and subsequent trimethylation of H3K9, specific DNA methylation of endogenous retroviral sequences occurs. These DNA methylation events are thought to encounter the indiscriminate genome-wide DNA methylation and permanently silence thousands of retroelements during development without the need for constant KRAB/TRIM28 recruitment (Quenneville et al., 2012; Rowe et al., 2013). KRAB/TRIM28 regulatory pathways are also implicated in metabolic control; for example, human ZNF224 represses the human aldolase A gene in KRAB/TRIM28-dependent manner and recruits arginine methyltransferase PRMT5 to the promoter that regulates the glycolysis pathway (Cesaro et al., 2009). Many KRAB-ZNFs are involved in the development of the adaptive immune system in humans (Santoni de Sio, 2014). Few studies pointed out the role of KRAB-ZNF proteins in neurodegenerative diseases. ZNF179 is a brain-specific protein and the inhibition of ZNF179 expression downregulates neuronal differentiation in P19 embryonal carcinoma cells and cerebellar granule cells (Pao et al., 2011). ZNF224 has been shown to influence intermediate phenotypes associated with Alzheimer's disease (Shulman et al., 2010), ZNF746 is implicated in Parkinson's disease pathology (see section 2.5.1).

Overexpression of several KRAB-ZNF proteins has been linked to oncogenesis; for example, the analysis of RNAseq profiling datasets of 16 different cancer types in The Cancer Genome Atlas (TCGA) project revealed a subset of 16 KRAB-ZNFs that are shown to be commonly upregulated (Machnik et al., 2019). Despite a large number of studies linking KRAB-ZNF proteins with biological and pathological pathways, the underlying molecular mechanisms are still largely unknown.

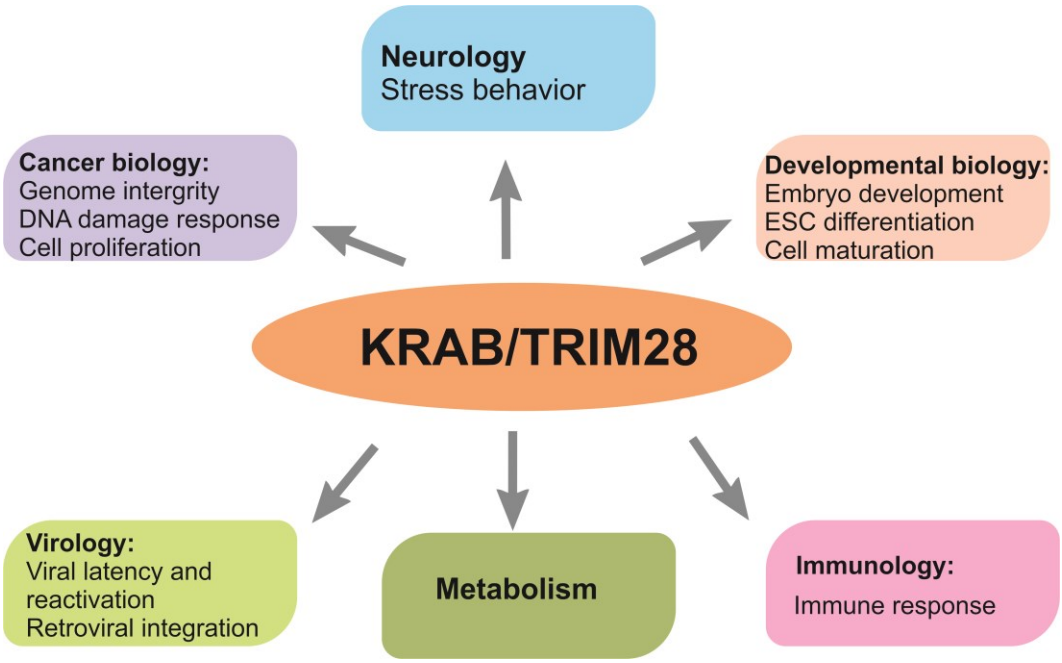

Fig. 4: Functions of the KRAB-ZNF/TRIM28 system or KRAB-ZNF in cellular physiology based on scientific literature.

2.4- The domain of Unknown function 3669 (DUF3669)

Conserved protein domains that have as yet unknown activities and biological properties are grouped in the Pfam database into subgroups of "domains of unknown function" using the prefix "DUF" followed by a number (a full list of DUFs can be found at http://pfam.xfam.org/family/browse?browse=d and https://www.uniprot.org/docs/upflist) (Mudgal et al., 2015). DUF3669 (protein domain identifier; PF12417) is described in InterPro database (https://www.ebi.ac.uk/interpro/beta/entry/InterPro/IPR022137/) as a 64 -80 amino acids domain. DUF3669-containing KRBA-ZNFs form a subfamily of six members that reside in the same cytogenetic location (7q36.1) in the human genome, which suggests that these genes have arisen from a common ancestral gene through internal duplication events.

It has been postulated that DUF3669-containing KRAB-ZNFs more likely bind to promotor regions (Imbeault et al., 2017). Note, all human DUF3669-containing KRAB-ZNFs are conserved between human and opossum. By excluding human SCAN-containing KRAB-ZNF genes, the genome-wide analysis showed that only three human KRAB-ZNF genes have orthologous proteins in nonmammalian amniote groups. These genes are DUF3669-containing KRAB-ZNFs (ZNF282, ZNF777, and ZNF783). In addition, multiple alignments of these orthologs revealed that their KRAB-A boxes lack the conserved LE residues within MLE sequence that are required for TRIM28 recruitment suggesting that DUF3669-containing KRAB-ZNFs might not be capable of TRIM28 binding. Furthermore, all human

DUF3669-containing KRAB-ZNFs are clustered in chr7; cytogenetic band 7q36.1(Liu et al., 2014).

Low homology sequences covering all metazoan taxa and fungi have been included to define DUF3669 consensus. The resulting domain model looks rather degenerated with a low number of strongly conserved amino acid residues just at the borders (https://www.ncbi.nlm.nih.gov/Structure/cdd/pfam12417).

2.5- Zinc finger protein 746 (ZNF746/PARIS)

Human ZNF746 is also known as PARIS (Parkin interacting substrate). Its protein configuration is characterized by an N-terminal DUF3669 followed by a KRAB domain with KRAB-A and KRAB-B subdomains and a C-terminal array of four zinc fingers (Fig. 10). The KRAB-A subdomain of the major ZNF746 isoform (designated ZNF746a throughout this thesis) has a 15aa truncation at its amino terminus. The repressive activity of ZNF746 is regulated by SUMOylation: *In vivo* and *in vitro* SUMOylation assays showed that ZNF746 is SUMOylated primarily at K189 and K286 (Nishida and Yamada, 2016). ZNF746 interacts specifically with Protein Inhibitor of activated STATy (PIASy) that functions as SUMO ligase E3 (Nishida and Yamada, 2016). Reporter assays showed that SUMOylation plays an essential role in the repression of PGC-1α promoter in a cell-dependent manner (Nishida and Yamada, 2016).

2.5.1- The role of ZNF746/PARIS in Parkinson's disease

Several studies pointed out the role of ZNF746 in mitochondrial biogenesis and glucose metabolism in dopaminergic neurons. ZNF746 overexpression has a deleterious effect on dopaminergic neurons and is a part of the physiological condition that regulates mitochondrial homeostasis, controlling both degradation and biogenesis.

ZNF746 binds the Insulin response sequence (IRS) region in the promoter of the peroxisome proliferator-activated receptor-gamma (PPARγ) coactivator-1a (PGC-1α) gene and represses the expression of PGC-1α in human cells, mouse brain, and the human brain. (Shin et al., 2011; Stevens et al., 2015). The downregulation of PGC-1α gene resulting from overexpression of ZNF746 in adult animal mice led to selective loss of dopaminergic neurons in the substantia nigra (Shin et al., 2011).

The levels of ZNF746 are regulated by the parkin ubiquitin-proteasome system through binding to and ubiquitination by the E3 ubiquitin ligase Parkin and subjected thus to

proteasomal degradation. The C-terminus of ZNF746 (Zinc finger motifs) binds to either the RING1 or RING2 domain of Parkin. Thus, PARKIN-ZNF746 seems to be an integral unit to maintain the mitochondrial homeostasis (Shin et al., 2011). It was reported that PTEN-induced putative kinase 1 (PINK1) directly phosphorylates serines 322 and 613 of ZNF746 and this phosphorylation is necessary for PARKIN-mediated ubiquitination of ZNF746 and thereby proteasomal degradation of ZNF746 (Fig. 5) (Lee et al., 2017). A reduction of mitochondrial number and abnormal appearing cristae in a significant fraction of mitochondria in adult conditional Parkin knockout mice. The knockdown of ZNF746 could substantially and significantly reverse and prevent these effects (Stevens et al., 2015).

In addition to PGC1α, Transketolase (abbrev. TKT; a key enzyme of pentose phosphate pathway glucose metabolism in all organisms) has been identified as an endogenous target of ZNF746. The overexpression of ZNF746 led to a significant reduction of TKT promoter activity in human cell lines that was reversed by parkin co-overexpression. The protein and mRNA levels of transketolase (TKT) specifically decreased in the dopaminergic neurons of the substantia nigra in mice but not in the cortical neurons of the cortex of AAV-ZNF746 mice. It was also reported that TKT mRNA levels decreased in the substantia nigra of Parkinson's disease patients. ZNF746 overexpression suppresses TKT transcription via the IRS-like motif in the TKT promoter (Kim et al., 2017). The ZNF746-mediated inhibition of the pentose phosphate pathway by repression of TKT resulted in the induction of HIF-1α that activates the glycolysis pathway in the substantia nigra and cortex in mice (Kang et al., 2018).

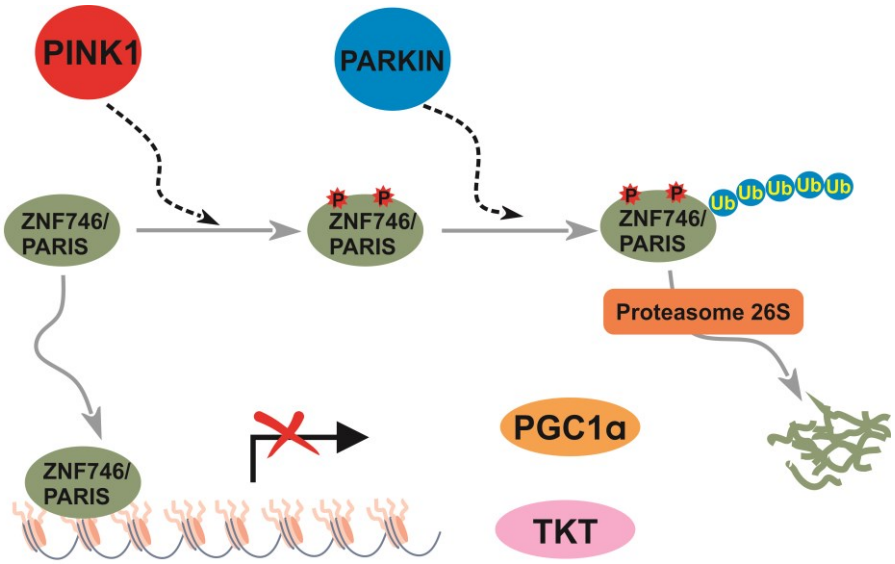

Fig. 5: ZNF746 plays a harmful role in dopaminergic neurons in the substantia nigra, and its levels are regulated by PARKIN-mediated ubiquitination*. ZNF746 downregulates the expression of transketolase and PGC-1α in dopaminergic neurons of the substantia nigra. Aberrant accumulation of ZNF746 results in impairment of mitochondrial biogenesis and energy metabolism leading to dopaminergic neurons death and development of Parkinson's disease. PINK-1 interacts directly with ZNF746 and phosphorylates it at S322 and S613. The phosphorylation events promote the interaction of ZNF746 with PARKIN E3 ubiquitin ligase. The ubiquitinated ZNF746 is eliminated by proteasomal degradation.*

2.5.2- The role of ZNF746/PARIS in cancer

ZNF746 is overexpressed in human colorectal cancer cells and colorectal cancer tissues. The overexpression of ZNF746 was correlated with reduced cellular survival. Accordingly, the depletion of ZNF746 suppressed the growth of colorectal cancer, lung, and breast cancer cells. ZNF746 interacts with the proto-oncogene protein c-Myc and co-localizes with c-Myc in the nucleus of human colon cancer cells (HCT116) (Jung et al., 2018). It was reported that the silencing of ZNF746 in H460 non-small cell lung cancer cells (H46 NSCLC) leads to inhibition of matrix metalloproteinase; MMP1, MMP2, and MMP9 that are known to promote tumor survival, invasion as well as metastasis (Kim et al., 2014). Chen and coauthors provided evidence on the role of ZNF746 in bladder cancer *in vivo* and *in vitro*. It was shown that the silencing of ZNF746 in bladder cancer cells resulted in significant suppression of their growth and invasion. In addition, the silencing of ZNF746 led to the downregulation of Protein kinase B (also known as AKT), and Matrix metalloproteinase 9 (MMP-9) (Y.-T. Chen et al., 2019).

2.6- Zinc finger protein 777 (ZNF777)

ZNF777 is a paralog of ZNF746 and shares a similar primary structure (DUF3669-KRAB AB- Zinc finger motifs), see Fig. 10. In comparison to ZNF746, the physiological functions and transcriptional repressive potential of DUF3669-KRAB-ZNF protein ZNF777 are far less understood. Quantitative RT-PCR and immunohistochemistry staining showed that ZNF777 is highly expressed in the human placenta (Liu et al., 2014). ZNF777 is localized to the nucleus through its zinc finger domain and shown to be responsible for inhibition of cell proliferation at low density (Yuki et al., 2015). The expression of ZNF777 resulted in decreasing levels of FAM129A (Family with sequence similarity 129 member A) at mRNA and protein levels which can partly be involved in inhibition of cell proliferation mediated by ZNF777 especially at low cell density (Yuki et al., 2015).

2.7. Haploid HAP1 cells – an efficient knockout model to study cellular functions

With the advent of CRISPR/Cas9-based genome engineering strategies, genes and gene segments can be removed or replaced. However, in diploid mammalian cells, the limitation of the mutagenesis methods is that the resulting mutants are usually heterozygotes, which can mask recessive phenotypes. The desired phenotype of a loss of function mutation on one chromosome is not displayed unless the sister chromatid is also mutated. This limitation is circumvented if haploid or nearly haploid cells are engineered utilizing CRISPR-Cas9 strategies.

HAP1 cells are a near-haploid cell line derived from the KBM-7 cell line that could be stably grown over at least three months. This cell line is characterized by the haploidy of all chromosomes with the exception of Chr.8 and a 30-megabases fragment of the Chr.15 (Kotecki et al., 1999). The haploidy property was harnessed to perform large-scale loss of function screens in human cells(Carette et al., 2009). Later on, the HAP1 cell line was generated with a haploidy of Chr.8. Such cells are not fully haploid, however, since they contain two copies of a portion of Chr.15, one of which is incorporated into the Chr.19 (Carette et al., 2011). This fragment was deleted from the Chr.19 in order to gain a fully haploid cell line using the CRISPR / Cas9 method (Essletzbichler et al., 2014).

Because HAP1 cells contain only one allele of each gene, they have the advantage of fast functional gene identification. Therefore, the HAP1 cells have been applied across a wide range of biological processes, such as DNA damage repair (Xing and Oksenych, 2019), drug sensitivity (Gerhards et al., 2018), and immune response (Keskitalo et al., 2019).

It has been observed that haploid HAP1 cells '' diploidized'' after several passaging circles through endoreduplication events (Essletzbichler et al., 2014), the loss of haploid cells is due to their limited viability. Thus, they replace themselves by an overgrowth of existing diploid cells in the cultures (Olbrich et al., 2017). Here, HAP1 cells deleted either of TRIM28 or SETDB1are employed and have been compared to wild-type HAP1.

3- Aim

This study focuses on the DUF3669-containing KRAB-ZNF protein subfamily. Two human members of this subfamily, ZNF746 and ZNF777, have been selected for a detailed study. Because of the clinical significance of ZNF746, better understanding of molecular mechanisms of action might open the door on new diagnostic and therapeutic aspects. ZNF777 belongs to the oldest evolutionarily conserved KRAB-ZNFs.

The aim of this work was to investigate the impact of the amino-terminal domain array of two DUF3669-containing KRAB-ZNF proteins; ZNF746 and ZNF777 on their transcription factor properties. In addition, this study purposed to evaluate the contribution of the canonical KRAB/TRIM28 pathway to the ZNF746- and ZNF777-mediated repression activities.

The experimental approaches in this study relied on classical DNA cloning methods, luciferase reporter assay, co-immunoprecipitation, and affinity pull-down assay.

4- Materials and methods

4.1- DNA techniques and cloning

4.1.1- RT-PCR (Reverse transcription PCR)

The coding sequences for several protein regions of ZNF746 and ZNF777 were synthesized by RT-PCR from RNA derived from the human fetal brain (Clontech #64094-1, 1mg/ml; pooled from 24 male/female Caucasians) and human testis (Clontech #64101-1, 1mg/ml; pooled from 45 Caucasians age 19-64) unless otherwise stated. High Capacity cDNA Reverse Transcription Kit with RNA inhibitors (Applied Biosystems, Cat. No. 4374966) was used.

Table 1: Preparation of reaction mixture for one reaction.

Components	Volume for 1x reaction
10x RT Buffer	2µl
25x dNTP Mix	0.8 µl
10x RT Random Primers	2µl
MultiScribe Reverse Transcriptase	1µl
RNase Inhibitor	1µl
molecular biology grade nuclease-free H2O	3.2 µl

The reaction mixtures were vortexed, quickly centrifuged at 4° C, and mixed with 10µl of 0.1 µg/µl total RNA solution (by pipetting). The reaction mixtures were incubated in the thermocycler at; 25°C for 10 min / 37°C for 120min / 85°C for 5min / 4°∞.

4.1.2- Specific PCR (DNA amplification)

Enzyme PfuUltra II Hotstart DNA polymerase (Stratagene/Agilent Cat. No. 600670) was used for PCR, together with the proprietary Pfu Ultra II reaction buffer and a 10mM dNTP mix (2.5mM each of deoxyadenosine/deoxyguanosine/ deoxycytidine/deoxythymidine triphosphate).

***Table 2**: Pipetting plan of PCR using PfuUltra II HS Stratagene.*

Component	Needed volume
dNTPs (stock 2.5mM)	6 µl
10 x PfuUltra II buffer	5 µl
PfuUltraII Fusion HS DNA polymerase (2.5 U/µl)	2 µl
cDNA (template)	Equiv. 200ng RNA or 10ng of plasmid
*Primers (Forward and reverse)	0.5µl+0.5µl
Nuclease-free H2O	Up to 50µl

* Primers are listed in **Table 4**

***Table 3**: Example on PCR cycle conditions*

Cycles	Temp [°C]	t
1	95	2 min
5	95	30sec
	62	30sec
	72	1 min
30	95	30 sec
	65	2 min
	72	1 min
End	72	5 min
	4	∞

4.1.3- Purification of PCR products

PCR amplicons with flanking restriction sites were purified via MinElute Reaction Clean up kit (50), Cat. No.28204;

- 300µl buffer ERC were added to the enzymatic reaction and mixed.
- The samples were applied to the MinElute column and centrifuged for 1 min/16060 x *g*.
- 750µl buffer PE were added to the MinElute column and centrifuged for 1 min/16060 x *g*.
- Flow-through was discarded and the column was centrifuged again for 1 min/16060 x *g*.
- Additional 1 min centrifugation step was done to ensure the removal of residual ethanol from buffer PE.
- In a clean 1.5 ml micro centrifuge tube, DNA was eluted by adding 10µl water to the center of the MinElute membrane and centrifuging 1 min at 16060 x *g*.
- The purified PCR amplicons were size-controlled on an agarose gel. (16060 x *g* = 13000 rpm; Heraeus biofuge fresco).

***Table 4: List of primers used for PCR.** Lowercase letters; overhangs that contain restriction sites, uppercase letters; coding sequence. The positions of the oligonucleotides in ZNF746 and ZNF777 exons are indicated relative to the nucleotide position; ZNF746 (isoform a: NM_152557.4 and isoform b: NM_001363517.1) and ZNF777 (NM_015694.3).*

Name	Position	PCR primers
Z746a/94-173	nt 551-790	5′-ccgctcgagtTGGATCCTGCGGCTGCC-3′ 3′-acgcgtcgactaTGGACTGGGGTCCACAGGAA-5′
Z746a/1-279	nt 272-1108	5′-ccgctcgagtATGGCCGAGGCGGTCG-3′ 3′-acgcgtcgactaTTCCGTGGAGGCTGCTGT-5′
Z746a/1-173	nt 272-790	5′-ccgctcgagtATGGCCGAGGCGGTCG-3′ 3′-acgcgtcgactaTGGACTGGGGTCCACAGGAA-5′
Z746a/94-279	nt 551-1108	5′-ccgctcgagtTGGATCCTGCGGCTGCC-3′ 3′-acgcgtcgactaTTCCGTGGAGGCTGCTGT-5′
Z746a	nt 272-2206	5′-ccgctcgagtATGGCCGAGGCGGTCG-3′

		3′-acgcgtcgacTCACATGTCCCCGCCAT-5′
Z746a/174-279	nt 791-1108	5′-ccgctcgagtGGCTCGGGGCC-3′ 3′-acgcgtcgactaTTCCGTGGAGGCTGCTGT-5′
Z746a/280-644	nt 1109-2206	5′-ccgctcgagtATGgATGTAAAAATTGTAATAAAAACAGA-3′ 3′-cgcgtcgacTCACATGTCCCCGCCAT-5′
Z746a/1-108	nt 272-595	5′-ccgctcgagtATGGCCGAGGCGGTCG-3′ 3′-acgcgtcgactaCTTAGGGGACTCCCCCTTGC-5′
Z746b/109-188	nt 598-837	5′- ccgctcgagtGTGCCCGTGACCTTTGATGATG -3′ 3′-acgcgtcgactaTGGACTGGGGTCCACAGGAA-5′
Z746b/1-294	nt 274-1155	5′-ccgctcgagtATGGCCGAGGCGGTCG-3′ 3′-acgcgtcgactaTTCCGTGGAGGCTGCTGT-5′
Z746b/1-188	nt 274-837	5′-ccgctcgagtATGGCCGAGGCGGTCG-3′ 3′-acgcgtcgactaTGGACTGGGGTCCACAGGAA-5′
Z746b/109-294	nt 598-1155	5′- ccgctcgagtGTGCCCGTGACCTTTGATGATG -3′ 3′-acgcgtcgactaTTCCGTGGAGGCTGCTGT-5′
Z777/283-362	nt 1110-1350	5′-ccgctcgagtGTCCCTGTCACATTTGATGATGT-3′ 3′-acgcgtcgactaAGCACTGGGATCTGTCGG-5′
Z777/1-362	nt 264-1350	5′-ccgctcgagtATGGAGAACCAACGCTCATC-3′ 3′-acgcgtcgactaAGCACTGGGATCTGTCGG-5′
Z777/1-282	nt 264-1109	5′-ccgctcgagtATGGAGAACCAACGCTCATC-3′ 3′-acgcgtcgactaCTTGGGAACTTCTCCATTGCT-5′
Z777/189-254	nt 828-1025	5′-ccgctcgagtTGGGCTGCCGTCCAA-3′ 3′-acgcgtcgactaCCGCCTCTGCAGCAGC-5′

4.1.4- Ligation

The T4 DNA ligase catalyzes the repair of single-stranded nicks in double-stranded DNA and joins double-stranded DNA restriction fragments having either blunt or sticky ends. The ligations were carried out for 1hr at RT using a molar ratio of 1:3 vector to insert.

DNA fragments were size-controlled by agarose gel electrophoresis and inserted into pKS restricted with *EcoR*V (pBluescript II KS (+) GenBank X52327, Stratagene) using blunt-end ligation protocol, and into expression vectors were performed using sticky end ligation protocol.

***Table 5**: Blunt- and Sticky-end ligation protocol.*

Solution	Composition	Blunt-end ligation	Sticky-end ligation
T4 DNA Ligase Fermentas, #EL0331	5U/µl	5units = 1 µl	5 units = 1 µl
10x T4 DNA Ligase Buffer Thermo Scientific # B69	400mM Tris-HCl,100mM $MgCl_2$, 100mM DTT 5mM ATP, pH 7.8	2 µl	2 µl
50% PEG solution	50% (w/v) polyethylene glycol 4000	2 µl	-
Insert	Purified PCR product	1:3 molar ratio of insert over vector DNA	1:3 molar ratio of insert over vector DNA
H2O	RNase free H2O	Up to 20 µl	Up to 20 µl

4.1.5- Competent bacteria and Transformation

4.1.5.1- Media

***Table 6**: Composition of media for cultivation of bacteria*

Medium Name	Composition
TY-Medium (Tryptone-Yeast medium)	For 1L: 15g NaCl, 10g Yeast extract, 16g Casein
LB-Agar (Lysogeny broth medium)	For 1L: 10g NaCl, 5g Yeast extract, 10g Casein, 15g Agar-agar

***Table 7**: selective antibiotic agents*

Antibiotic	Concentration in TY Medium	Concentration on LB Plates
Ampicillin	100µg/ml	300µg/ml
Tetracycline	10µg/ml	30µg/ml
Kanamycin	50µg/ml	150µg/ml

4.1.5.2- Preparation of Competent Bacteria

To make the bacteria permeable to DNA, a high concentration of Ca^{2+} ions was used to create pores in the bacterial cell membrane. *E.coli* XL1-Blue strain (Stratagene Catalog #200249; tetracycline resistant) was used for transformations and Blue-White screening. XL1-Blue cells are endonuclease (endA) deficient, which greatly improves the quality of mini-preparation of plasmid DNA, and are recombination (recA) deficient to improve the stability of the insertion. The hsdR mutation prevents the cleavage of cloned DNA by the *Eco*K endonuclease system. The lacIq ZΔM15 gene on the F′ episome allows the blue-white color screening.

Genotype: recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F′ proAB lacIqZΔM15 Tn10 (Tetr)].

A single colony was picked from the LB agar plate and placed in 2ml TY medium supplemented with tetracycline (end concentration 10µg/ml). The bacteria were grown by overnight incubation in a shaker at 200 rpm and 37°C. Next day, the bacterial culture was expanded to 100ml by adding fresh TY–medium supplemented with 10µg/ml tetracycline in a Schott flask and incubated in a shaker (37°C, 200rpm) to an OD_{600nm} ~ 0.3. The bacteria suspension was split into 2x 50ml falcon tubes and centrifuged for 10min at 2777x *g* (4000 rpm; Heraeus, Megafuge® 1.0R). The supernatants were discarded, and the pellets were resuspended with 12ml cold 0.1M CaCl2. The tubes were incubated on ice for 10 min and then centrifuged again for 10min at 2777 x *g*, 4°C. The supernatants were discarded, and the pellet in one falcon was resuspended with 1ml cold 0.1M CaCl2. The suspension of the 1st tube was pipetted into the 2nd tube. The competent cells were stored for 2 days at 4°C.

4.1.5.3- Transformation

Chemical transformation protocol was used to transform DNA plasmids into competent *E-coli* (Hanahan, 1983). The ligation mixture (20µl) was mixed with 100µl of freshly prepared competent cell suspension (XL1). 3 cycles of; 10min on ice incubation/ 5 min heating at 37°C were done. Transformation mixtures were cooled down on the ice for 1-2 minutes and then 1 ml TY-medium was added. Cells were incubated for 15 minutes at 37 °C while shaking. Some or all of the transformation mixture were plated onto a 10 cm LB agar plate containing the appropriate antibiotic (usually, 3 plates for each transformation mixture, 200 µl for the first plate, 400 µl for the second, and the rest for the third plate) using a metal loop that was flamed beforehand. Plates were incubated upside down overnight at 37 °C.

4.1.6- Blue-White Screening

Blue-white screening is an efficient and rapid technique that is used to identify recombinant bacteria. It relies on the activity of β-galactosidase in *E. coli*, which cleaves lactose into glucose and galactose. The used cloning plasmid pKS (pBluescript II KS (+) GenBank X52327; Stratagene) carries a *LacZα* gene, which encodes α-peptide (N-terminal part of β-Galactosidase). The *E.coli* XL1-Blue strain has lacZΔM15 deletion mutation, which generates ω-peptide. When the plasmid vector is taken up by these cells, due to the α-complementation process, a functional β-galactosidase enzyme is produced. The multiple cloning site (MCS) of pKS plasmid is located within the *LacZα* gene. When the coding sequence is successfully inserted into the plasmid, it disrupts this gene such that it is unable to produce a functional β-galactosidase enzyme. X-Gal, a chromogenic substrate, is added to the agar plate. If β-galactosidase is produced, X-gal is hydrolyzed and spontaneously dimerizes to produce an insoluble blue pigment (5, 5'-dibromo-4,4'-dichloro-indigo). 100 µl IPTG (Isopropyl β-D-1-thiogalactopyranoside) and 100 µl of X-Gal were added on LB agar plates with the appropriate antibiotic. The plates were allowed to dry in the incubator, and the bacteria solution (transformation mixture) was added to the plates. After the drying of excess fluid, the plates were then incubated at 37˚C overnight. Bacteria with recombinant colonies will appear in white colonies, while the bacteria with non-recombinant plasmids will appear in blue colonies.

4.1.7- Sanger sequencing

While Taq polymerase extends the primer, ddNTPs (2', 3' dideoxynucleotides) terminate the extension reactions because they fail to form a phosphodiester bond with the next deoxynucleotide. This results in randomly fragments that differ in length by one base. Each of

the four ddNTPs is attached with a different fluorescent dye and emits light at a different wavelength when excited by a laser. Taq Cycle products can be separated by capillary gel electrophoresis, which differ from each other with one fluorescent base and analyzed by ABI PRISM® 310 Genetic Analyzer according to the manufacturer's instructions (Applied Biosystems). Dye-terminator chemistry and a thermostable polymerase (BigDye Terminator Prämix v1.1, Applied Biosystems, #4337450) were used. In a PCR tube, the following mixture was prepared; 2µl primer, 3µl BigDye, 2µl DNA (miniprep), up to 15µl RNase Free H2O.

Taq Cyle condition: 96°C/1min- 94°C/20sec- 56°C/30sec-60°C/4min (25X) - 4°C ∞.

DNA fragments were purified via DyeEx spin columns (DyeEx 2.0 Spin Kit, QIAGEN, # 63204) as follows:

- Gently vortex the spin column to re-suspend the resin
- Loosen the cap of the column a quarter turn.
- Snap off the bottom closure of the spin column, and place the spin column in a 2 ml collection tube.
- Centrifuge for 3 min at 595 x *g* (2800 rpm; Heraeus biofuge fresco) at room temperature.
- Carefully transfer the spin column to a clean centrifuge tube. Slowly apply the sequencing reaction (10–20 µl) to the gel bed.
- Centrifuge for 3 min at 595 x *g* at room temperature.
- Remove the spin column from the microcentrifuge tube (The eluate contains the purified DNA)
- Dry the sample in a vacuum centrifuge (BACHOFER vacuum concentrator).
- Dissolve the pellet with 15µl highly deionized formamide (Life Technologies #4311320).
- Denature the DNA by heating the tube at 96˚ C for 2 min and then cooling the tube down by putting it on ice.
- DNA sequence chromatograms were visualized by Chromas softwareVersion 2.6.5.

Sequencing primers:

M13 forward	5′-ACGACGTTAAAACGACGGCCAG	(Sigma)
M13 reverse	5′-ATTTCACACAGGAAACAGCTA	(Sigma)
Z746_40F	5′-CCCCAGACCTCTTGATGC	(Eurogentec)
Z746_1247R	5′-CTTGCGGAGGTGGTCCTT	(Eurogentec).

4.1.8- Restriction digestion of DNA

A combination of 10x buffers and restriction enzymes (one or more) were used according to the supplier's instructions for analytical and preparative digestion. For analytical purposes, the subcloned plasmids were restricted at certain restriction sites and the resulting fragments were then size controlled by agarose electrophoresis, whereas preparative digestion was used to cut out the desired dsDNA fragment from residual vector to be purified later from agarose gels. In addition, restriction digestion was applied to prepare plasmids for receiving specific fragments for subcloning. In a typical reaction, for analytical purposes, five units of restriction enzyme were used to digest about 400ng of plasmid DNA. For preparative purposes, 40 units of restriction enzyme were used to digest 10µg of plasmid DNA. The buffers have been chosen according to the manufacturer's instructions.

***Table 8**: A list of used restriction enzymes for analytical and preparative purposes.*

Restriction enzyme	Composition	Restriction site
*Xho*I (Thermo Scientific,#ER0691)	10 U/µl, 2000U	5′-C↓T C G A G-3′ 3′-C A G C T↑C-5′
*Sal*I (Thermo Scientific, #ER0645)	10 U/µl, 1500U	5′-G↓T C G A C-3′ 3′-C A G C T↑G-5′
*EcoR*V (Fermentas, #ER0301)	10 U/µl, 2000U	5′-G A T↓A T C-3′ 3′-C T A↑T A G-5′
*Hind*III (Fermentas, #ER0501)	10 U/µl, 5000U	5′-A↓A G C T T-3′ 3′-T T C G A↑A-5′
*BamH*I (Thermo Scientific, #ER0051)	10 U/µl, 4000U	5′-G↓G A T C C-3′ 3′-C C T A G↑G-5′
*EcoR*I (MBI-Fermentas, #ER0271)	10 U/µl, 4000U	5`-G↓A A T T C-3′ 3′-C T T A A↑G-5′
*Sac*I (Fermentas,#ER1131)	10 U/µl, 1200U	5′-G A G C T↓C-3′ 3′-C↑T C G A G-5′
*Bgl*II (Thermo Scientific, #ER0082)	10 U/µl, 2500U	5′-A↓G A T C T-3′ 3′-T C T A G↑A-5′
Pvu II (Boehringer Mannheim, 642 690)	10 U/µl, 1000U	5′-C A G↓C T G-3′ 3′-G T C↑G A C-5′
*Eco*311(*Bsa*I) (Thermo Scientific, #ER0291)	10 U/µl, 1000U	5′-G G T C T C (N)$_1$↓-3′ 3′-C C A G A G (N)$_5$↑-5′
*Xba*I (MBI Fermentas, #ER0681)	10U/µl, 1500U	5′-T↓C T A G A-3′ 3′-A G A T C↑T-5′
*Sma*I (Fermentas, #ER0661)	10U/µl, 1200U	5′-C C C↓G G G-3′ 3′-G G G↑C C C-5′
*Nde*I (Fermentas, #ER0582)	10U/µl, 2500U	5′-C A↓T A T G-3′ 3′-G T A T↑A C-5′
*Nae*I (Fermentas, #ER1522)	10U/µl, 1000U	5′-G C C↓G G C-3′ 3′-C G G↑C C G-5′
*Nhe*I (Fermentas, #ER0972)	10U/µl, 2500U	5′-G↓C T A G C-3′ 3′-C G A T C↑G-5′

***Table 9**: A list of used buffers*

Buffer	Composition
Blue buffer (Thermo Scientific, #BB5)	10 mM Tris-HCl (pH 7.5 at 37°C), 10 mM MgCl2,0.1 mg/ml BSA
Green buffer (Thermo Scientific, #BG5)	10 mM Tris-HCl (pH 7.5 at 37°C), 10 mM MgCl2,0.1 mg/ml BSA
Orange buffer (Thermo Scientific, #BO5)	50 mM Tris-HCl (pH 7.5 at 37°C), 10 mM MgCl2,100 mM NaCl, 0.1 mg/ml BSA.
Red buffer (Thermo Scientific, #BR5)	10mM Tris HCl,10mM MgCl2,100mM KCl,0.1mg/mL BSA, pH 8.5
Tango buffer (Thermo Scientific, #BY5)	33 mM Tris-acetate (pH 7.9 at 37°C),10 mM magnesium acetate,66 mM potassium acetate,0.1 mg/ml BSA.

***Table 10**: Analytical and preparative digestion protocol*

Component	Analytical digestion	Preparative digesion
DNA	~400ng	~10µg
Buffer	3µl	10µl
Restriction enzyme 1	0.5µl	4µl
Restriction enzyme 2	0.5µl	4µl
H_2O	Up to 15µl	Up to 100µl
Incubation	1h at 37°C	4-16h at 37°C

4.1.9- Agarose Gel Electrophoresis

Electrophoresis in horizontal agarose gels is a widely used technique for separation, purification, and identification of plasmid DNA and DNA fragments. Due to its negatively charged phosphate groups, DNA migrates towards the anode in an electric field. Under constant voltage, linear double-stranded DNA fragments migrate inversely proportional to their molecular weight within a certain range. The size of a DNA fragment can be determined by comparison with standard DNA marker fragments of identified size.

Depending on the size of the DNA molecules, the agarose concentration chosen was between 1 % and 1.5 % (w/v). DNA-samples were mixed with DNA loading buffer and applied into the wells of the gel. TBE (Tris-borate-EDTA) or TAE (Tris-acetate-EDTA) buffers were used to dissolve agarose and as an electrophoretic buffer. The electrophoretic separation was done at 90 volts for 7.5x5.5cm gels, and at 140 volts for 18x11 cm gels. Separated fragments were visualized by UV-light.

***Table 11**: Solutions of agarose gel*

Solution Name	Composition
TBE 1x for analytical gels (Tris-Borate-EDTA)	In 1l: 10.8g Tris Base, 5.5g Boric acid 4ml EDTA 0.5M, pH=8.0
TAE 1x for preparative gels (Tris-acetate-EDTA)	In 1l: 4.84g Tris Base, 1.142g Boric Acid, 2ml EDTA 0.5M,pH=8.0
Loading buffer 7x	0.175% (w/v) bromophenol blue, 35mM EDTA, 35% (v/v) Glycerol
λ DNA *EcoRI*/*Hind*III marker (Thermo Scientific. # SM0191) Φ X174 DNA *Hae*III marker (Thermo Scientific. # SM0251)	0.08 µg/µl DNA, 10mM Tris-HCl 0.03% Bromophenol blue 0.03% Xylene cyanol FF 60% Glycerol 60mM EDTA pH 7.6

- Weigh agarose, for 1% or 1.5% gel.
- Add 30ml 1x TBE / TAE buffer for 8 well-gel or 100ml for 28 well-gel, and 20-30ml H2O, and boil in the microwave.
- Add 2µl for 8 well-gel or 6µl for 28 well-gel ethidium bromide, mix, and cast gel.
- Electrophoresis chamber filled with 1x TBE (for analytical gel) or TAE (for preparative gel) buffer covering gel, add 20µl ethidium bromide for 8 well-gel or 40µl for 28 well-gel.
- The samples were mixed with 7x loading buffer and loaded into agarose wells. The electrophoresis was run at 90V for 8-well gels, and at 140V for 28-well gels.
- Visualization and documentation: Azure Biosystem C200

4.1.10- DNA gel extraction

DNA fragments with flanking restriction overhangs were purified from agarose gel via QIAEX® II, Gel Extraction Kit 150, #20021

- Excise the DNA bands from the agarose gel with a clean and sharp scalpel.
- Weigh the gel slices in colorless tubes, add QX1 buffer to the gel (300µl/100 mg gel).
- Re-suspend the QIAEXII by vortexing 30 sec. add QIAEXII to the sample (30µl).
- Incubate at 50˚C for 10 min to solubilize the agarose and bind the DNA, mix by vortexing every 2 min to keep QIAEX II in suspension.
- Centrifuge for 30 sec at 16060 x g (13000 rpm; Heraeus biofuge fresco).
- Wash the pellet with 500 µl of buffer QX1, vortex and centrifuge 30sec at 16060 x *g*.

- Wash the pellet with 500 µl of buffer PE, vortex and centrifuge 30sec at 16060 x *g*. Repeat this step.
- Air dry the pellet until the pellet becomes white.
- Elute DNA with free nuclease H2O and suspend the pellet, incubate 5 min at RT.
- Centrifuge for 30 sec at the highest speed, the supernatant then contains the purified DNA.

4.1.11- DNA precipitation

For efficient ligation and transformation, the digestion mixture of linearized plasmids should be liberated from high salt concentrations.

- Add Isopropanol and sodium acetate as follows: 1 volume digestion mix + 1 volume isopropanol + 0.1 volume 3M sodium acetate, vortex
- Centrifuge at 16060 x *g* (13000 rpm; Heraeus biofuge fresco) for 15 min.
- Discard the supernatant by pipetting; add 200 µl Ethanol 70% and briefly vortex.
- Centrifuge at 16060 x *g* for 6 min. Repeat this washing step again.
- Discard the supernatant by pipetting, and let the DNA to dry.
- Dissolve the DNA in 40 µl ddH2O.
- Detect on an agarose gel.

4.1.12- Plasmid Isolation

4.1.12.1- Mini-preparation of plasmid DNA from E. coli

The plasmid mini-preparation provides a small scale, quick method to isolate low amounts of sufficiently purified plasmid DNA for restriction and sequence analysis. An alkaline lysis buffer is used to lyse the cultivated bacteria, which contains detergent SDS (to denature bacterial proteins) and strong base NaOH (to denature chromosomal and plasmid DNA). Followed by adding neutralizing buffer with sodium acetate leading to rapid reannealing of plasmid DNA, and precipitation of the chromosomal DNA and bacterial proteins. The soluble plasmid DNA was precipitated with isopropanol, and the DNA pellet was washed with ethanol 70% (Birnboim and Doly, 1979).

***Table 12**: buffers and solutions for minipreparation plasmid isolation method*

Solution Name	Composition
Plasmid solution I (Resuspension buffer)	50mM Glucose, 10mM EDTA , 25mM Tris-HCl buffer, pH 6-8
Plasmid solution II (Lysis buffer)	0,2M NaOH, 1% SDS
Plasmid solution III (Neutralization buffer)	3M Ca-Acetate, 11,5% (v/v) Acetic acid
RNAse	10mg/ml

- Pick a single colony from LB agar plate and grow the bacteria in 2 ml TY medium (with appropriate antibiotic).
- Overnight culture, 37°C, shake 200rpm
- Pour the Bactria suspension into 1.5 ml tubes.
- Centrifuge at RT, 9520 x *g* (10 000 rpm; Heraeus biofuge fresco) for 1min.
- Discard the supernatant.
- Re-suspend the pellet with 100 µl of plasmid solution I (by rubbing over a grid).
- Add 200 µl of plasmid solution II, mix by repeated gentle inversion.
- Add 150 µl of plasmid solution III, vortex.
- Centrifuge at RT, 16 060 x *g* (13 000 rpm; Heraeus biofuge fresco) for 3min.
- Transfer the supernatant into new tubes.
- Add 250µl Isopropanol, mix by inverting the tubes.
- Centrifuge at RT, 16 060 x *g* for 5min.
- Discard supernatant, centrifuge briefly and discard the rest of supernatant.
- Rinse the pellet with 200µl of Ethanol 70%, brief vortex.
- Centrifuge at RT, 16 060 x *g* for 1min.
- Discard the supernatant by using a pipette.
- Air dry.
- Dissolve the DNA in tubes with 50µl H2O + RNase [0.4µg/µl]. Incubate on a shaker at low speed about 5 min to remove the RNA.

4.1.12.2- Midipreparation of plasmid DNA from E. coli

For purifying plasmid DNA of higher purity and quantity, QIAGEN midi preparation kit was used. The method is based on QIAGEN Anion-Exchange Resin. QIAGEN plasmid buffer set (Qiagen #19046) and QIAfilter plasmid midi kit (Qiagen #12243, containing QIAGEN-tip 100 anion-exchange tips and QIAfilter Midi cartridges) were used for Midi preparations.

***Table 13**: QIAGEN midi prep buffers*

Name	Composition
Plasmid Solution I	50mM Tris-HCl, 10mM EDTA, 100µg/ml RNAse A, pH 8.0
Plasmid Solution II	0,2M NaOH, 1% SDS
Plasmid Solution III	3.0 M KAc pH 5.5
Buffer QBT	750mM NaCl ,50mM MOPS, 15% Isopropyl alcohol (v/v) ,0.15 % Triton® X-100 (v/v), pH 7.0
Buffer QC	1.0 M NaCl, 50mM MOPS,15% Isopropyl alcohol (v/v), pH 7.0
Buffer QF	1.25M NaCl, 50mM Tris-HCl, 15% Isopropyl alcohol (v/v), pH 8.5

- Inoculate 5µl of starter bacteria suspension in 50ml TY medium/antibiotic, incubate with shaking 200rpm at 37°C

- Collect the bactria culture in 50 ml tube, centrifuge at 4°C, 2778 x *g* (4000 rpm Heraeus. Megafuge® 1.0R) for 10min Discard the supernatant.
- Add 4ml of buffer I to re-suspend the cells (Resuspension buffer).
- Add 4ml of buffer II (lysis Buffer), mix (not vortex), wait about 5 min until the mixture becomes clear.
- Add 4ml of buffer III (neutralization buffer), mix by repeated gentle inversion.
- Incubate in ice for a few minutes (about 5 minutes).
- Centrifuge at 5°C, 6240 x *g* for 5 minutes.
- Filtrate the supernatant by using funnel and glass wool, put the filtrate in ice.
- Setup the Qiagen tip 100.
- Equilibrate a QIAGEN-tip 100 by applying 4 ml Buffer QBT, and allow the column to empty by gravity flow.
- Apply the supernatant (filtrate) to the QIAGEN-tip and allow it to enter the resin by gravity flow.
- Wash the QIAGEN-tip with 2 x 10 ml Buffer QC.
- In a new 50 ml tube, elute DNA with 5 Buffer QF.
- Precipitate DNA by adding 3.5 (0.7 volumes) room-temperature isopropanol to the eluted DNA. Mix and centrifuge immediately at 6240 x *g*, for 15 min at RT. Carefully decant the supernatant.
- Wash the pellet with 2 ml of 70% ethanol, brief vortex.
- Centrifuge immediately at 6240 x *g* (6000 rpm; Heraeus. Megafuge® 1.0R), for 10 min at RT.
- Discard the supernatant, brief centrifuging and pipette the rest of supernatant.
- Air-dry the pellet, and dissolve the DNA in 500 µl ddH2O
- Measure DNA concentration via nanodrop® ND-1000 Spectrophotometer.

4.1.12.3- Endotoxin-free mini preparation of plasmid DNA from E. coli

Plasmids with low endotoxin content, intended for transfection of eukaryotic cells, were prepared with the QIAGEN EndoFree plasmid buffer set (Qiagen; #19048) and QIAfilter plasmid midi kit (Qiagen; 12243, containing QIAGEN-tip 100 anion-exchange tips and QIAfilter Midi cartridges) were used for midi preparations.

***Table 14**: QIAGEN midi prep endotoxin free buffers*

Solution Name	Compostition
Plasmid solution I Endotoxin-free	50mM Glucose, 10mM EDTA, 25mM Tris HCl buffer, pH 6-8
Plasmid solution II	0,2M NaOH, 1% SDS
Plasmid solution III Endotoxin-free	5M CaAc, 4,31% (v/v) Acetic acid
RNAse	10mg/ml
Buffer ER Endotoxin free	Proprietary
Buffer QBT Resin equilibration	750mM NaCl, 50 mM MOPS buffer, pH 7.0, 15% Isopropyl alcohol (v/v), 0.15 % Triton® X-100 (v/v)
Buffer QC Wash buffer	1.0 M NaCl, 50 mM MOPS, pH 7.0, 15% isopropyl alcohol (v/v)
Buffer QN Elution buffer	1.6 M NaCl, 50 mM MOPS, pH 7.0, 15 % isopropyl alcohol (v/v)

- Inoculate 5µl of starter bacteria suspension in 50ml TY medium/antibiotic, incubate with shaking 200rpm at 37°C.
- Harvest the bacterial cells by centrifugation at 2778 x *g* (4000 rpm; Heraeus. Megafuge® 1.0R) for 15 min at 4°C.
- Re-suspend the bacterial pellet in 4 ml Buffer P1 by pipetting up and down until no cell clumps remain.
- Add 4 ml Buffer P2, mix gently but thoroughly by inverting 4-6 times, and incubate at room temperature for 5 min.
- During the incubation prepare the QIAfilter Cartridge: Screw the cap onto the outlet nozzle of the QIAfilter Midi Cartridge to prevent dripping. Place the QIAfilter Cartridge into a convenient tube.
- Add 4 ml chilled Buffer P3 to the lysate, and mix immediately but gently by inverting 4-6 times. Proceed directly to step 7. Do not incubate the lysate on ice.
- Pour the lysate into the barrel of the QIAfilter Cartridge. Incubate at room temperature for 10 min. Do not insert the plunger. Do not agitate the QIAfilter Cartridge during this time
- Remove the cap from the QIAfilter outlet nozzle. Gently insert the plunger into the QIAfilter Midi Cartridge and filter the cell lysate into a convenient tube.
- Add 1ml (0.1 volumes) Buffer ER to the filtered lysate, mix by inverting the tube approximately 10 times, and incubate on ice for 30 min.
- Equilibrate a QIAGEN-tip 100 by applying 4 ml Buffer QBT, and allow the column to empty by gravity flow.

- Apply the cleared lysate to the QIAGEN-tip and allow it to enter the resin by gravity flow.
- Wash the QIAGEN-tip with 2 x 10 ml Buffer QC.
- In new clean 50 ml tube elute DNA with 5 ml Buffer QN.
- Precipitate DNA by adding 3.5 ml (0.7 volumes) of room-temperature isopropanol to the eluted DNA. Mix and centrifuge immediately at 6240 x *g* (6000 rpm; Heraeus. Megafuge® 1.0R) for 30 min at 4°C. Carefully decant the supernatant.
- Wash DNA pellet with 2 ml of endotoxin-free room-temperature 70% ethanol and centrifuge at 6240 x *g* for 10 min at RT.
- Air-dry the pellet, and redissolve the DNA in a suitable volume of endotoxin-free Buffer TE. Do not dissolve with TE buffer, Dissolve with 250µl nuclease-free H2O
- Measure the DNA amount via Nanodrop® ND-1000 Spectrophotometer.

4.1.13- Cloning Strategies

4.1.13.1- Plasmids for expression in eukaryotes

PCR products were cloned blunt-end into cloning vector pKS (pBluescript II KS (+) GenBank X52327; Stratagene) and Sanger sequencing was used for verification of all PCR amplicons. The desired cassettes were inserted as *Xho*I/*Sal*I fragments into the pM3 vector (Sadowski et al., 1992) restricted with *Sal*I to generate a coding sequence for a fusion protein with the DNA binding domain of the yeast transcription factor GAL4 (GAL4-DBD) at its N-terminus. Correspondingly, eukaryotic expression vectors encoding fusion proteins with glutathione-S-transferase from *Schistosoma japonicum* (GST) were generated by insertion of the *XhoI*/*Sal*I fragments into pN2-GST (Born et al., 2014) digested with *Sal*I (Fig. 6). The sequence encoding full-length ZNF777 (amino acids 1-831) was assembled in the pM3-ZNF777/1-362 construct by inserting through the *Sac*I site the respective fragment from a commercial truncated ZNF777 plasmid (BioCat GmbH/Heidelberg; # BC023985-seq-TCHS1003-GVO-TRI, GenBank accession BC023985). The DNA sequence for full-length ZNF746a was purchased from BioCat GmbH/Heidelberg (BC068505-seq-TCHS1003-GVO-TRI). The coding sequences of mutant protein regions (Z746b/109-188_R141L, Z746b/109-188_G142E, Z746b/109-188_RG/LE, and Z746a/94-279-K189R) with flanking *Xho*I/*Sal*I overhangs were purchased as *in-vitro*-synthesized DNA fragments that were cloned into the *EcoR*V site of the pUC57 vector (Eurogentec). *Xho*I-*Sal*I fragments were subcloned into pExpr-IBA105 (IBA

GmbH #2-1905-000) digested with *Xho*I to generate a coding sequence of N-terminal One Strep-tagged proteins.

To construct a coding sequence of a chimeric fusion protein between DUF3669 and the KRAB-AB domain of ZNF10, the coding sequences of DUF3669 domains of both ZNF777 and ZNF746 without stop codon (Z746a/1-108 and Z777/189-254, respectively) were first cloned as *Xho*I/*Sal*I fragment into pM3 restricted with *Sal*I. Then the resulting plasmids were linearized with *Sal*I and the coding sequence of the KRAB domain of ZNF10 as *Xho*I/*Sal*I fragment was inserted in this site (Fig. 7). *EcoR*I/*Sal*I DNA fragments for Z777/189-254 and Z777/1-282 were subcloned into pEGFP-C1 (Clontech #6081-1) restricted with *EcoR*I/*Sal*I to generate a coding sequence for fusions with an N-terminal green fluorescent protein (GFP) fusions. To clone GST and GST-Z746a/1-279 into pcDNA5-FRT-TO expression vector (Invitrogen #V6520-20). The respective pN2-GST plasmid was restricted with *Hind*III/*Sal*I restriction enzymes, and the resulting fragments were inserted into pcDNA5-FRT-TO plasmid restricted with *Hind*III/*Xho*I restriction enzymes.

4.1.13.2- Plasmids for expression in prokaryotes

For prokaryotic expression, the GST-Z746a/1-108 encoding *Bgl*II-*Sal*I cassette was subcloned into bacterial expression vector pRSFDuet1 (Merck/Novagen #71341-3) restricted with *Bgl*II-*Xho*I. To construct a coding sequence for maltose binding protein fusions (MBP), the *Sac*I-*Sal*I fragments were subcloned into pMALc2x (New England Biolabs, addgene #75286) restricted with *Sac*I-*Sal*I. To generate 5xHis tagged proteins, *Xho*I/*Sal*I fragments were subcloned into pRSET-B (Invitrogen # V351-20) restricted with *Xho*I.

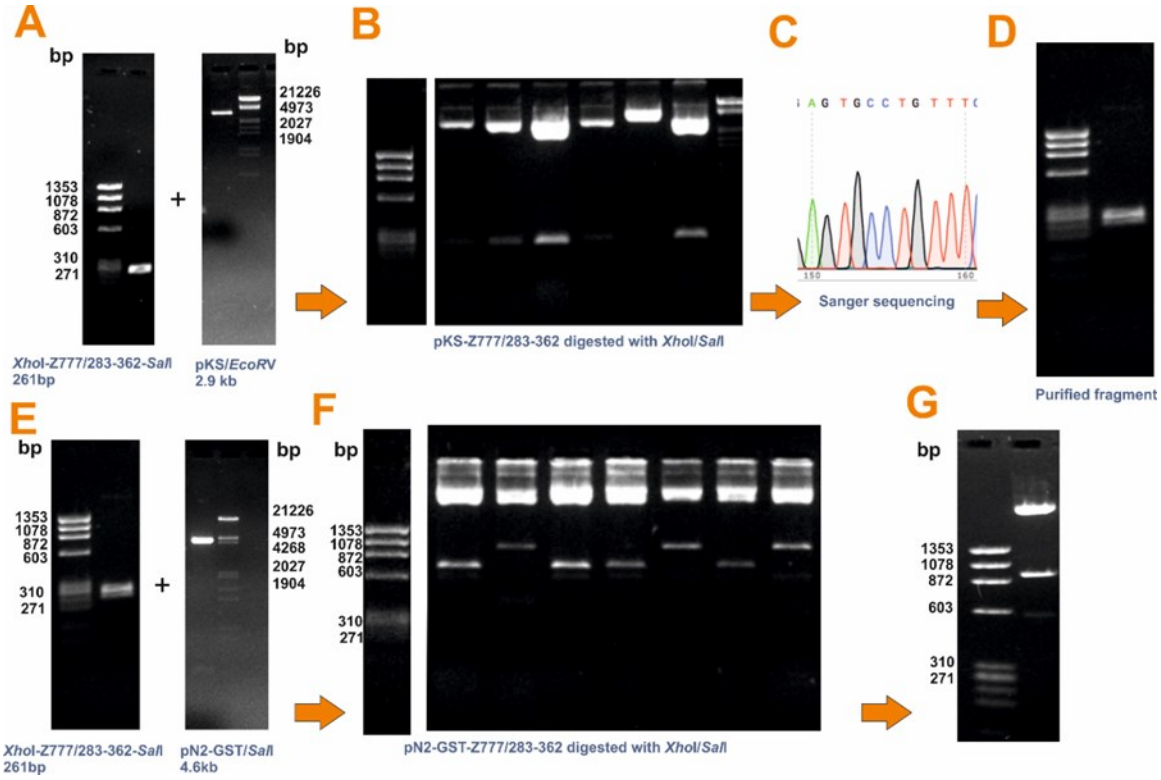

Fig. 6: Plasmid construction milestones. Exemplary plasmid construction milestones of pN2-GST-ZNF777/283-362. *Images depict ethidiumbromide-stained agarose gels.* ***A****: A blunt dsDNA PCR amplicon encoding ZNF777/283-362 with XhoI/SalI sites at its 5' or 3' end, respectively, was synthesized by RT-PCR using RNA derived from human testis as initial source (left image) and cloned blunt-end into cloning vector pBluescript KS linearized by EcoRV (right image).* ***B****: After transformation blue-white screening, plasmid mini preparation and analytical digestion with XhoI/SalI, clones containing the desired plasmid (here identified by the excision of a 249bp fragment) were identified by electrophoresis.* ***C****: Sanger DNA sequencing method was performed to verify the DNA sequences.* ***D****: XhoI-ZNF777/283-362-SalI was isolated from pKS vector by digestion with XhoI/SalI restriction enzymes followed by separation on agarose gel and purification.* ***E:*** *Subcloning into expression vector pN2-GST linearized with SalI.* ***F:*** *Minipreparation analysis to identify the sucloned colonies (subcloned vector gives 932bp fragment, and non-cloned vector gives 638bp fragment).* ***G:*** *Expansion of the subcloned colony followed by isolation of the subcloned plasmid using endotoxin-free midipreparation plasmid isolation method.*

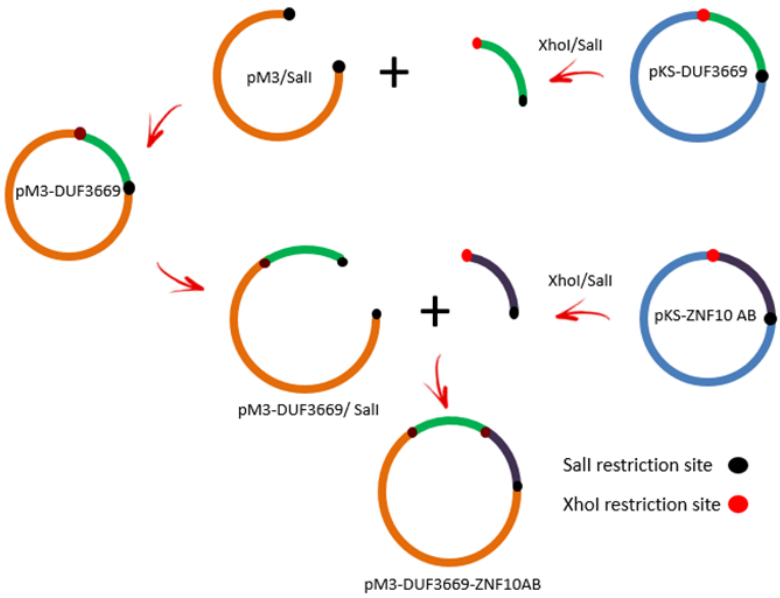

Fig. 7: The workflow of construction of chimeric DUF3669-ZNF10AB encoding sequences (ZNF746a/1-108-ZNF10AB and ZNF777/189-254-ZNF10AB) in pM3 plasmid. *pKS plasmid constructs with DUF3669 or ZNF10-AB sequences serve as donor plasmids. The desired fragment wer excised with XhoI and SalI restriction enzymes. DUF3669 was inserted into pM3 plasmid linearized with SalI. Then, the recombinant plasmid was restricted again with SalI and ZNF10AB was inserted at 3' end of DUF3669 resulting in a coding sequence of chimeric N-terminal DUF3669-ZNF10AB protein.*

Table 15*: Plasmids that were used as either cloning or expression vectors*

Plasmid	Description
pKS	pBluescript II KS(+) GenBank X52327 (Agilent/Stratagene # 212205)
pM3	Dr. Lorenz, Immunology Rostock (Sadowski et al., 1992)
pN2-GST	Dr. Lorenz, Immunology Rostock Modified from pEGFP-N2 (Clontech #. 6081-1) (Born et al., 2014)
pM3-ZNF10AB	Dr. Lorenz, Immunology Rostock (Born et al., 2014)
pN2-GST-ZNF10AB	Dr. Lorenz, Immunology Rostock (Born et al., 2014)
pExpr-IBA105	IBA GmbH # 2-1905-000_ 5471 bp
pcDNA5-FRT-TO	Invitrogen, #V6520-20
pOG44	Invitrogen, #V6005-20
pRSFDuet1	Merck/Novagen #71341-3; 3829bp
pUC57Z746AB345	Eurogentec, Order 1000849727 Project: U2844CI250-1
pBR-ZNF 746	BioCat GmbH/Heidelberg (#BC068505-seq-TCHS1003-GVO-TRI).
pOTB7-BC023985	BioCat GmbH/Heidelberg (# BC023985-seq-TCHS1003-GVO-TRI , Gene ID 27153)
pUC57-Z746NT1a_K189R	Eurogentec, Project: U9686DD110-1, Clone ID_M89815
pGL3c-5'Gal5	(Born et al., 2014)
pRLuc/TK	Promega
pEGFP-C1	Clontech #6081-1, 4737 bp
pMal-c2X	New England Biolabs. Add gene, #75286, 6645 bp
pRSET-B	Invitrogen # V351-20, 2900 bp

4.2- Cell Culture

4.2.1- Cell lines and media

All cell culture works were performed under a cell culture hood type II to maintain an aseptic work area (Herasafe™ KS). Media were pre-warmed immediately before use.

Table 16: *Cell lines and media used in this work*

Cell line	Origin	source	Medium
HeLa	human epitheloid cervix carcinoma cells	German Cancer Research Center, Heidelberg	IMDM Iscove's Modified Dulbecco's Medium Gibco®, #12440-053
Flp-In™ T-REx™ 293	293 human embryonic kidney cells (Graham et al., 1977)	Thermo Fisher Scientific #R78007	
HAP1 Wild type	male chronic myelogenous leukemia (CML) cell line KBM-7	Horizon Genomics #C631	DMEM+GlutaMAX™ Dulbecco's Modified Eagle Medium Gibco ®, #31966-021
HAP1 TRIM28 ko		Horizon Genomics # HZGHC000293c001	
HAP1 SETDB1 ko		Horizon Genomics # HZGHC001331c001	

Media were supplemented with 10% inactivated fetal calf serum FCS (Biochrom FBS Superior, #S 0615), and 45 units/ml penicillin and 45 µg/ml streptomycin (Gibco®, #15070-063). Flp-In™ T-REx™ 293 cells were in addition kept in the presence of 100µg/ml Zeocin (Invitrogen, #64-0509) and 15µg/ml Blasticidin (PAA Laboratories GmbH, #P11-017)

4.2.2- Freezing and Thawing human Cell lines

The cells were frozen in freezing medium (50%complete growth medium + 40%FBS + 10% DMSO) in liquid nitrogen. The cryovial containing the frozen cells was taken from liquid nitrogen storage and immediately placed into a 37°C water bath for ~1min. The cells were re-suspended with 5ml pre-warmed complete growth medium in a 15 ml falcon tube and centrifuged at 210 x *g* (1100 rpm; Heraeus, Megafuge® 1.0) for 5min at room temperature. The supernatant was aspirated, and the pellet was re-suspended with 15ml pre-warmed complete growth medium, transferred to 250ml cell culture flask, and allowed to grow in a humidified atmosphere containing 5 % CO2 at 37 °C (BINDER).

4.2.3- Cell passaging and seeding

At confluency of maximally 90%, the medium was discarded and cells were briefly washed with 5 ml sterile 1x DPBS (Gibco, #14/90-094) and trypsinized with a 0.05 % Trypsin-EDTA solution (Gibco, # 25300-054) for ~5 minutes at 37 °C. When the cells were being suspended, 5ml medium were added and the cell suspension was mixed thoroughly. The total number of

cells was determined using a hemocytometer (Neubauer Improved counting chamber, #T729.1). Appropriate volumes of cell suspensions were pipetted into new cell culture vessels and allowed to grow in the incubator.

Table 17: *Cell culture plastic wares*

24-Well Cell Culture Plate sterile, with lid	Greiner bio-one, #662 166
12-Well Cell Culture Plate sterile, with lid	Greiner bio-one, #665180
6-Well Cell Culture Plate sterile, with lid	Greiner bio-one, #657 160
6cm diameter cell culture dishes	Greiner bio-one, #628 160
10cm diameter cell culture dishes	Greiner bio-one, #633171
CELLSTAR® Filter Cap Cell Culture Flasks, 75cm^2, 250ml	Greiner bio-one, #658175
CELLSTAR® Filter Cap Cell Culture Flasks, 25cm^2, 50ml	Greiner bio-one, #690175

4.2.4- Transient transfection

The lipid-based transfection method was used for all transfections. Commercial reagents that contain a proprietary blend of lipids and other components and supplied in 80% ethanol were applied according to manufacturer's instructions. The cells were seeded in cell culture plates or dishes and maintained until reaching the desired cell confluency. All transfections were conducted at about 70-90% cell confluency. For dual luciferase assay; HAP1 wild type and knockout cells were transfected using TurboFectin 8.0 reagent (OriGene, #TF81001), whereas HeLa cells were transfected using X-tremeGENE 9 (Roche, #06365787001). For immunoprecipitation and affinity pull-down assays, HeLa cells and Flp-In™ T-REx™ 293 cells were transfected with FuGENE® HD (Promega, E2311). The media were changed just prior to transfection. Transfection complexes were formed at 3µl transfection reagent per 1 µg DNA in Opti-MEM® I medium (Gibco, #51985026) during a 15-20min incubation at room temperature and added then to the cells dropwise. The dishes and the plates were then gently rocked back and forth and from side to side to ensure thorough mixing.

4.2.5- Generation of stable cell lines

The Flp-In T-Rex 293 cell line (short FITR293) was used for the generation of isogenic lines that inducibly and stably overexpress the gene of interest (Karimi et al., 2007). Flp-In T-Rex 293 cells contain an integrated FRT (Flp recombination target) locus known to be permissive for robust transcription. This locus expresses a lacZ-zeocin fusion as a marker. Further, the cells stably express a Tet-OFF Tet repressor protein under the control of the human CMV promoter and blasticidin as a selection marker. The recombination into the target locus is accomplished by using two plasmids. Plasmid pcDNA5/FRT/TO (Invitrogen, #V6520-20) plasmid contains the gene of interest under the control of a CMV/TetO2 promoter and the

hygromycin resistance gene with an FRT site embedded in the 5′ coding region. The hygromycin resistance gene does not have a promoter and the ATG startcodon. Vector pOG44 (Invitrogen, #V6005-20) constitutively expresses the Flp recombinase under the control of the human CMV promoter. Incorporation of pcDNA5/FRT/TO into FRT site of the genome makes the SV40 promoter and the ATG start codon (from pFRT/lacZeo) near and frame with the hygromycin resistance gene, and at the same time disrupts the lacZ-Zeocin fusion gene. Thus, stable Flp-In T-REx expression cell lines can be characterized for the following phenotypes; Blasticidin resistance, Hygromycin resistance, Zeocin sensitivity, and Lack of β-galactosidase activity

Flp-In T-Rex 293 cells were grown in 6cm dishes under the following selective medium; [(DMEM) supplemented with 10 % fetal calf serum (FCS) (Biochrom FBS Superior S 0615), 100 units/ml penicillin and 45µg/ml streptomycin (Gibco®, #15070-063) + 45µg/ml Zeocin (Invitrogen #64-0509) + 15µg/ml Blasticidin (PAA laboratories GmbH_ #P11-017)].

Directly before transfection, the selective medium was replaced with a new selective medium that lacks Zeocin. Cells within each dish were co-transfected with either pcDNA5-FRT-TO-GST or pcDNA5-FRT-TO-ZNF746a/1-279 (expresses Hygromycin B resistance and respective GST fusion protein), and with pOG44 (expresses Flp recombinase). The total amount of DNA is 5µg), 3µl/1µg DNA of transfection reagent was deployed (FuGENE® HD).

24h post-transfection, the cells were transferred from 6cm dish to 10 cm dish (Greiner bio-one, #633171) (each in 2x 10cm dishes). 48 h later, the medium was replaced with second selective medium that is composed of; [DMEM+10% FCS + 45 units/ml penicillin+45 µg/ml streptomycin] + 15µg/ml Blasticidin + 250µg/ml Hygromycin B. Two weeks later, the visible foci were picked and grown in 24 cell culture well plates under the condition of the selective medium that contains 250µg/ml Hygromycin B. Colonies were screened for expression of the desired protein by indirect immunofluorescence and Western blotting (Fig: 13). The expression of the transgene was induced by 2µg/ml tetracycline for 24 hours before analysis. The positive colonies were expanded in 75cm^2 cell culture flasks. The cells were trypsinized and counted. Then, 1 million cells were re-suspended in 1 ml freezing medium (50% growth medium + 40%FBS + 10% DMSO) and put in 1 ml cryogenic vials and stored at -70 ºC for few days. Vials were transferred to the liquid nitrogen tank (-180 ºC) for long-term storage.

4.3- Production of recombinant proteins in E. coli

E. coli strain BL21 (DE3) contains the λDE3 lysogen that contains T7 RNA polymerase gene under the control of the lacUV5 promoter. Isopropyl β-D-1-thiogalactopyranoside (IPTG) induces the expression of the T7 RNA polymerase that enhances the overexpression of the recombinant genes cloned downstream of a T7 promoter.

This strain is deficient in Lon protease (cytoplasm) and OmpT protease (outer membrane), resulting in a higher yield of intact recombinant proteins.

Genotype: fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS

λ DE3 = λ sBamHIo ΔEcoRI-B int (lacI PlacUV5 T7 gene1) i21 Δnin5

Here, an optimized commercial strain, SoluBL21, that has undisclosed mutations after screening for higher solubility of recombinant proteins, was used and obtained as competent bacterial stock (Amsbio # C700200). Transformation proceeded with 10ng of 5xHis tagged protein expression vector (pRSET-B-Z746a/1-108, pRSET-B-Z777/189-254, and pRSET-B-Z777/1-282), GST (pGEX6P-1), GST-Z746a/1-108 (pRSFDuet1-GST-Z746a/1-108), MBP (pMalc2x) and MBP-Z777/189-254 (pMalc2x-777/189-254). 49µl of competent bacteria were mixed with an equivalent volume of 10ng plasmid, the heat-pulse of the mixture was conducted at 42° C for 45 seconds, and then 250µl of TY (Tryptone Yeast) medium were added to the mixtures, and incubated for 1h at 37°C with shaking. The transformation mixtures were plated onto a 10 cm LB (lysogeny broth) agar plate using a metal loop, which was flamed beforehand. The plates were incubated upside down overnight at 37°C.

For the starter culture, a single colony was picked from each plate and incubated in 2ml of glucose-rich growth medium (RMG) supplemented with an appropriate antibiotic at 37°C overnight with shaking 200rpm. For expansion, 50µl of the overnight culture were diluted in 5ml of fresh RMG medium and cultured under the same conditions until reaching OD600 nm≈0.5. The bacteria were pelleted by centrifuging for 10min at 2777x *g* (4000 rpm; Heraeus. Megafuge® 1.0R) at 4°C; washed with 5ml PBS and re-suspended in 5 ml SB-medium with an appropriate antibiotic.

Expression was induced by adding 0.1mM IPTG (5 µl of 0.1M IPTG stock = 23.8 mg/ml in H2O) to SB bacterial suspension and the bacterial mixture was incubated with shaking for 4h at 37 °C. The bacteria were then pelleted by centrifuging for 10min at 2777x *g* (4000 rpm;

Heraeus. Megafuge® 1.0R), 4°C, the pellets were washed with 5ml of 50mM TRIS/HCl (PH:8), 300mM NaCl.

The bacteria were re-suspended with buffer lysis buffer (Table 22). Cells were lysed by four 15-second sonication cycles. The lysates were cleared by centrifugation for 30min at 16060 x g (13 000rpm; Heraeus fresco biofuge) at 4°C. The expression levels and quality of the different recombinant proteins in these crude extracts were analyzed by Western blotting (Fig. 14) and served as input for pull-down assays. shows western blot results of fusion proteins after probing with respective antibodies.

Table 18: *Media used for in vitro translation*

Medium	Composition
RMG	for 1 liter: 10 g Tryptone, 5 g yeast extract, 5 g NaCl, 2 g Glucose
SB	component A for 1L: 12 g peptone from casein, 24g yeast extract, 4ml glycerol, component B for 1L: 340 mM KH2PO4, 1.44M K2HPO4; ready-to-use medium: 950ml component A+50ml B

4.4- Immunofluorescence

Indirect immunofluorescence staining was used to control for expression of a construct and for qualifying the intracellular distribution of the encoded protein.

Table 19*: Immunofluorescence solutions*

Solution	Composition
PAFA fixative	1xPBS, 4% (w/v) paraformaldehyde, pH=7.3
Permeabilization buffer	1x PBS, 0.5% (w/v) Triton X-100
Antibody buffer	1x PBS, 1% (v/v) Normal goat serum, 1% (w/v) BSA
DAPI staining buffer	1x PBS, 0,5µg/ml DAPI
PBS	137mM NaCl, 12mM Phosphate, 2.7mM KCl, pH=7.4
Mounting medium	90% (v/v) glycerol, 10% (v/v) PBS, 0.1% (w/v) p-phenylenediamine, adjusted with 0.1M carbonate bicarbonate buffer to pH=8.5
Carbonate-bicarbonate buffer	184mM Na_2CO_3, 16mM $NaHCO_3$

The coverslips (20 x 20 mm) were flamed and placed into cell culture wells. If Flp-In T-REx 293 cells were used, the coverslips were coated with poly-L-lysine (SIGMA, #25988-63-0, 0.01% solution) as follows:

- Place a drop of poly-L-lysine on each coverslip and leave for 15min.
- Aspirate off the poly-L- lysine and rinse the coverslips 2x with PBS.
- Aspirate off the PBS and let the coverslips dry for 1-2 h.

The cells were grown on glass coverslips placed in 6-well cell culture plates. The transfection or expression induction was conducted when the cells reached approximately 70-90% of

confluency. 24h post-transfection, the medium was aspirated and the cells were rinsed briefly with cold 1x PBS, fixed by incubating the cells with 2ml PBS (4%/w/v) paraformaldehyde for 15 min at RT without shaking, and permeabilized by incubating cells in 1ml PBS, 0.5% (w/v) Triton X-100 for 5 min at RT.

The cells' sides of coverslips were incubated with 45µl of primary antibody solution for 30min at RT (Table 20). The coverslips were then rinsed briefly with PBS, washed 3x 5min with PBS, and incubated with 45µl secondary antibodies solution for 30min at RT. Then the coverslips were rinsed and washed 1x 5min with PBS. The nuclei were stained with 2 ml 4′,6-diamidino-2-phenylindole (DAPI) for 5min and washed with PBS 1x 5 min. The cell side coverslips on object slide were mounted in mounting medium and sealed with nail polish. The slides were stored at -20°C.

Images were acquired with a confocal microscope (Leica TCS2 AOBS) using a 63× oil immersion lens (numerical aperture of 1.32) or Zeiss LSM 780 Confocal Microscope with the PlanApochromat 63x/1.40 Oil DIC M27 objective. Image Studio Lite 5.2 and ZEISS ZEN Imaging Softwares were used for image processing.

Table 20: *Primary and secondary antibodies used for immunofluorescence staining*

Antibody	Stock concentration	Dilution
Rabbit anti-GAL4 (Santa Cruz, # sc-577)	200µg/ml	1:100
Monoclonal antibody anti TIF1beta/KRIP-1 (BD 610680/K57620)	250µg/ml	1:200
Rabbit anti TRIM28/TIF1-beta (Santa Cruz, # sc-33186)	200µg/ml	1:50
Monoclonal mouse anti GST (Abcam, #ab92)	1mg/ml	1:100
Rabbit anti GST (Santa Cruz, #sc-459)	200µg/ml	1:200
Goat anti-rabbit-IgG-fp488 (Interchim Fluoprobes, #FP-GARBTTGY488)	0.5mg/ml	1:400
Goat anti-mouse-IgG-Alexa633 (Life Technologies, #A-21052)	2mg/ml	1:200
Goat anti-rabbit-IgG-fp642 (Interchim Fluoprobes, #FP-GARBTTGY642)	0.5 mg/ml	1:200
Goat anti mouse-IgG-Alexa488 (Thermo Fisher, #A-11029)	2mg/ml	1:400

4.5- Immunoprecipitation and Pulldown Assays

4.5.1- Lysis of eukaryotic cells

Testing for the presence of the presumable protein interaction partners in the same protein complex was done using co-immunoprecipitation or affinity pull-down assays based on ligands for the tags of the interrogated proteins after transfection of respective constructs. The

scale for co-immunoprecipitation was usually 0.5-1 million cells per 6-cm dish grown for 2 days, transfected with a total amount of 5µg expression vector(s) and harvested 24 hours post-transfection. Affinity pull-down assays used the same scale or single wells of 6-well plates transfected with a total of 2.5µg DNA. The media were aspirated and the cells were briefly rinsed with ice-cold 1x PBS and washed with 3ml 1 x TST2 buffer. The cells were incubated with 1.5 ml TST2 complete for 5 min on ice. The cells were then scraped with a cell scraper (TPP Cell scraper 30cm, 99003). The raw extracts were transferred into 2ml reaction tube, vortexed, and the extracts were sheared through QiaShredder columns (Qiagen, #79656) and centrifuged for 5 min at 4°C/16060 x *g* (13000rpm, Heraeus Biofuge Fresco). The flowthrough extracts were centrifuged (~1500 µl) for 10min at 16060 x *g*, 4°C. The supernatants were transferred into a fresh tube.

4.5.2- Immunopreciptation and pulldown assays in eukaryotic cells

All buffers and tubes were ice cold before use. The beads were prepared by aspirating the storage buffer and washing the beads with TST2 buffer, for each sample appropriate amount of beads were pipetted into 1.5ml tube (Table 22).

For immunoprecipitation, the beads were suspended in 25µl TST2 complete + 1µl antibodies against TRIM28; (roughly 1mg/ml concentration; mAb 1Tb1A9, (Remboutsika et al., 1999); a kind gift of Pierre Chambon and Régine Losson; Illkirch, France). The tubes were incubated at RT with shaking for 30min.

1000µl of crude cell extracts were mixed to beads/beads loaded with antibodies and incubated overnight with end-over-end mixing at 4° C, 500 µl of extract for each sample were stored at -20°C. The beads were washed with 3 x 0.5 ml each with TST2+0.5% TritonX-100. The beads were then washed again with 1 ml with TST2. The beads were eluted with respective elution buffer. The eluates and the whole-cell extracts (TST2 extract) were tested on 12% SDS-polyacrylamide gel electrophoresis and Western blotting.

4.5.3- Pulldown assay with NiNTA agarose beads

An appropriate aliquot (70 -200 µl) of a 5xHis-tagged recombinant protein extract was combined in vitro with an aliquot of another extract containing GST- and MBP fusion proteins (40 - 75 µl) in different combinations and mixed with 20µl Ni-NTA magnetic bead suspension in extract buffer. After overnight incubation at 4°C on a rotating mixer, the beads were washed 5 x 5 min with wash buffer (Table 22) followed by washing for 3 x 5min with

the same buffer without BSA. The proteins were eluted and the eluted proteins were separated on 18% SDS-PAGE gel.

***Table 21**: The beads used for immunoprecipitation and pulldown assays.*

Beads	Used volume/sample	Elution buffer
Protein G Mag-Sepharose Beads GE Healthcare, #28-9440-08)	25µl/sample	30µl 1x SDS sample buffer (Table 23)
Strep Tactin-Superflow IBA GmbH, #2-1206-010	50µl/sample	40µl 3mM D (+)-Biotin in TST2 Carl Roth GmbH #3822.1
Glutathione Magnetic Agarose Beads Pierce, #78601	50µl/sample	40µl 50mM reduced Glutathione in TST2 Sigma-Aldrich #G4251-1G
Ni-NTA magnetic bead Qiagen, #36111	20µl/sample	250mM imidazole in 50mM TRIS/HCL, 300mM NaCl (PH:8)

***Table 22**: Buffers used for immunoprecipitation and pulldown assays*

Buffer	Composition
TST2	20 mM TRIS/HCl pH7.5; 60 mM KCl, 15 mM NaCl, 10 mM MgCl2, 1mM CaCl2, 250 mM Sucrose
TST2 complete	TST2 buffer, 0.5% Triton X-100, cOmplete ULTRA Tab. Mini EDTA free (Roche,#0589279100),1mM DTT, 1 mM Na3VO4, 50 mM NaF, 40 mM beta-Glycerophosphate.
Lysis buffer of prokaryotic cells	50mM TRIS/HCl (PH:8), 300mM NaCl, protease inhibitor (cOmplete ULTRA Mini EDTA free, Roche_ # 05 892 970 001, 1tablet/10ml), 20% glycerol, 0.1% bovine serum albumin (BSA), 20mM imidazole, and 0.05% tween 20
Wash buffer for NiNTA beads	50mM TRIS/HCl (PH:8), 300mM NaCl, 0.1% bovine serum albumin, 20mM Imidazole and 0.05% Tween20

4.6- SDS-PAGE (sodium dodecyl sulfate-polyacrylamide gel electrophoresis)

One dimensional gel electrophoresis under denaturing conditions in the presence of 0.1% SDS was performed to test for protein expression and co-associated proteins. The gel is cast as separating denaturing sodium dodecyl sulfate-polyacrylamide gel (10%-18%) and topped by ~0.5-1cm of 4.5% staking sodium dodecyl sulfate-polyacrylamide gel. The proteins were denatured and solubilized by heating in the SDS loading buffer at 95°C for 5min.

Samples preparation:

For expression test, 24h post transfected cells were extracted with sodium dodecyl sulfate (SDS). The extracts were sheared through QiaShredder columns (Qiagen, 79656) and

centrifuged for 5 min at 16 060 x *g* (13 000rpm) at 4°C. SDS samples were denatured by heating 5min at 95°C. For purified proteins, the eluates were diluted with SDS loading buffer and denatured by heating 5min at 95°C. SDS samples were loaded into sodium dodecyl sulfate-polyacrylamide gel and the electrophoresis run was performed at constant 30 mA. Electrophoresis was stopped once the downmost sign of the protein marker almost reaches the foot line of the glass plate.

***Table 23**: SDS-PAGE buffers and solutions*

Solution	Composition
Tris/HCl buffer pH=8.9	1M Tris/HCl
Tris/HCl buffer pH=6.8	1M Tris/HCl
Acrylamide / Methylenbisacrylamide (BisAA) (Roth Rotiphorese Gel 30, #3029.1)	30%/0,8% (v/v)
Laemmli SDS-Sample Buffer 5x	312.5mM TRIS/HCl, 10% SDS, 325mM DTT, 0.08 Bromophenol blue, 50% Glycerin.pH=6.8
Tetramethylethylendiamine (TEMED) , Serva, #35930	99.0%
Ammonium persulfate (APS) Serva, #13375.03	10 % (w/v) in H2O, freshly prepared
Running buffer	0.25M TRIS BASE, 1%SDS, 1.92M Glycine

***Table 24**: SDS- polyacrylamide gel composition*

Solution	10% separation gel	12% separation gel	15% separation gel	18% separation gel	4,5% stacking gel
AA/BisAA	7ml	8ml	10ml	12ml	1.5ml
Tris/HCl buffer pH=8,9	8 ml	8ml	8ml	8ml	---
Tris/HCl buffer pH=6,8	---	---	----	---	1.25ml
H_2O	5 ml	4ml	2ml	---	7ml
10% SDS	216 µl	216µl	216µl	216µl	100µl
10% APS	160µl	160lµl	160lµl	160µl	100µl
TEMED	20 µl	20µl	20µl	20µl	10µl

4.7- Western blotting

Electroblotting of separated proteins from polyacrylamide gels on membranes

Semi-dry Western blotting was performed according to the method of Kyhse-Anderson (Kyhse-Andersen, 1984).

After electrophoresis, the separated proteins were electroblotted onto low-fluorescence polyvinylidene difluoride (PVDF) membrane (Immobilon®-FL, #IPFL00010). To allow the migrated protein to transfer from the gel to PVDF membrane, several layers of humidified absorbent filter papers were used.

18 filter papers and PVDF membrane were cut to the dimension of the SDS/PAGE gel, and they were assembled in a transfer cell as follows:

- 6x filter papers were saturated with HT and placed onto the platinum anode.
- 3x filter papers were saturated with LT.
- Equilibrated PVDF membrane (wetted with isopropanol, rinsed with dH2O, and wetted with LT buffer).
- Equilibrated SDS gel (by ε-aminocaproic acid buffer).
- 9x filter papers were saturated with a ε-aminocaproic acid buffer
- Stainless cathode
- After each step, roll out the air bubbles.
- The transfer cell was run at 25V for 2h (Bio-Rad Trans-Blot SD Semi-Dry Transfer Cell).

Transfer efficiency was checked by reversibly staining with Ponceau S dye (0.3 % Ponceau S in 3 % TCA), the membrane was then rinsed gently with dH2O water. An image was taken for Ponceau stained membrane.

The membrane was wetted again with isopropanol and rinsed with PBS. Non-specific antigen binding was blocked by Odyssey blocking buffer (LI-COR, #927-40000) at 4°C overnight. Transfer unit: Bio Rad Trans-Blot SD Semi-Dry Transfer Cell

Table 25: *western blot solutions*

Solution	Composition
ε-aminocapric acid buffer	0.04M ε-aminocapric acid, 0.025M TRIS base 20% Methanol, pH = 9.4
HT buffer	0.3M TRIS base, 20% Methanol, pH = 10.4
LT buffer	0.025M TRIS base, 20% Methanol, pH = 10.4
Ponceau S solution	0.3% Ponceau S (w/v), 3% Trichloro acetic acid
PBS	137mM NaCl, 12mM Phosphate, 2.7mM KCl, pH = 7.4

The probing of the transferred proteins with relevant antibodies was done as follows: The blot was incubated with antibodies solution for 2h. at RT with gentle shaking. Afterward, the blot was washed 4x5min with wash buffer and incubated with secondary antibodies solution for 1h at RT with shaking. The membrane was washed again 4x5min and rinsed with PBS. The membrane was stored in PBS, and protected from light until scanning. Primary antibodies were detected by fluorescently-tagged secondary antibodies against rabbit or mouse IgG (LI-COR # 926-32211, # 926-32220) and signals were acquired using an LI-COR Odyssey® CLx imager. The indicated image panels in figures were always subjected “in one piece”

simultaneously to any contrast or brightness adjustment within the LI-COR Image Studio Lite 5.2 software before export as black-and-white or pseudo-color image files.

***Table 26**: Buffers for staining with antibodies*

Blocking buffer	Odyssey® Blocking buffer (LI-COR # 927-40000) 1:2 diluted with filtrated 1x PBS
Wash buffer	1x filtrated PBS supplemented with 0.1 % Tween 20
Antibodies dilution buffer	Odyssey® Blocking buffer (LI-COR # 927-40000) 1:2 diluted with filtrated 1x PBS supplemented with 0.1 % Tween 20

***Table 27**: Primary and secondary antibodies used for detection of the blotted proteins*

Antibody	Stock concentration	Dilution
Rabbit anti-GAL4 (Santa Cruz, #sc-577)	200µg/ml	1:1000
Rabbit anti-GST (Santa Cruz, #sc-459)	200µg/ml	1:1000
Monoclonal antibody anti-GAPDH (Abcam, #ab8245)	500µg/ml	1:5000
Rabbit anti TRIM28/TIF1-beta (Epitomics, #3509-1)	not specified	1:2000
Monoclonal anti-One II (IBA, #2-1507-001)	100µg	1:1000
Monoclonal antibody anti-TRIM28 (BD Bioscience, #K57620/610681)	250µg/ml	1:1000
Rabbit anti MBP (Santa Cruz, #No.sc-808)	200µg/ml	1:2000
Mab anti GFP, (Santa Cruz, #sc-808).	400µg/ml	1: 1000
Monoclonal Penta His antibody, BSA free (Qiagen, #34660)	0.2mg/ml	1:1000
Goat anti-mouse-IgG-IRDye 680CW (Li-COR, #926-68070	1mg/ml	1:10 000
Goat anti-rabbit-IgG-IRDye 800CW (Li-COR, #926-32211)	1mg/ml	1:10 000

4.8- Dual-luciferase reporter assay

Dual-luciferase reporter assay is a heterologous reporter assay, in which the expression of the luciferase reporter gene is regulated by the GAL4-DBD fusion protein. Firefly luciferase (*Photinus pyralis*) reporter expressing vector pGL3control-(5'Gal4)5 (referred as pGL3c-5´Gal5 is derived from pGL3control (Promega, #E1741) and modified by inserting five DNA-binding sites for the Gal4 DBD at the *Bgl*II/*Hind*III restriction site upstream of a strong SV40 promoter (Born et al., 2014). In addition to firefly luciferase plasmid (pGL3c-5´Gal5), the cells were transfected with Renilla luciferase-expressing gene pGL 4.74 (hRluc/TK, Promega #E692A) which lacks GAL4-DBD DNA binding site was used for normalizing. Both luciferase reporter plasmids were transfected together with effector plasmid (expresses either GAL4 alone or GAL4 fusion protein). Three independent experiments at least with duplicates within each experiment were performed for each effector plasmid. Promega's Dual-Luciferase

Reporter Assay System (Promega, #E1910) was used to perform dual-luciferase assays in accordance with the manufacturer's instructions.

Cells were seeded in 6-well cell culture plates as follows; $2.5x10^5$ cells/well for HeLa, $3x10^5$ cells/well for HAP1 wild type and HAP1-SETDB1 knockout cells, and $6x10^5$ cells/well for HAP1-TRIM28 knockout cells. The cells were incubated for 48h until reaching 70-90% confluency and transfected with; 750ng effector plasmid, 250ng pGL3c-5´Gal5, and 10ng pGL 4.74 (hRluc/TK).

The medium was discarded, and the cells were briefly washed with 1x PBS and lysed by incubating with 500µl passive lysis buffer (supplied in the kit) on a shaker for 15min at RT. The lysates were then collected in 1.5ml tubes and centrifuged for 5min at RT/ 16060 x *g* (13000 rpm; Heraeus Biofuge fresco).

The activities of firefly and Renilla luciferases of each sample were sequentially measured after all samples were ready using single-tube measurements in Berthold single-channel luminometer (Lumat LB9501) as follows: Firefly luciferase measurements were started by adding 20 µl of the cell lysate to 100µl Luciferase Reagent II in luminometer plastic tube. The tube was immediately placed into the luminometer and measurements were carried out for 10 sec. Then, 100 µl of Stop & Glow® reagent were added to stop the luciferase activity and catalyze the Renilla reaction, and then Renilla activity was measured for 10 sec. Firefly luciferase activities were normalized with respective Renilla luciferase activities for each biological sample. The luciferase activity in the presence of the effector plasmid expressing the GAL4 DNA binding domain alone represents the reference transcriptional capacity. Fold repression was calculated experiment-wise by dividing the normalized luciferase activity of that reference through normalized luciferase activities of the sample with the respective test effector plasmid.

4.9- Compilation of DUF3669-containing KRAB-ZNFs

According to the InterPro database, the human genome encodes six KRAB-ZNF proteins with the “domain of unknown function 3669” (DUF3669) at their amino termini (http://www.ebi.ac.uk/interpro/entry/IPR022137/proteins-matched). First, we compiled all orthologs of the six human DUF3669-containing proteins from the TreeFam database release 9. 207 DUF3669-containing proteins were identified, all of them are ZNF proteins and can be categorized into six major orthologous protein groups, ZNF212, ZNF282, ZNF398, ZNF746, ZNF777, and ZNF783.

For each identified protein we then extracted the DUF3669-containing polypeptides including all other residues encoded by the same single exon from Ensembl Release 95 (January 2019). We then employed multiple alignments (https://www.genome.jp/tools-bin/clustalw) followed by using profile Hidden Markov Model (HMM) software (HMMbuild of HMMER version 2.3.1; installed on a local server) to generate profile consensus sequence matrices for DUF3669 (hmm file). Using the protein BLAST server of NCBI and Uniprot (https://blast.ncbi.nlm.nih.gov/Blast.cgi and https://www.uniprot.org/blast/) with default parameters, we searched for more DUF3669-containing proteins in the animal phyla. DUF3669-containing proteins were trimmed according to HMMER result to retrieve DUF3669 sequences. The total number of DUF3669-containing proteins tested in this study was 505. Finally, ClustalW alignments of the conserved sequences were used as input to generate sequence logo representations for the DUF3669 sequences (http://weblogo.threeplusone.com/create.cgi) (Crooks et al., 2004).

5- Results

The canonical function of the KRAB-ZNF proteins is the transcriptional repression. However, transcriptional repression activities of only a limited number of KRAB-ZNF proteins have been investigated in detail so far. Both, ZNF746 and ZNF777 are evolutionarily conserved proteins (Liu et al., 2014) and conform to the canonical structure of the KRAB-containing zinc finger protein family regarding the presence of an amino-terminal KRAB domain followed by a tandem array of contiguous zinc finger motifs. However, they display an additional exceptional feature, the domain of unknown function DUF3669 at their far amino termini. As indicated by its acronym, the function of DUF3669 is largely unknown.

5.1- Six human KRAB-ZNFs have an extremely evolutionarily conserved DUF3669 and specific to amniotes

A new domain model that was based on the DUF3669 sequences of the six human members and their orthologs in other species was generated. As described in the method section, all orthologs were compiled using public database resources and reciprocal BLAST searches, see additional file 1 in Al Chiblak et al., 2019. Interestingly, the DUF3669 domains are only found in the genomes of amniotes, but not in amphibians (Table 28).

Table 28: *The number of the DUF3669-containing KRAB-ZNF orthologs analyzed in this study.*

	ZNF212	ZNF282	ZNF398	ZNF746	ZNF767	ZNF777	ZNF783	unassigned	Total
Mammalia	48	62	33	59	1	64	33	26	326
Reptilia	1	10	0	0	0	7	0	61	79
Aves	0	27	0	0	0	19	0	55	101
Amphibia	0	0	0	0	0	0	0	0	0
Sum	**49**	**99**	**33**	**59**	**1**	**90**	**33**	**142**	**506**
no. of DUF3669	49	99	31	59	1	90	33	176	538
no. of KRAB-A	49	99	33	58	0	90	33	162	524
no. of KRAB-B	4	60	1	58	0	88	31	119	361
no. of proteins with ZnF (at least 1)	48	79	33	54	0	87	24	97	422

An overall consensus DUF3669 domain logo, as well as individual ortholog group logos were built (Fig. 8). The sequence analysis identified an extremely conserved polypeptide of 95 aa

that has several absolute conserved amino acids in all compiled proteins and might be a consensus of DUF3669 which is longer than that of PF12417.

As described for the DUF3669 domain (Fig. 7), an overall consensus KRAB-A logo as well as consensus logos for the individual orthologous groups were generated. The KRAB-A subdomain of ZNF10 was used as a standard KRAB-A subdomain (Fig. 9). Several previous studies demonstrated that MLE motif in the canonical KRAB-A subdomain plays an essential role in the interaction with TRIM28 and thereby the repressive activity (Margolin et al., 1994; Witzgall et al., 1994; Friedman et al., 1996; Peng et al., 2009; Murphy et al., 2016). Of note, all human DUF3669-containing zinc finger proteins and their orthologs lack this canonical MLE motif in their KRAB-A subdomain. Instead, they often display amino acid stretches like “MKG”, “MRG”, and “VKE”. This suggests that the KRAB domain of DUF3669-containing KRAB-ZNFs might not be able to form stable complexes with TRIM28 as well as show different behavior with respect to transcriptional regulation compared to canonical KRAB domains.

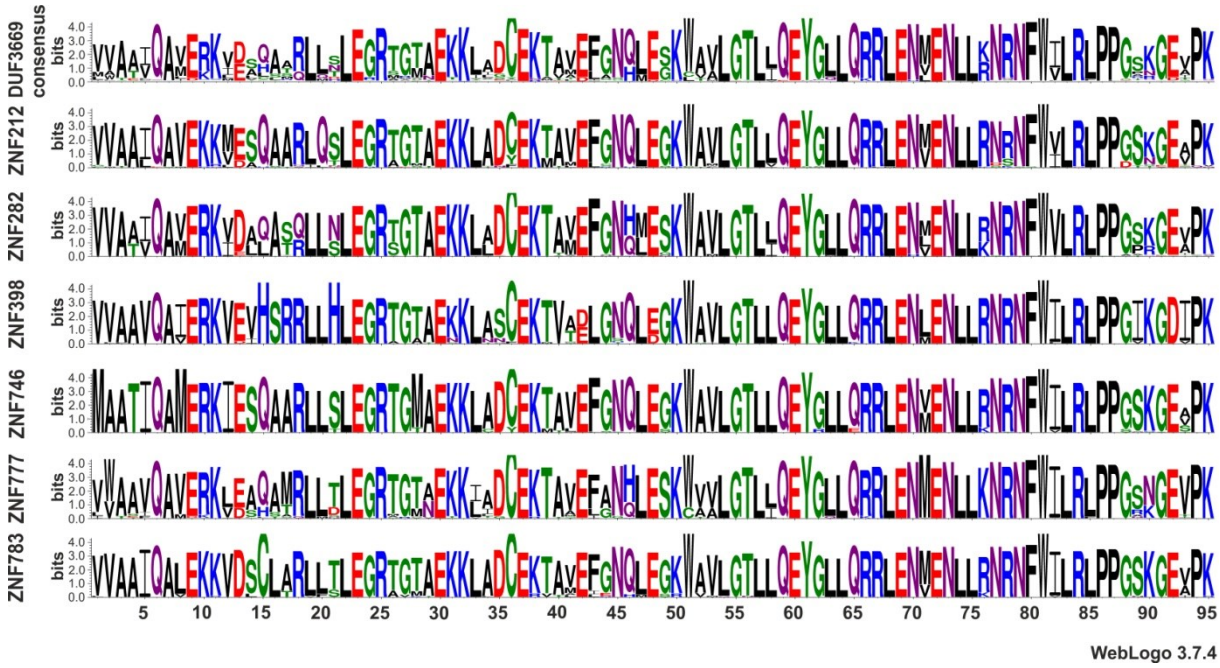

Fig. 8: Compilation of amino acid sequence logos for DUF3669 domain-encoding genes from mammals, birds, and reptiles*. The logos were generated with WebLogo 3.7.0 (http://weblogo.threeplusone.com/). The x-axis indicates the position within the domain and the y-axis denotes the information content (bits) for each amino acid. The height represents the frequency of each amino acid at a certain position within the tested sequence set. The overall DUF3669 consensus sequence at the top was derived from 505 DUF3669-containing proteins from amniotes. The other DUF3669 domain logos represent the ortholog groups of the six DUF3669-encoding KRAB-ZNFs (ZNF212 n=49; ZNF282 n=99; ZNF398 n=33; ZNF746 n=59; ZNF777 n=90; ZNF783 n=33). Orthologs were obtained from TreeFam database release 9, UniProt release 2019_02, and NCBI database.*

Orthologs of the human DUF3669-containing proteins were only found in amniote species from reptiles and birds to mammals, but not in amphibians. Sauropsids were thus the evolutionary oldest taxa encoding DUF3669 sequences according to our definition (Al Chiblak et al., 2019, additional file 1). Although named ZNF767, the seventh gene potentially gives only rise to a protein just consisting of DUF3669 without any other known domain. In public databases, ZNF767 is considered to be a pseudogene.

Interestingly, ZNF746 and ZNF212 orthologs were not detected outside mammals, whereas the others were found in mammals, birds, and reptiles. This suggests that ZNF746 and ZNF212 emerged more recently than other DUF3669-containing KRAB-ZNFs. Most ZNF746 orthologous proteins contain truncated KRAB-A subdomain at their amino termini. In addition, the amino acids in zinc finger motifs of the six human DUF3669-containing proteins that have the major impact on DNA binding specificity “recognition code” are clearly specific, this suggests that these proteins recognize different endogenous genomic targets and execute distinct biological functions, see supplement S2.

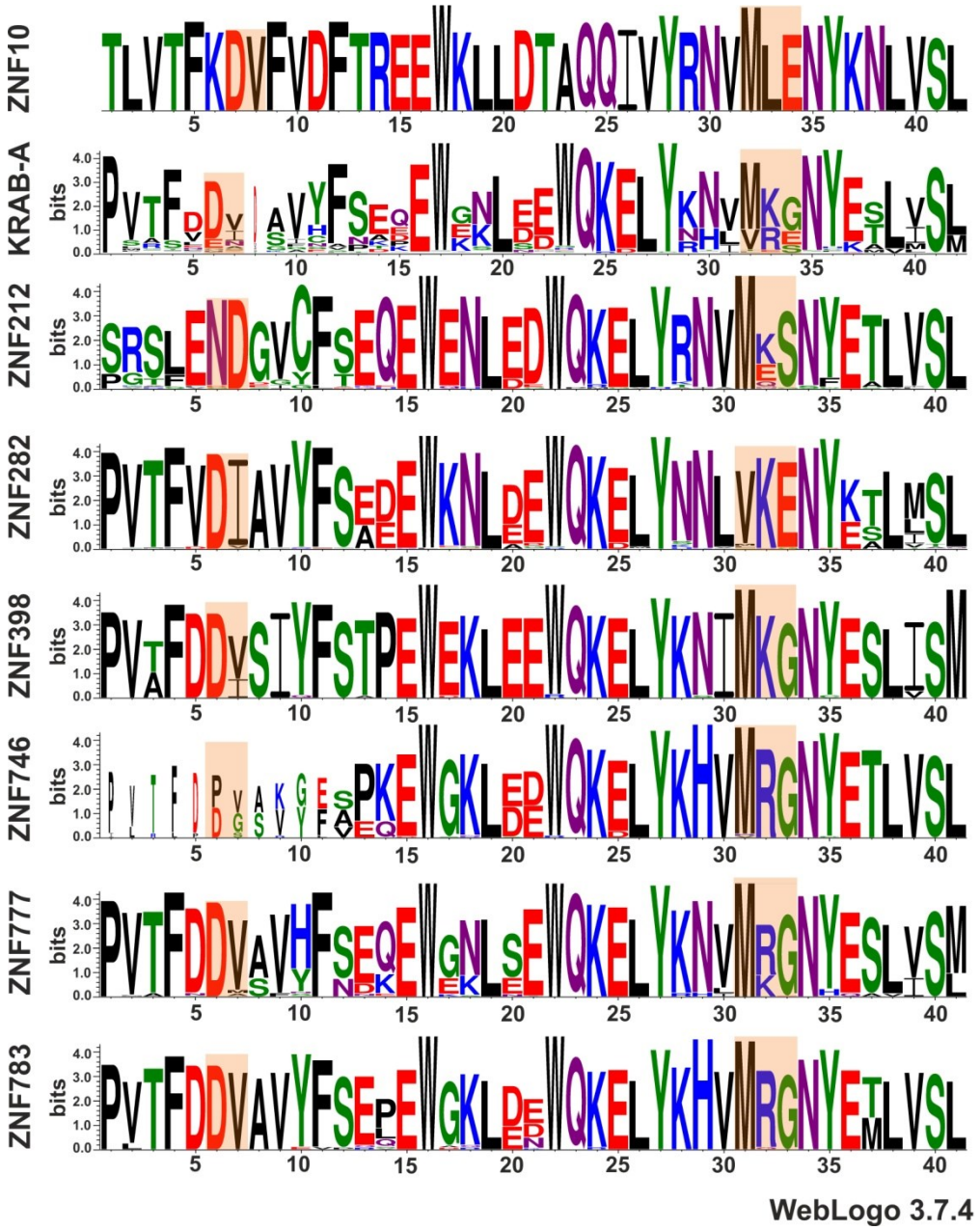

Fig. 9: Amino acid sequence logos for KRAB-A boxes in 505 DUF3669-containing zinc finger proteins and in each orthologous group as in Fig. 8*. The highlighted amino acids represent the position of amino acids thought to be essential for the interaction with TRIM28 and repressive activities (DV and MLE). The amino acid sequence of the KRAB-A subdomain of ZNF10 is used for comparison (on the top).*

5.2- Protein constructs and expression verification

A number of dedicated expression vectors were constructed to investigate the contribution of the KRAB and DUF3669 domains of ZNF746 and ZNF777 to transcriptional modulation and to study the formation of protein complexes. The expressed protein fragments contain KRAB domains, DUF3669 domains, zinc finger motifs, and combinations of them (Fig. 10) were used aminoterminaly to GAL4, GST and other tags as reported. The major isoform of

ZNF746 (PARIS, given name: ZNF746a_UniProtKB: Q6NUN9) has a KRAB-A subdomain that is 15 amino acids truncated at its amino terminus as compared to the human KRAB-A consensus (Supplement S1). However, sequence databases contain information on another isoform that has a complete KRAB-A. This isoform was designated ZNF746b and is represented by UniProtKB database entry A0A2R8YDQ5 (Fig. 16). The existence of this isoform was confirmed by cloning the respective constructs from an initial ZNF746 amplicon obtained by RT-PCR from the human fetal brain and human testis RNA. We compensated for the truncated region of the KRAB-A box in ZNF746a by adding 15aa from the adjacent amino-terminal part (Z746a/94-173) to get a relatively similar protein size to its counterpart of ZNF746b isoform.

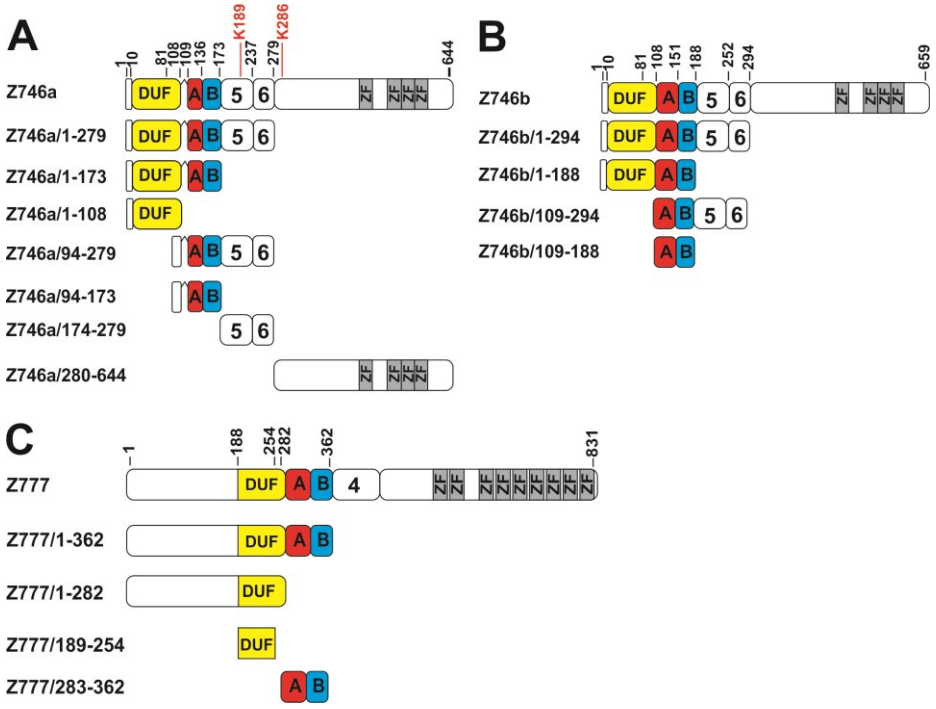

Fig. 10: Schematic representations of the primary protein structures encoded by the various constructs employed in this study*. Shown are the full-length proteins and the truncated versions along with the names of the constructs/protein pieces used throughout the manuscript. The different protein segments were expressed as fusion proteins with various partners depending on the experimental approach and indicated in the respective figure legends. Each curved round corner box denotes the protein region encoded by a separated exon. The numbers correspond to the amino-acids positions in the respective full-length protein. The known protein domains are labeled, i.e. DUF3669 (DUF, yellow), KRAB-A ("A", red), KRAB-B ("B", blue), C2H2 zinc finger motif (ZF, grey).* ***(A)*** *Major ZNF746 isoform designated Z746a (represented by UniProtKB: Q6NUN9) containing a truncated KRAB-A domain indicated by a gap in the depiction for highlighting.* ***(B)*** *ZNF746 isoform designated Z746b (represented by UniProtKB: A0A2R8YDQ5) with a full-length KRAB-A domain.* ***(C)*** *ZNF777 protein designated Z777 (UniProtKB: Q9ULD5-2).*

The expression of GAL4 fusions in HeLa cells (Fig. 11) and HAP1 cells (Fig.12) was evaluated by Western blotting before conducting reporter assays.

All transfected cells showed reliable expression bands, the amino parts of both ZNF746 isoforms (Z746a/1-279 and Z746b/1-294) showed degradation bands in HeLa cells (Fig. 11B).

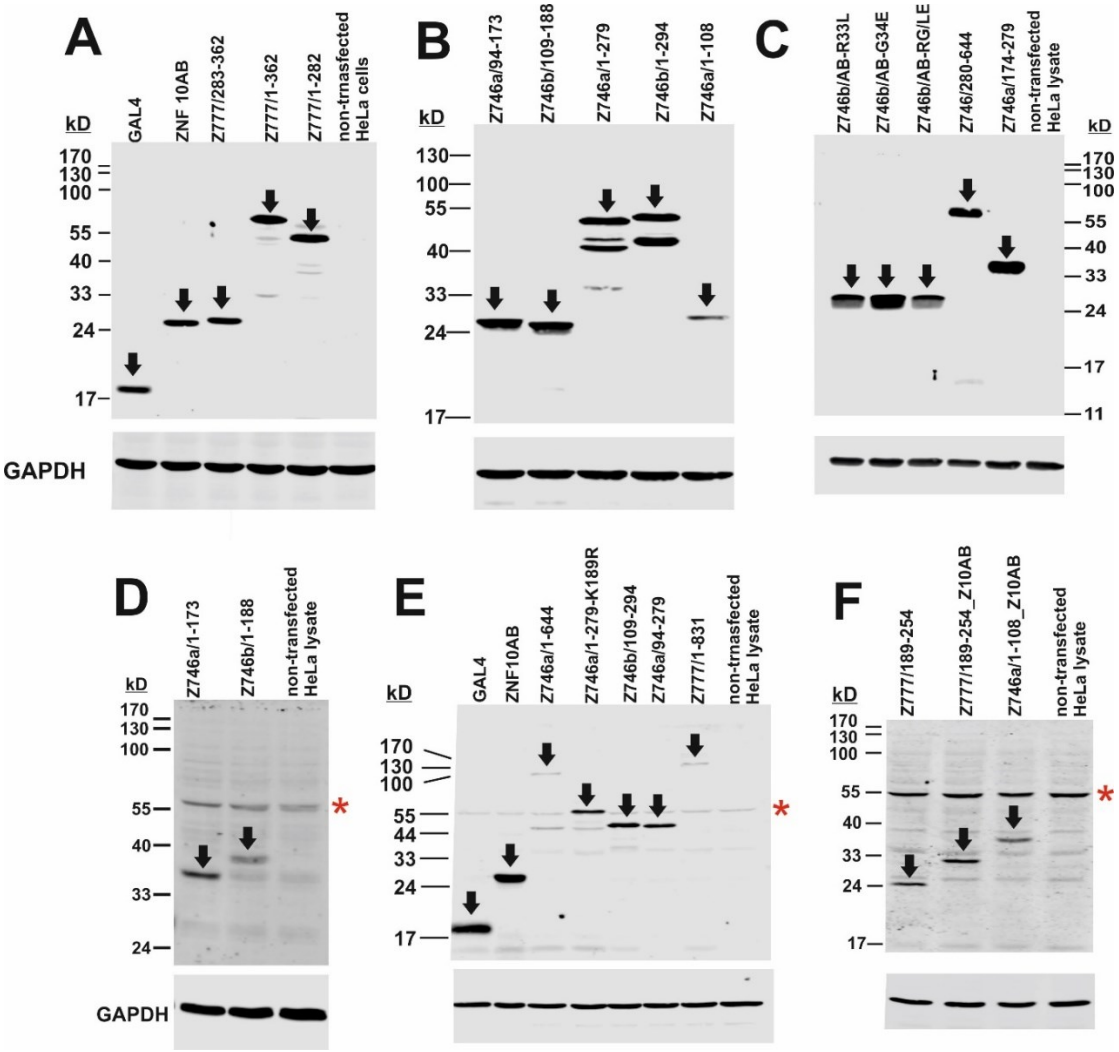

Fig. 11: Verification of Gal4 fusion protein expression of expected size encoded by the constructed pM3 expression vectors by Western blotting in HeLa cells*. The bands show the protein expression of 2µg effector plasmid (pM3 vector). HeLa extracts made 24 hours post-transfection with 1 x SDS sample buffer were probed with rabbit polyclonal antibodies against GAL4 and monoclonal antibodies against endogenous GAPDH. Non-transfected HeLa cell lysates are used as negative controls. Note that the higher molecular weight bands for GAL4-Z746a/1-279 and GAL4-Z746b/1-294 correspond to the computed molecular weight for GAL4-Z746a/1-279 and GAL4-Z746b/1-294 (48 and 50 kD, respectively). Moreover, full-length ZNF746a and full-length ZNF777 show significant weaker bands as compared to other bands. Arrowheads point to bands of expected size; * indicates unspecific bands recognized by the antibodies.*

It is worth mentioning that some bands of higher molecular weight appear in some lanes (indicated with blue asterisks in Fig. 12). These bands could be posttranslationally modified products e.g. SUMOylation that was reported for ZNF746 (Nishida and Yamada, 2016) and might occur in ZNF777 as well.

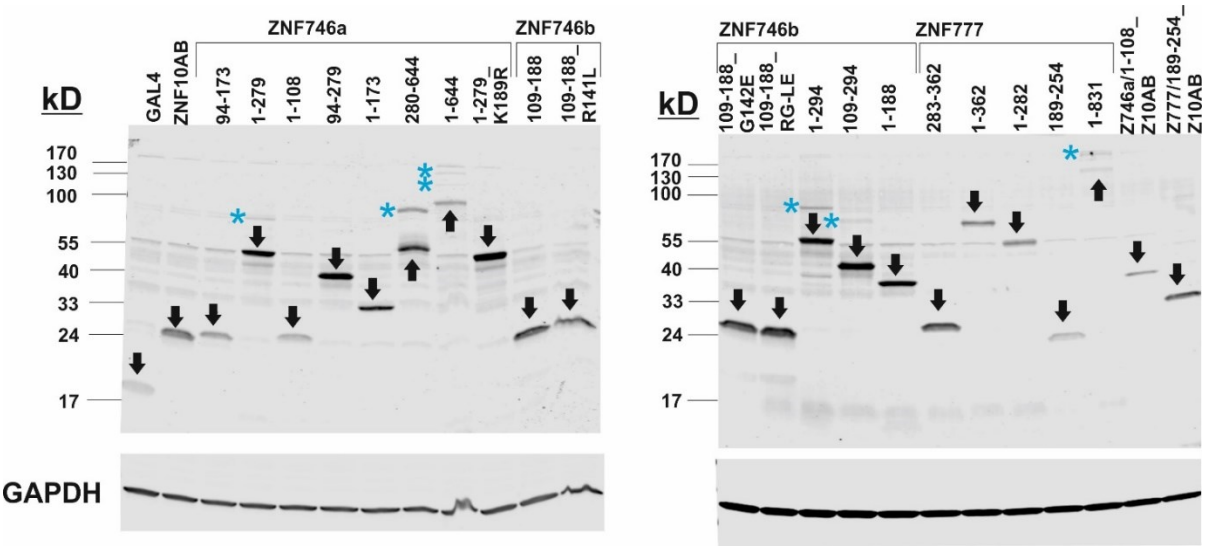

Fig. 12: Verification of Gal4 fusion protein expression of expected size encoded by the constructed pM3 expression vectors by Western blotting in HAP1 wildtype cells*. The cells were transfected, lysed, and analyzed as described in Fig.11. The bands indicated with blue asterisks might be modified proteins by SUMOylation.*

Stable cell lines expressing GST-Z746a/1-279 and GST were generated (see methods). The expression levels and spatial distribution of induced proteins were evaluated by Western blotting and immunofluorescence (Fig. 13). The stable cell lines could efficiently express GST and GSTZ746a/1-279 as compared with transiently transfected cells. In addition, as observed in HeLa cells for GAL4-Z746a/1-279, multiple degradation bands were observed.

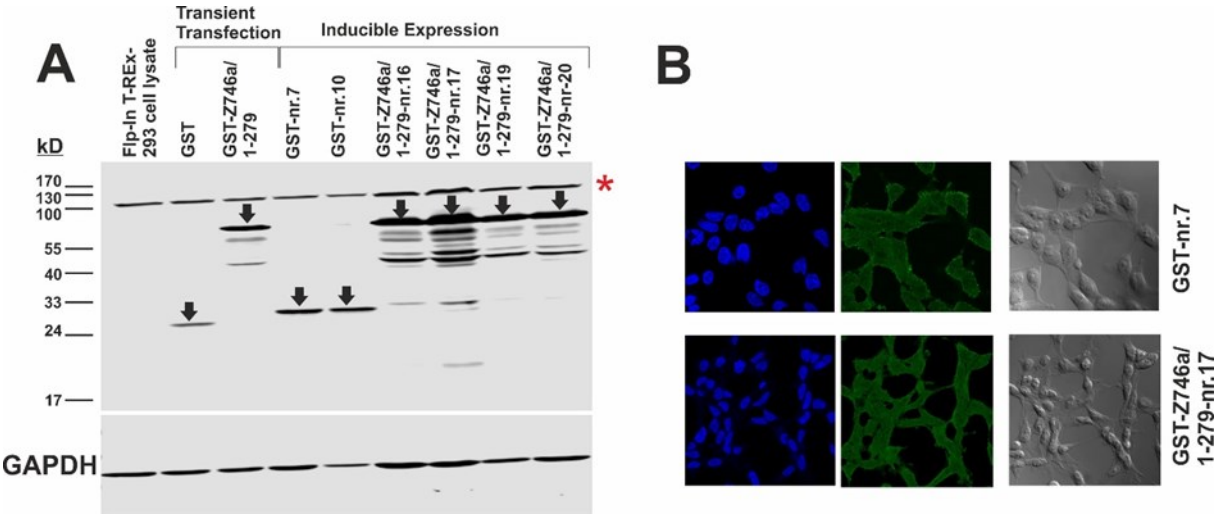

Fig. 13: Analysis of stable HEK293 cell lines expressing GST alone or GST-Z746a/1-279. (A) *Western blot analysis of total protein extracts from six clones that survived under the selection of hygromycin B (2 express GST and 4 express GST-Z746a/1-279) (see methods) alongside extracts from parent cell line Flp-in T-REx-239 cells (negative control) and transiently transfected parent cells expressing GST or GST-Z746a/1-279 (positive controls). Extracts made with 1x SDS sample buffer after a 24-hour induction of expression with 2µg/ml Tetracycline and immunostaining with anti-GST (upper blot) and anti-GAPDH (lower blot) antibodies. Protein bands of the expected size are indicated by arrows. * indicates unspecific bands recognized by the antibodies.* ***(B)*** *indirect immunofluorescence staining for GST in two positive colonies; nr.7 expressing GST (above) and nr.17*

expressing GST-Z746a/1-279 (bottom). The cells were allowed to grow onto glass coverslips that pre-coated with poly-L-lysine. 24h post-induction, the cells were fixed and stained with monoclonal antibodies against GST (in green), polyclonal antibodies against GST (in red), and DAPI (4'-6-diamidino-2-phenylindole) to stain the cell nuclei (in blue). The fluorescence images show that GST and GST-Z746a/1-279 localize to cytoplasm and nuclei almost homogeneously.

For the expression of recombinant proteins in E. coli, the bacteria were transformed with tagged DUF3669 of ZNF746 and ZNF777. The Western blotting analysis was used to evaluate expression levels. Different volumes of crude extracts were loaded into the gel to optimize the volumes needed for pulldown assays. All tagged proteins showed reliable bands but the expression efficiencies obviously varied.

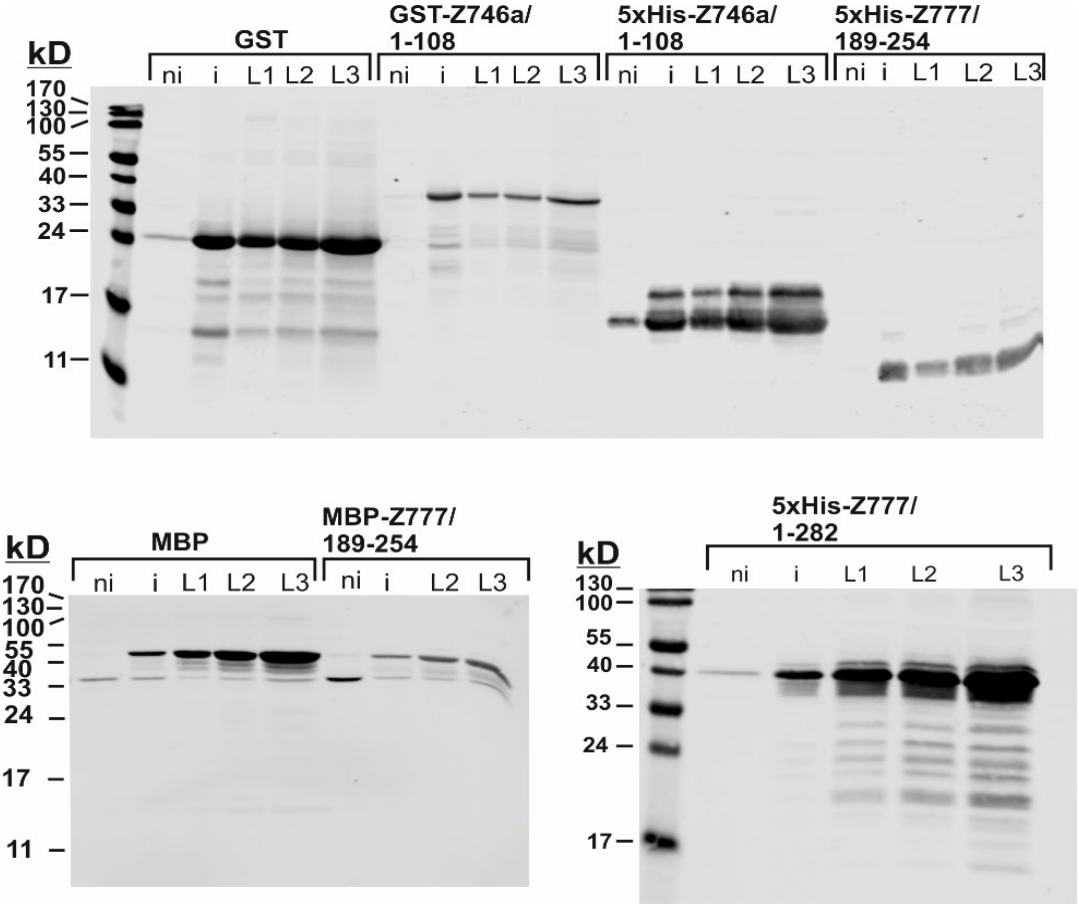

Fig. 14: Expression of recombinant DUF3669 in E. coli. *SoluBL21 (DE3) competent E. coli were transformed with 10ng of the indicated prokaryotic expression vector. The protein expression was induced by adding 0.1mM IPTG to the bacterial suspensions at OD600= 0.4-0.6. Non-induced bacteria were lysed with 1xSDS sample buffer (ni lanes, 25µl), bacteria suspension after 4h induction with IPTG were lysed with 1xSDS sample buffer (i lanes, 25µl). The induced bacteria were briefly washed with 1x ice-cooled PBS, resuspended with lysis buffer and lysed by sonication. Different amounts of bacterial lysates were loaded into SDS-polyacrylamide gel (L1: 4µl, L2:8µl, L3:20µl).*

5.3- Contribution of the protein domain setup of ZNF746 to its transcriptional repression activities

To assess the impact of the N-terminal domain setup on transcriptional repression specific expression constructs of ZNF746 (two isoforms with respect to truncated or complete KRAB-A subdomain) were tested using a classical heterologous reporter assay. ZNF10 (also known as KOX1) was used as a positive control because it is well known as a robust transcriptional repressor through its KRAB-A box (Margolin et al., 1994). For simplicity, the term "**amino half**" of ZNF746 will be used to denote ZNF746 without zinc finger motifs corresponding to the configuration; DUF3669-KRAB-Exons 5 and 6. Tables contain repression factors of all tested constructs used to generate the bar charts in this study are found in supplements S3-S20.

The truncated KRAB domain of ZNF746a (aa 94-173) did not exhibit any repression activity, while the amino half (aa 1-279) displayed the highest repression activity among all tested ZNF746a-derived fragments indicating the necessity of all amino-terminal domains for reaching the efficient repression activity (Fig. 15B). In addition, the DUF3669-containing segment of ZNF746a without further sequences (aa 1-108) caused downregulation of 2.7-fold. The protein region of ZNF746a corresponding to DUF3669 followed by the KRAB domain caused slight transcriptional repression (1.4-fold) and the ZNF746a portion that hooked up KRAB to the exons 5/6 caused downregulation of 2.7-fold. It is noteworthy that full-length ZNF746a exhibited transcriptional repression of 3.1-fold in HeLa and 3.9-fold in HAP1, which is less than the fold changes caused by its amino half (aa1-279). One reason might be expression disparities of effector plasmids in regard to protein levels since the expression analysis of 2µg pM3-Z746a by Western blotting showed a weaker band than that of pM3-Z746a/1-279 (Fig. 11 and Fig. 12). Moreover, the protein region of zinc finger motifs caused a weak downregulation of 2-fold.

On the other side, the KRAB domain isoform with non-truncated KRAB-A of ZNF746b (aa109-188) caused downregulation of 3.1-fold (Fig. 15C). The amino half of ZNF746b (aa 1-294) exhibited the most potent transcriptional repressive activity amongst all tested effector plasmids of this isoform (12.5-fold). It reached about 70% of repression potential achieved by the robust repressor KRAB domain of ZNF10 alone. Significant enhancement in repression activities was also shown by the protein regions of ZNF746b corresponding to DUF3669-KRAB and KRAB-exons 5/6, 5.2- and 8-fold respectively.

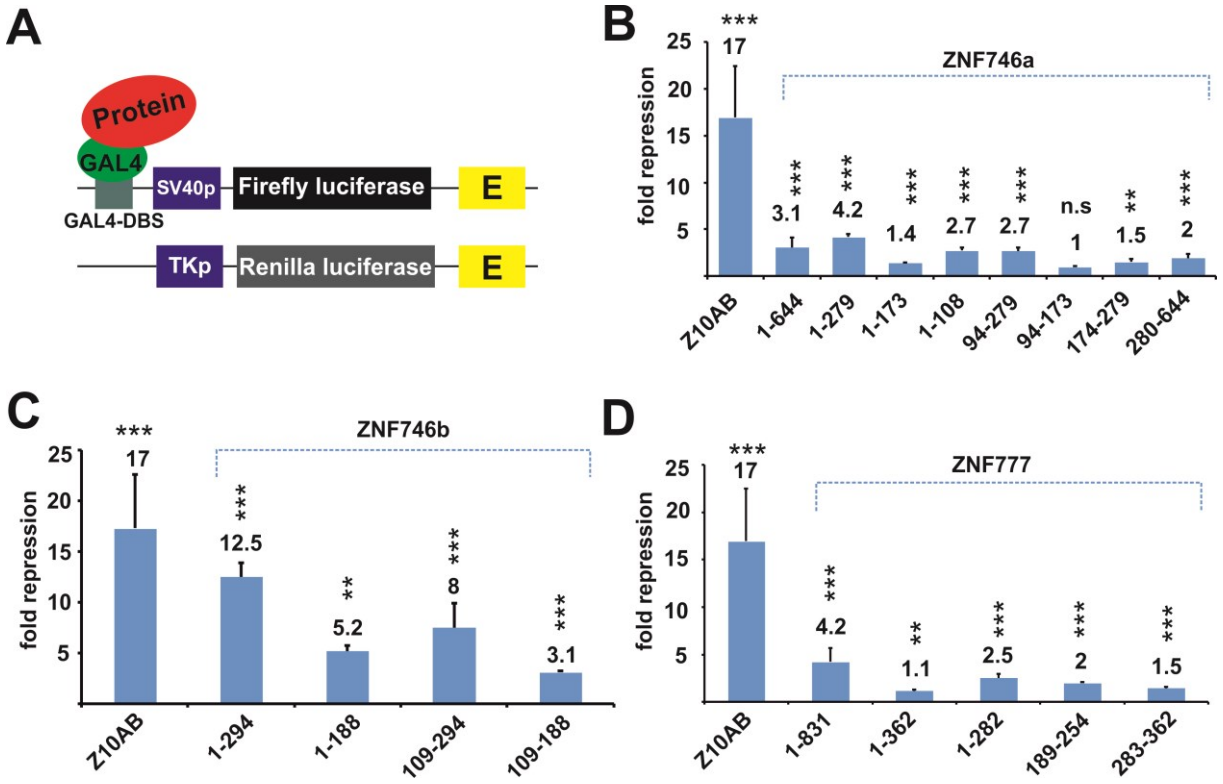

Fig. 15: Contribution of different protein domains to the transcriptional repressor activities of ZNF746 and ZNF777. *Heterologous reporter assay data using fusions of the indicated protein segments with the Gal4 DBD in human HeLa cells. The fusion of the ZNF10 KRAB domain (KRAB-AB) was used as a positive control and reference for potent repressor activity. This reference is included in each subfigure plot.* ***(A)*** *Illustration of assay; only the firefly luciferase reporter plasmid harbors upstream Gal4 DNA-DBS and its expression was normalized to the expression of Renilla luciferase plasmid that does not have GAL4-DBS.* ***(B)*** *Repressor potency of isoform ZNF746a and its selected segments.* ***(C)*** *Repression factors of segments derived from ZNF746b.* ***(D)*** *Repressor activity of ZNF777 and its selected segments. Bar plots represent normalized mean repression factor values ± STDEV of six biological replicates from three independent experiments relative to experiment-specific values for the Gal4-DBD alone. Asterisks indicate results of a two-tailed unpaired t-test (one asterisk * $p<0.05$, two asterisks ** $p<0.01$ and three asterisks *** $p<0.001$).*

These findings showed that the overall repression activity of ZNF746 is not attributed to the KRAB domain alone. Even though the KRAB domain of ZNF746a is inactive as a transcriptional repressor; the presence of DUF3669 and the protein region encoded by exons 5 and 6 lead to the establishment of an N-terminal domain array with moderate transcriptional repression capacity. Full-length KRAB domain of ZNF746 is a functional transcriptional repressor, and the extension by DUF3669 and/or exons 5/6 regions exhibited significant increments of transcriptional repressive potential compared to their ZNF746a counterparts.

In brief, reporter assay data showed that the whole N-terminal configuration of DUF3669, KRAB, and exons 5/6 segment interdependently contribute to the repression potential of ZNF746. It is quite clear that the repressive isoform of the KRAB domain is more influenced by the presence of DUF3669 and/or exons 5/6 than the non-repressive KRAB isoform.

5.4- Impact of the ZNF777 domain organization on transcriptional repression

ZNF777 contains the amino-terminal DUF3669 in addition to a KRAB domain with a complete KRAB-A subdomain and nine zinc finger motifs (Fig. 10C). The KRAB-AB domain of ZNF777 (aa283-362) caused a neglectable downregulation even though its KRAB-A subdomain is non-truncated and matches the KRAB-A of ZNF746b (Fig. 15D, Fig. 17C blue bars). Furthermore, the configuration DUF3669-KRAB (aa 1-362) in ZNF777 caused a weak downregulation of 2-fold only in HAP1 cells (Fig. 17C). The DUF3669-containing segment (aa1-282) caused a weak repression activity of 2.5-fold in HeLa and 4.1-fold in HAP1. The part of the amino terminus that matches the consensus of DUF3669 (PF12417) (aa 189-254) caused a weak downregulation of 2-fold in HeLa cells and 1.4-fold in HAP1 cells. Full-length ZNF777 (aa1-831) showed with 4.2- (HeLa) and 6.3-fold (HAP1) a clearcut and compared to ZNF10AB moderate transcriptional repression potential although its KRAB-AB alone was a very poor transcriptional repressor. The last result indicates again (as shown for ZNF746) that ZNF777 requires the DUF3669 domain and maybe the exon 4 encoded residues for its repressive activity.

DUF3669 of ZNF746 and ZNF777 are transcriptional repressors. By comparison, the protein region of ZNF777 that matches the consensus of DUF3669 (PF12417) (aa189-254) failed to preserve the repression potency of the longer version of DUF3669-containing fragment (aa1-282). This indicates that this sequence is insufficient to reach a full-function as a transcriptional repressor, Fig.24 shows alignment of all human DUF3669 domains.

The poor repression activity of the KRAB domain of ZNF777 although its KRAB-A subdomain is similar to the KRAB-A subdomain of ZNF746b (Fig 16), could be explained by the differences in individual residues, maybe in the KRAB-B domain that can stimulate KRAB-A-repressive activity (Vissing et al., 1995).

```
Z746a_KRAB-A      WILRLPPGSKGESPKEWGKLEDWQKELYKHVMRGNYETLVSLD
Z746b_KRAB-A      VPVTFDDVAVYFSEQEWGKLEDWQKELYKHVMRGNYETLVSLD
Z777_KRAB-A       VPVTFDDVAVHFSEQEWGNLSEWQKELYKNVMRGNYESLVSMD
                    : :   :   * :***:*.:*******:*******:***:*

Z10_KRAB-A        TLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSLG
                  . ***.** * *:.:**  *.  *: :*::**  **:.***:.
```

```
Z746a_KRAB-B        YAISKPEVLSQIEQGKEPCNWRRPGPKIPDVPVDPSP
Z746b_KRAB-B        YAISKPEVLSQIEQGKEPCNWRRPGPKIPDVPVDPSP
Z777_KRAB-B         YAISKPDLMSQMERGERPTMQEQEDSEEGETPTDPSA
                    ******:::**:*:*:.*    .: ..:   :.*.***.

Z10_KRAB-B          YQLTKPDVILRLEKGEEPWLVEREIHQETHP------
                    * ::**::: ::*:*:.*    .:    :   .
```

Fig.16: Multiple alignment of amino acid sequences of KRAB-A subdomains and KRAB-B subdomains of ZNF746b, ZNF777, and ZNF10. *Multiple alignment was conducted using Clustal tool, ''W'' version. (https://www.genome.jp/tools-bin/clustalw). The color code according to chemical properties; polar in green (G, S, T, Y, C), neutral in purple (Q, N), basic in blue (K, R, H), acidic in red (D, E), and hydrophobic in black (A, V, L, I, P, W, F, M). The highlighted polypeptide in yellow displays the amino acids from the neighbouring protein region to fill the truncated region of the KRAB-A in the isoform ZNF746a.*

5.5- Evaluation of the dependency of ZNF746 and ZNF777 repressor potential on TRIM28 and SETDB1

The prevailing view is that interaction with TRIM28 is indispensable for a KRAB domain to function as a transcriptional repressor. TRIM28, in turn, recruits SETDB1 (histone methyltransferase) and NuRD complex (contains histone deacetylase activity) to the genomic locus target and subsequently confers epigenetic modulation and transcriptional repression (Fig. 3).

This stage of work aimed to investigate to what extent the transcriptional repression activities of constructed protein regions of ZNF777 and ZNF746 are dependent on the presence of TRIM28 and SETDB1. Dual reporter assays were performed in HAP1 wildtype, HAP1-TRIM28 knockout, and HAP1-SETDB1 knockout cells (Fig. 17). The canonical KRAB domain of ZNF10 is used as a positive control in all cell lines because ZNF10 was originally utilized to isolate TRIM28, and the interaction with TRIM28 has long been well-characterized (Friedman et al., 1996). Lists show the repession factors of all tested constructs in HAP1 cells are found in supplements S7-S15. As expected, the robust transcriptional repression caused by the KRAB domain of ZNF10 in HAP1 wildtype was completely abolished in HAP1-TRIM28ko cells and drastically declined in HAP1-SETDB1ko.

In terms of ZNF746a-derived protein fragments, as observed in HeLa cells, the highest repression activity was caused by the amino half that corresponds to DUF3669-KRAB-exons5/6 (7.2-fold, Fig. 17A). This repression activity decreased by almost half in TRIM28 knockout cells (3.8-fold) suggesting that their transcriptional repression function dependent on the presence of TRIM28, but not completely. The dependency on SETDB1 can also be

observed by a significant decrease in repression activities in SETDB1 knockout cells (2.3-fold). Unlike ZNF746a-amino half, the transcriptional repressive activity of full-length ZNF746a appears to be independent of TRIM28 and SETDB1. The repression activities caused by the ZFs-containing fragment of ZNF746 (aa 280-644) appear to be independent of TRIM28 but dependent on SETDB1. Moreover, DUF3669 of ZNF746 and ZNF777 displayed dependency on TRIM28 and SETDB1.

The full-length KRAB domain of ZNF746 showed complete dependency on TRIM28 and SETDB1 (Fig. 17B). Moreover, the repression activities caused by all amino-terminal segments of ZNF746b declined significantly in TRIM28ko and SETDB1ko cells suggesting that their repressive activities are dependent on the presence of TRIM28 and SETDB1. Of note, robust repressive activity caused by amino half of ZNF746b did not reach the basal expression levels in knockout cells. Interestingly, despite demonstrating distinct transcriptional repressive capacities in wild type cells, the amino halves of both isoforms (Z746a/1-279 and ZNF746b/1-294) showed very comparable values in knockout cells (Fig. 17A and B). In addition, both amino halves of ZNF746 showed higher repression activity in TRIM28ko than in SETDB1ko cells. Altogether, the results in HAP1 cells suggest a TRIM28-independent mechanism might also function, in addition to the canonical KRAB/TRIM28 pathway.

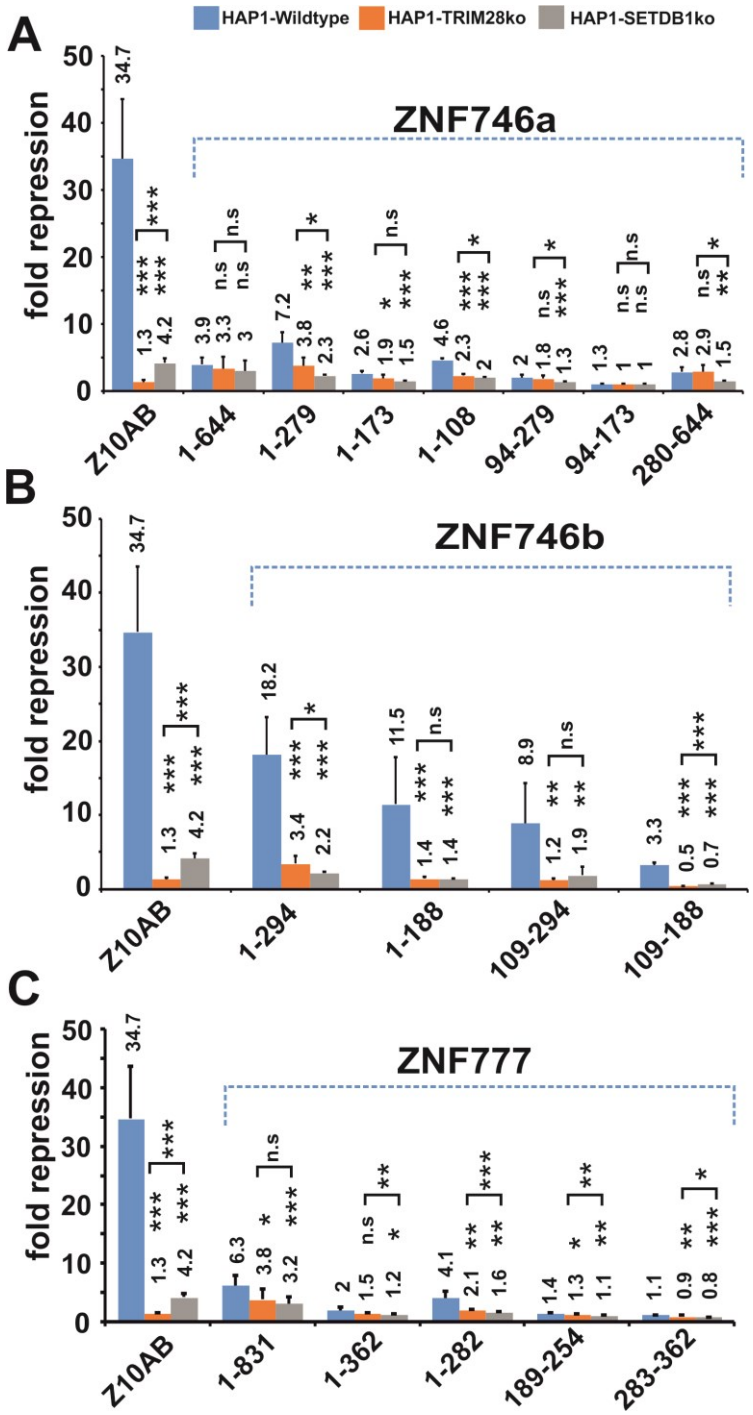

Fig. 16: Dependency of the transcriptional repressive activities of ZNF746- and ZNF777-derived protein segments on TRIM28 and SETDB1*. Results of dual-luciferase reporter assays using Gal4-DBD fusions of the indicated protein segments. The respective ZNF10 KRAB domain (KRAB-AB) construct served as a positive control for a KRAB repressor dependent on TRIM28 and SETDB1.* ***(A)*** *ZNF746a-derived protein fragments* ***(B)*** *ZNF746b-derived protein fragments* ***(C)*** *ZNF777-derived protein fragments. For each GAL4 fusion protein, the transcriptional repressive activities were evaluated in human HAP1 wild type (blue bars), human TRIM28 knockout HAP1 cells (orange bars), and human SETDB1 knockout HAP1 cells (grey bars). Bars depict normalized mean repression factor values ± STDEV of six biological replicates from at least three independent experiments relative to experiment-specific and cell-type-specific values for the Gal4-DBD alone. Asterisks indicate results of a two-tailed unpaired t-test (one asterisk * $p<0.05$, two asterisks ** $p<0.01$ and three asterisks *** $p<0.001$; n.s. = no statistical significance, i.e. $p>0.05$) when comparing knockout with wildtype data of a protein fragment.*

Full-length ZNF777 caused downregulation of the reporter gene by 6.3-fold in HAP1 wildtype cells. ZNF777 caused also repression activities in TRIM28ko cells (3.8-fold repression) and SETDB1ko cells (3.2-fold repression) suggesting that ZNF777-mediated repression activity is partially dependent on TRIM28

In brief, the presence of both DUF3669 and exons 5/6 in the amino halves of ZNF746 isoforms caused a significant enhancement in the repression activities not only in TRIM28-dependent manner but also through TRIM28-independent mechanism. Obviously, full-length ZNF746a and ZNF777 behave differently from their respective protein regions. Full-length ZNF746a appears to be independent of TRIM28 whereas full-length ZNF777 is partially dependent on TRIM28 and SETDB1. These data are not easy to interpret since full-length ZNF746a and ZNF777 contain functional zinc finger domains that bind to chromosomal regions whereas the GAL4-DNA binding domain binds to the GAL4 binding sites present in the transiently transfected reporter plasmid. Furthermore, DUF3669 containing protein variants could even theoretically highjack functional endogenous TRIM28/SETDB1 complexes by forming DUF3669 containing homodimer or heterodimer complexes (see section 5-10).

5.6- Investigation of the interaction between ZNF746 or-/ ZNF777 protein segments and endogenous TRIM28.

Reporter assay data showed that the repression activities of several ZNF746- and ZNF777-derived protein regions significantly declined or even disappeared in the absence of TRIM28. Therefore, the next logical step was to investigate whether these fragments were able to form stable complexes with TRIM28.

First, a colocalization assay for all tested KRAB domains was performed (KRAB domains of ZNF10, ZNF746a, ZNF746b, and ZNF777). It relies on the appearance of nuclear foci for Gal4-KRAB fusions that can easily be checked for the enrichment of endogenous TRIM28 by fluorescence microscopy (Born et al., 2014). Bright foci in the same image pane of GAL4-KRAB of ZNF10 and ZNF746b (in green) and cellular TRIM28 (in red) were observed, but not for the KRAB domains of ZNF746a and ZNF777 (Fig. 18). This suggests that the KRAB domains of ZNF746b and ZNF10 can recruit TRIM28.

To investigate whether TRIM28 associates in the same complex with any of ZNF746- or ZNF777-derived protein regions, co-immunoprecipitation assay using protein G magnetic beads loaded with monoclonal antibodies against endogenous TRIM28 was performed. None

of GST-ZNF777 and GST-ZNF746a fusions co-precipitated with cellular TRIM28 (Fig. 19A and Fig.19C).

The ZNF746b-derived GST fusions corresponding to the KRAB domain (aa 109-188), DUF3669-KRAB (aa 1-188) and KRAB-exons 5/6 (aa 109-294) co-associated with endogenous TRIM28 (Fig. 19B). However, the extent of enrichment was clearly much lower than for the ZNF10 KRAB domain arguing that the stability of the complex and/or affinity of the binding partners were inferior for ZNF746b. Conversely, the amino half of ZNF746b (aa 1-294) could not form a stable complex with TRIM28 although its shorter versions showed significant association with TRIM28. This result was reproducible and confirmed when anti-GST antibodies were used as a bait.

Despite multiple tries with different transfection methods, it was not possible to get sufficiently efficient expression of full-length GST-ZNF746a and GST-ZNF777 in HeLa cells. Therefore, it was not possible to assess their interaction with TRIM28 in this study.

It stands to reason that the truncated KRAB-A with ZNF746a fragments failed to form a stable complex with TRIM28 because it lacks the well-conserved DV and LE motifs that are critical for interaction with TRIM28 whereas the KRAB domain of ZNF746b with complete KRAB-A showed a weak interaction with TRIM28 likely because of the presence of DV motifs (Margolin et al., 1994).

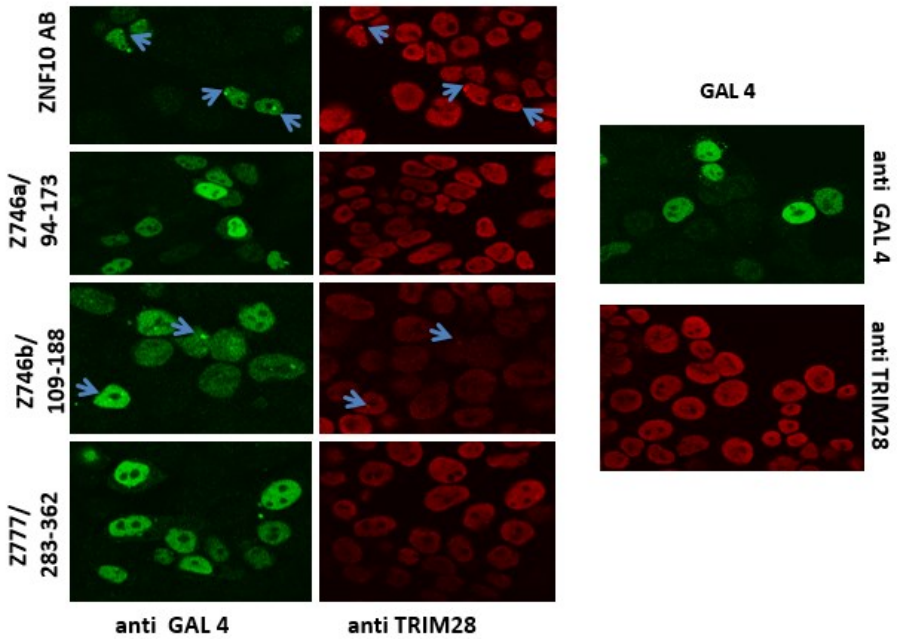

***Fig. 17**: **Co-localization analysis of ectopically expressed GAL4-KRAB versions and cellular TRIM28 in HeLa cells using immunofluorescence microscopy.** 24h post-transfection, HeLa cells*

were fixed in PBS (4%/w/v) paraformaldehyde, and permeabilized in PBS, 0.5% (w/v) Triton X-100. Subsequently, the cells were stained for GAL4-KRAB fusion proteins (in green) and cellular TRIM28 (in red). Both GAL4-KRAB and TRIM28 predominantly localize to the nuclei.

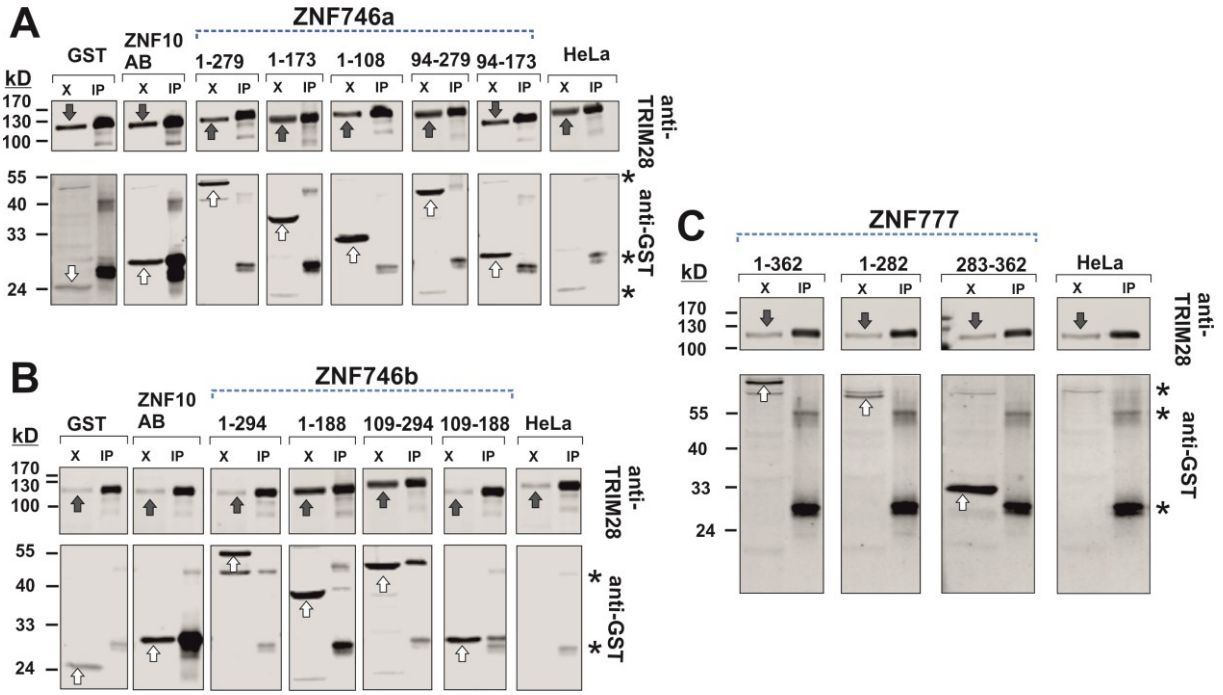

Fig. 18: Analysis of the interaction between ZNF746 and ZNF777 protein fragments and endogenous TRIM28. *Extracts from HeLa cells after transient expression of GST/GST-ZNF fusion constructs were subjected to immunoprecipitation using TRIM28 as bait. Input protein extracts ("X") and eluted immunoprecipitates ("IP") were analyzed by Western blotting. The blots were probed with monoclonal antibodies against TRIM28 and rabbit polyclonal antibodies against GST, followed by respective secondary antibodies with different fluorescent tags. Images represent black/white representations of cropped regions (see Methods) informative to evaluate TRIM28 precipitation (upper part) and GST/GST-fusion protein co-immunoprecipitation (lower parts) from the same lanes., White block arrows indicate GST/GST-fusion protein signals whereas the black block arrows point to TRIM28 bands in the input cell lysate lanes. GST alone was used as a negative control whereas the KRAB domain of ZNF10 served as a positive control. Further controls were obtained by mock immunoprecipitation from non-transfected HeLa lysates (label "HeLa") and protein-G beads eluted with SDS sample buffer. (**A**) Results for Z746a-derived GST fusions, (**B**) data of Z746b-derived GST fusions, (**C**) analysis of Z777-derived GST fusions. Asterisks indicate unspecific bands.*

5.7- Lack of a conserved glutamic acid residue in ZNF746 KRAB-A explains the weak repressor activity and insufficient interaction with TRIM28.

The canonical KRAB-A subdomains display conserved amino acid motifs that have been found to be essential for their function to confer transcriptional repression by TRIM28 recruitment. Accumulated evidence indicated that DV and MLE conserved sequences in the human KRAB-A consensus are essential for TRIM28 binding and thereby transcriptional repression in ZNF10 (Margolin et al., 1994; Friedman et al., 1996), Kid 1; rat zinc finger

protein (Witzgall et al., 1994); in KRAB-O (Peng et al., 2009), and mouse ZFP568 (Murphy et al., 2016).

Sequence analysis revealed that MLE motif is replaced by MRG in the KRAB-A of the isoform ZNF746b. The KRAB domain of ZNF746b (Z746b/109-188) displayed a weak downregulation of about 3-fold and weak interaction with TRIM28. Thus, we asked whether the absence of leucine (L), glutamic acid (E), or both are responsible for its weak transcriptional repressive capacity as well as its weak interaction with TRIM28. Substitution mutations were generated; Z746b/109-188_R141L, Z746b/109-188_G142E, and Z746b/109-188_RG/LE. The transcriptional repressive activity and the interaction with TRIM28 were tested with all mutant KRAB domains (Fig. 20).

The results showed that the mutant Z746b/109-188_R141L did not exhibit any improvement in either repression activity or its ability to interact with TRIM28. In contrast, the replacement of glycine with glutamic acid in the mutant Z746b/109-188_G142E caused a drastic increase in transcriptional repression activity to a factor of 41-fold repression and concomitantly significant recruitment of TRIM28. The double mutant Z746b/109-188_RG/LE did not improve transcriptional repressive activity or TRIM28 interaction further. The results suggested that the glutamic acid residue at its specific position within the MLE conserved motif was the most essential for KRAB-mediated repression activity by improving the interaction with TRIM28. One explanation might be its negative charge that can pair with positively charged amino acids to stabilize KRAB-TRIM28 interaction.

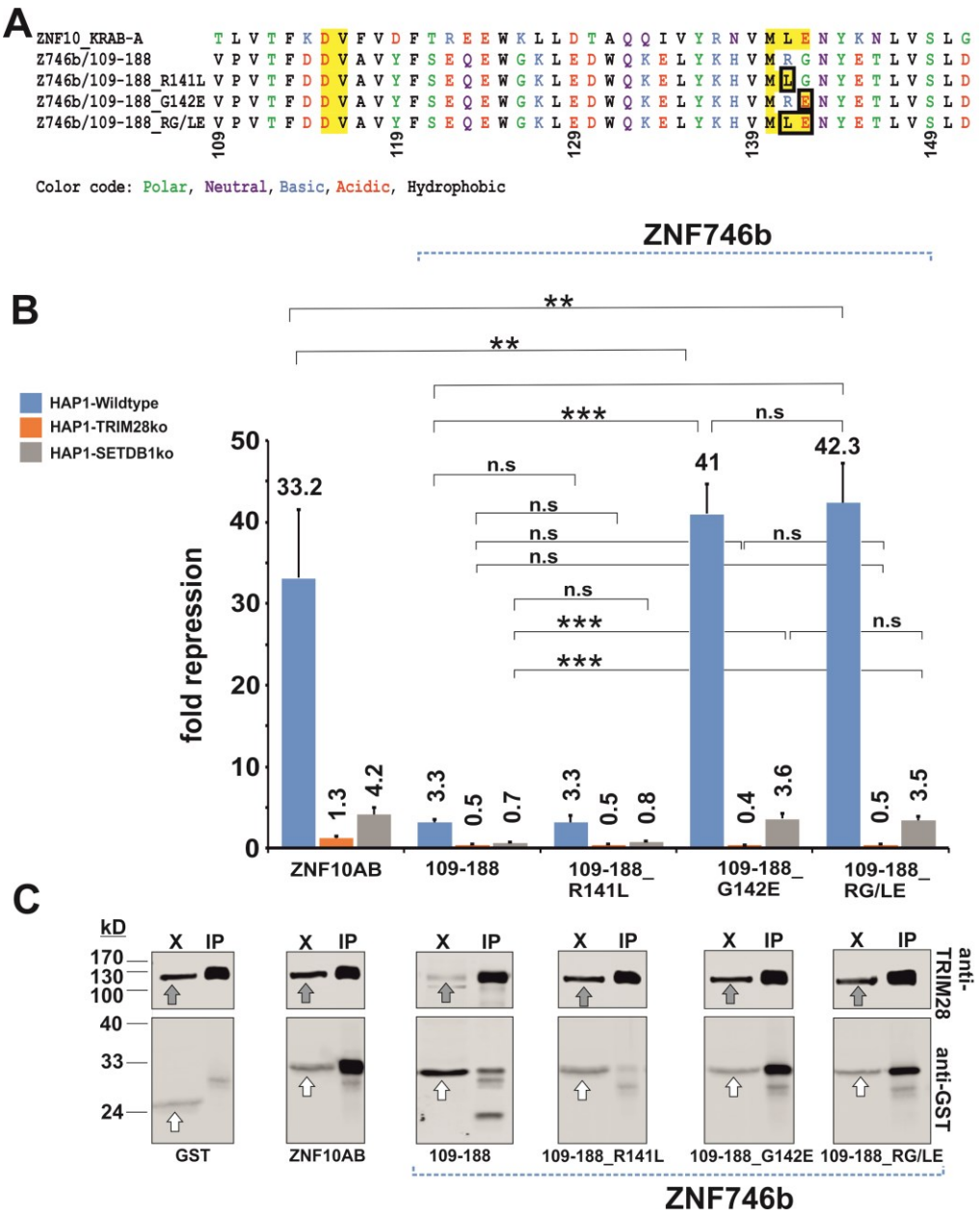

Fig.20: Mutation to a conserved acidic residue in KRAB-A strongly improved the repression activity and TRIM28 interaction of the ZNF746b KRAB-AB domain*. **(A)** Multiple alignment of the KRAB-A amino acid residues of ZNF10, ZNF746b and three ZNF746 mutants, R141L, G142E, and RG/LE. The amino acids highlighted in yellow indicate agreement with the amino acids that have been previously confirmed to be necessary for potent transcriptional repression activity and stable interaction with TRIM28. Mutated residues in ZNF746b are highlighted by black boxes. **(B)** Comparison of transcriptional repression activities of wild type KRAB-AB of ZNF746b, and the three configurations with the indicated mutated residues in KRAB-A in HAP1 wild type (blue bars), TRIM28ko (orange bars), and SETDB1ko cell lines (grey bars). ZNF10-AB is used as a positive control. Results of dual-luciferase assays presenting mean repression factor values ± STDEV of six biological samples from three independent experiments. Asterisks indicate results of a two-tailed unpaired t-test (* $p<0.05$; ** $p<0.01$; *** $p<0.001$; n.s.= not significant, i.e. $p>0.05$) **(C)** Analysis of complex formation of GST-KRAB-AB fusion proteins with TRIM28 in HeLa cells using co-immunoprecipitation and endogenous TRIM28 as bait as described in Fig. 19. Grey block arrows point to endogenous TRIM28 in the lysate lanes (X), and white block arrows point to the GST/GST fusions in the lysate lanes (X).*

5.8- Lack of transferability of the DUF3669 effect on repression to a canonical KRAB domain

The protein region of ZNF746b corresponding to DUF3669-KRAB (aa 1-188) showed a substantial enhancement in transcriptional repression as compared to its respective KRAB domain (Fig. 15C and Fig.17B); this level of repression improvement was not observed when the KRAB domain was not active as a transcriptional repressor as shown in ZNF746a (aa 1-173) and in ZNF777 (aa 1-362). This might imply that DUF3669 might fine-tune the transcriptional repressive potential of KRAB domains, particularly when the KRAB domain is a transcriptional repressor. Furthermore, transcriptionally inactive KRAB domains might even be favored in concert with DUF3669 domains to introduce and gain extended regulatory potential.

To examine whether DUF3669 could inherently support the transcriptional repressive potential of a potent canonical KRAB domain, the DUF3669 of ZNF746 (aa1-108) and ZNF777 (aa189-254) were chimerically linked to the amino terminus of the KRAB-AB domain of ZNF10. The transcriptional repression potentials of DUF3669-Z10AB were evaluated in HAP1 cells. As shown in Fig. 21A, the chimeric proteins did not show any significant changes in transcriptional repression potential. This finding suggests that DUF3669 functions as an enhancer of KRAB domain functionality in its authentic configuration. Of note, Z746a/1-108-Z10AB showed some residual repression activity in HAP1-TRIM28ko which might suggest that DUF3669 of ZNF746 can attenuate the dependency of transcriptional repressive activities on TRIM28.

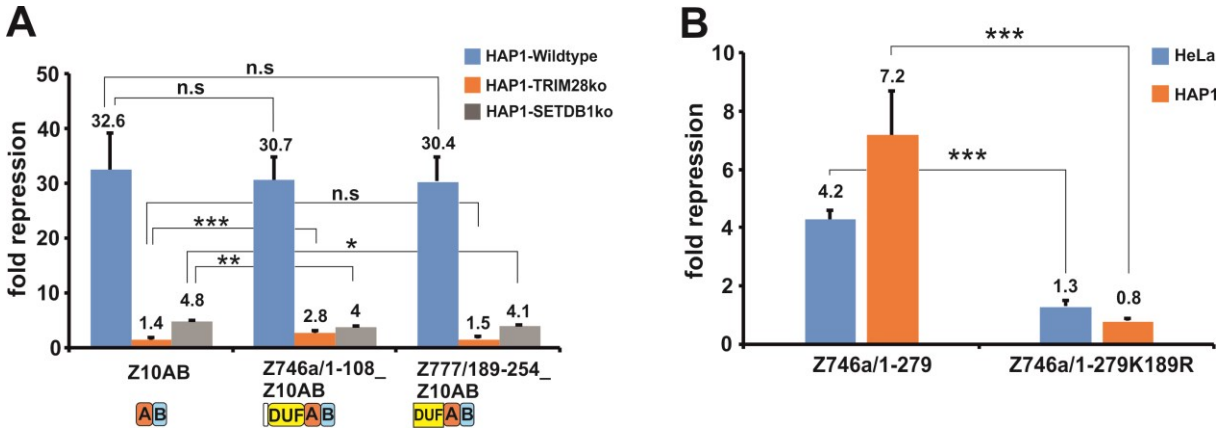

Fig.21: (A) *Impact of DUF3669 domains from ZNF746a and ZNF777 on the repressor activity of ZNF10-KRAB-AB. Wildtype ZNF10-AB and fusions of ZNF10-AB to DUF3669 domain segments from ZNF746a or ZNF777 were analyzed by dual-luciferase activities in HAP1 wildtype, SETDB1 and TRIM28 knockout cells as in Fig. 17.* ***(B)*** *Mutation of a major SUMO-acceptor lysine (K189R) in the Z746a/1-279 fragment abolished its repressor activity in HeLa cells (blue bars) and HAP1 wildtype*

*cells (orange bars). Dual-luciferase reporter assay as in Fig. 15. Bars visualize mean repression factor values ± STDEV of six biological samples from three independent experiments. Asterisks indicate results of a two-tailed unpaired t-test (*** $p<0.001$; n.s. = not significant, i.e. $p>0.05$).*

5.9- SUMOylation at K189 in ZNF746a is critical for its transcriptional repression activities

Nishida and Yamada (2016) reported that ZNF746 has two primary SUMOylation target sites (K189 and K286 in isoform a) and that the SUMOylation of ZNF746 regulates its repressive activity in a cell type-dependent manner (HeLa and HAP1cells were not studied) (Nishida and Yamada, 2016). K189 is located within the protein region of ZNF746 encoded by exon 5, whereas K286 is located within the amino terminus of the ZFs-containing fragment encoded by exon 7. Reporter assay results showed that the protein regions of KRAB-exons 5/6 in ZNF746a and ZNF746b caused a significant enhancement in their repressive activities.

To evaluate if the SUMOylation at K189 affects the transcriptional repression of ZNF746a, a K to R amino acid replacement mutation was generated at K189 in the ZNF746a/1-279 fragment. The transcriptional repression activities caused by this mutant protein were evaluated in HeLa and HAP1 cells (Fig. 21B). Compared to the wildtype ZNF746a/1-279 portion that caused reporter repression of 4.2-fold in HeLa and 7.2-fold in HAP1, the ZNF746a/1-297K189R mutant failed to exert a clearcut repressive activity (residual repression factor of 1.3-fold in HeLa and 0.8 in HAP1 cells). This strongly suggests that SUMOylation at K189 is crucial for ZNF746a repression activities in the studied cell models. The mutant ZNF746a/1-279K189R failed to achieve repressive activity although it still contains an intact DUF3669 domain. The abolishment of repression that resulted from the disruption of SUMOylation at lysine 189 predominates over the repression enhancement that is probably caused by the DUF3669 domain. Further investigation is still required to examine if the SUMOylation plays the same role in the other isoform of ZNF746.

5.10- Investigation of DUF3669 Homo- and Hetero-oligomerization

To investigate whether DUF3669 is able to oligomerize, we first interrogated data sets determined by BioPlex 2.0 for shared complexes between members of the DUF3669-KRAB-ZNF proteins. Both, ZNF746 and ZNF777, were found to associate with two DUF3669-containing KRAB-ZNF proteins (ZNF212 and ZNF 398) in large-scale affinity purification-mass spectrometry studies in HEK293 cells (Fig. 22A) (Huttlin et al., 2017). Since neither

ZNF746 nor ZNF777 was used as bait in these studies, existence in the same complex was not formally proven.

Self-association of ZNF212 has been previously reported (Gao et al., 2008). Importantly, the tandem-affinity purification method showed that ZNF746 interacts with ZNF212, ZNF282, ZNF746, and ZNF783. These ZNF proteins are DUF3669-containing proteins (Kang and Shin, 2015) (Fig.25). As these data suggested, we decided to prove whether these associations occur through DUF3669.

The stable cells expressing either GST or GST-Z746a/1-279 were tranfected with either One Strep tagged-Z777/1-362 (OST- Z777/1-362) or One Strep taggd-Z777/1-282 (OST-Z777/1-282) that harbor the amino acid sequence of DUF3669. Reciprocal pulldown assays were conducted using either Protein G magnetic beads loaded with polyclonal antibodies against GST or Strep-Tactin resin. As depicted in Fig. 22B, the amino half of ZNF746a reciprocally co-associated with ZNF777/1-362 and with ZNF777/1-282. The enrichments using Strep-Tactin resin were more efficient than the immunoprecipitations using protein G magnetic beads. Since DUF3669 is present in all tested fusions and the KRAB is not known to elicit oligomerization, these co-associations most likely occurred through DUF3669.

To consolidate this finding, HeLa cells were transiently transfected with expression plasmids of DUF3669 containing segments of ZNF746a (aa1-108) and ZNF777 (aa189-254 and aa1-282) (Fig. 22C). Self-association of ZNF746a/1-108 encompassing DUF3669 and co-association between ZNF746a/1-108 and both DUF3669-containing segments of ZNF777 (aa1-282 and aa189-254) were observed. In contrast, DUF3669 of ZNF777 failed to show self-association. To ascertain that the inability of DUF3669 of ZNF777 to form homo-oligomers was not due to the tag, an alternative experiment was performed that replaced the OST- with a GFP-tag (Fig. 22D). In conclusion, the results demonstrated that DUF3669 of ZNF746 is able to form homo- and heterooligomers, whereas DUF3669 of ZNF777 contributed to hetero-oligomers but failed to form homo-oligomers

In order to show physical direct homo- and hetero-association without the presence of additional eukaryotic cell proteins or the need for post-translational modifications, pull-down assays were done with recombinant DUF3669 protein extracts from E. coli. In agreement with the previous results from eukaryotic cells, DUF3669-containing segment of ZNF746 (aa1-108) was able to homo- and hetero-associate with other DUF3669-containing segments of ZNF777 (aa1-282 and aa189-254) (Fig. 23). Further, DUF3669 of ZNF777 was once more not able to form homo-oligomers.

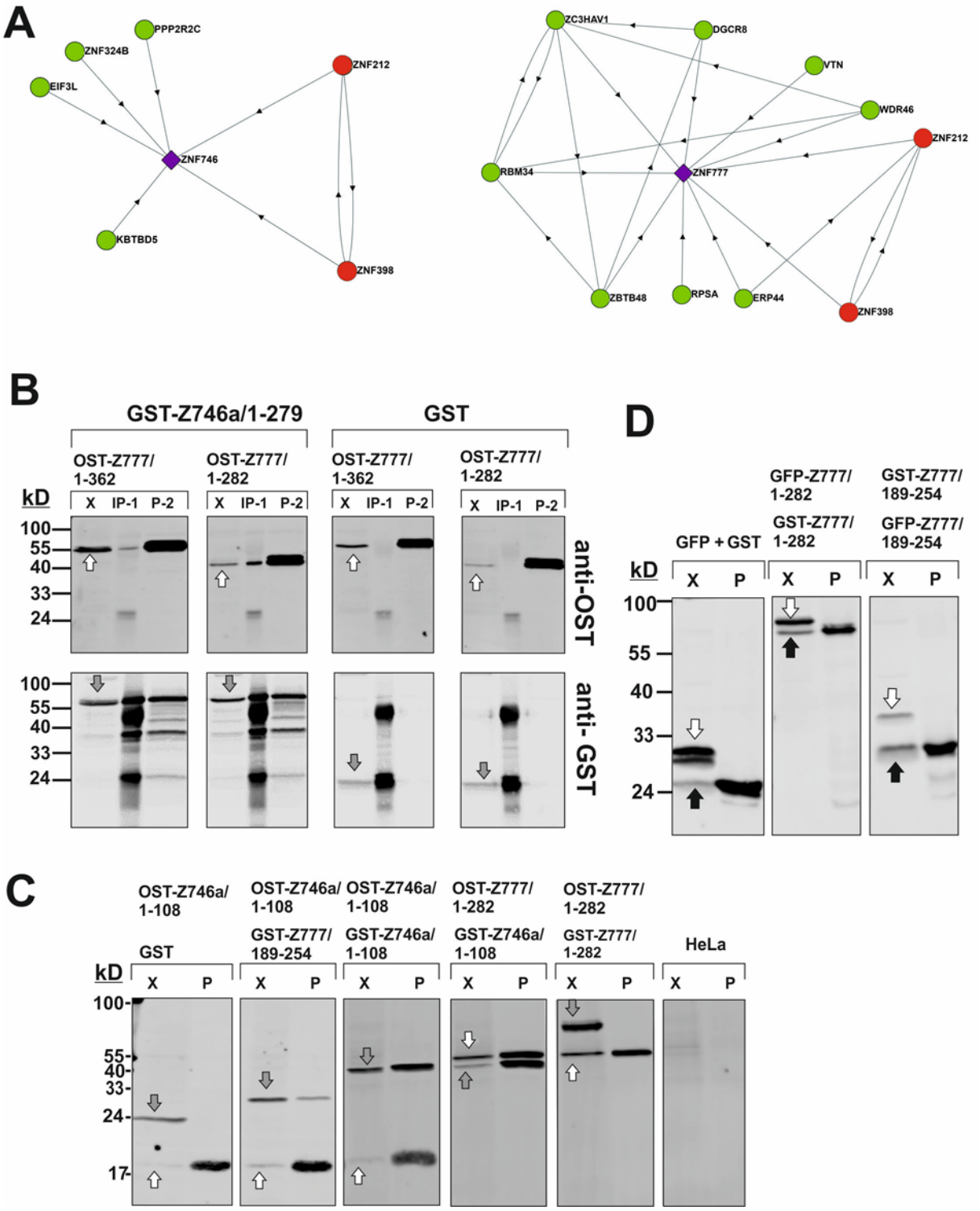

Fig. 22: In vivo evaluation of homo- and hetero-oligomerization of DUF3669 of Z746 and Z777. ***(A)*** *First-neighbor networks from the BioPlex2.0 project (Huttlin et al., 2017) for ZNF746 (left) and ZNF777 (right) indicate shared complexes of different DUF3669-containing ZNF proteins in HEK293 cells. A purple rhombus denotes the center of the network (ZNF746 and ZNF777, respectively), and the circle represents a bait node. DUF3369 -containing ZNF proteins are represented in red circles. Note that both, ZNF746 and ZNF777, were found to form complexes with DUF3669-containing ZNF proteins ZNF212 and ZNF 398.* ***(B)*** *Analysis of oligomerization between N-terminal fragments of ZNF746 and ZNF777. HEK293 cells that stably express either GST or GST-Z746a/1-279 were transiently transfected with either Z777/1-362 or Z777/1-282 tagged with a One-Strep-Tag (OST). Proteins enriched by either immunoprecipitation with anti-GST rabbit polyclonal antibodies (lanes*

labeled "IP") or by pull-down using Strep-Tactin via the OST (lanes "P") were analyzed by Western blotting alongside an aliquot of the input extract (lanes "X") with GST and OST antibodies as indicated. White arrows point to GST bands whereas grey arrows point to OST bands in the input lanes (X lanes). ***(C)*** *Narrowing down of the protein oligomerization domains and analysis of homo-oligomerization. HeLa cells were transfected with the indicated GST/GST-fusion and OST-fusion constructs and subjected to Strep-Tactin pulldown via the OST-tagged proteins. Western blots interrogated the input extracts and the enriched fractions (lanes labeled "X" and "P", respectively) using anti-GST and anti-OST antibodies in different color channels.* ***(D)*** *An alternative analysis of ZNF777 homo-oligomerization using HeLa cells transfected with GST/GST-fusion and GFP/GFP-fusion constructs. Enriched proteins from GST pulldown via glutathione beads (lanes "P") were probed with anti-GST and anti-GFP antibodies. White arrows point to GST bands whereas black arrows point to GFP bands in the input extracts (lanes "X").*

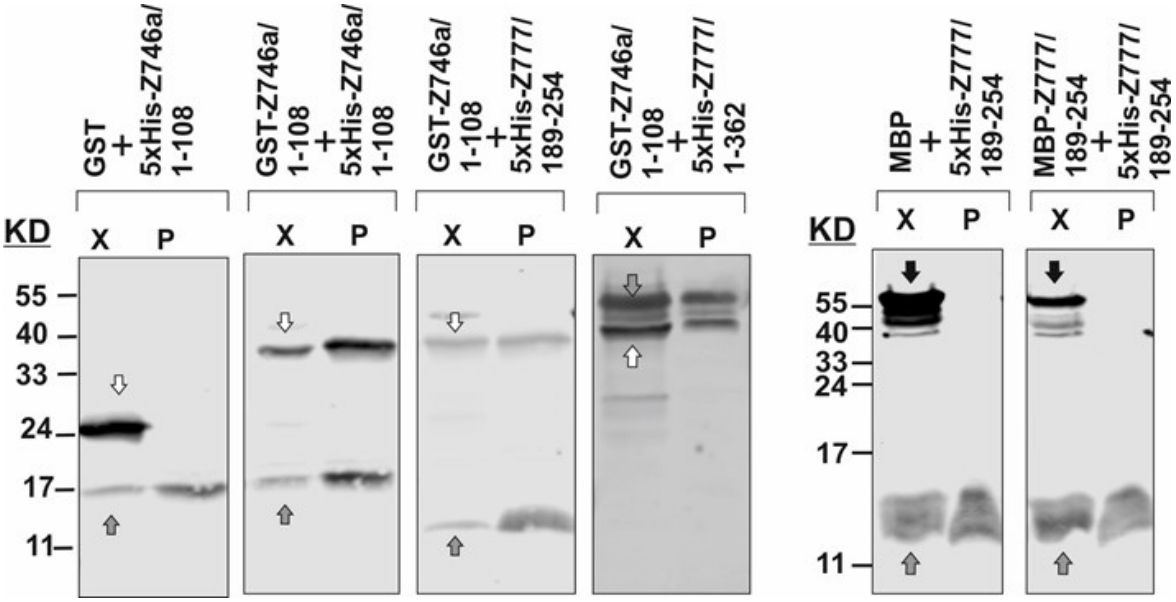

Fig. 19: Evaluation of homo- and hetero-dimerization of DUF3669-containing fragments of ZNF746 and ZNF777 in E. coli. *Bacterial extracts with the indicated recombinant proteins were mixed and Ni-NTA beads were used to pull-down His-tagged proteins. Western blots probed with anti-His tag antibodies and GST or MBP tag antibodies were used to visualize the proteins in the input extracts (lanes "X") and to identify the enriched recombinant proteins. GST alone and MBP alone were used as negative controls. Grey arrows denote His-tagged proteins, white arrows point to GST and GST fusion bands and black arrows point to MBP and MBP fusion proteins.*

These findings establish DUF3669 as a novel protein oligomerization domain and suggest that it has a role in regulating the assembly of complexes with other proteins that contain DUF3669.

6- Discussion

KRAB-ZNFs are known for their powerful transcriptional repressor potentials (Margolin et al., 1994; Pengue et al., 1994; Witzgall et al., 1994; Pengue and Lania, 1996); however, their endogenous targets and biological functions appear to be essential in tetrapod physiology. Although the canonical KRAB/TRIM28 model has been well characterized, the current knowledge about the transcriptional repression activities of individual KRAB-ZNFs is limited to a few well-characterized KRAB-ZNF proteins. Conversely, several studies reported the involvement of several KRAB-ZNF proteins in transcriptional activation (W. Chen et al., 2019; Hallen et al., 2011; Losson and Nielsen, 2010; Mysliwiec et al., 2007). Moreover, some members also failed to stably interact with TRIM28 despite trans-repressor activity (Itokawa et al., 2009; Murphy et al., 2016).

This study investigated two members of DUF3669-containing KRAB-ZNF subfamily; ZNF746 and ZNF777 in terms of the repression activities of their unusual amino termini including their protein oligomerization behavior. The reporter assay results showed that DUF3996 present in ZNF746 and ZNF777 behave as intrinsic transcriptional repressors. In accordance with the results in this study, the protein region that is now known as corresponding to DUF3669 in the paralogous ZNF282/HUB1 was reported to have transcriptional repression activity (Okumura et al., 1997). Our analysis suggests that the repression activity of DUF3669 is a conserved feature of this domain.

Here, this work established DUF3669 amino acid sequences as protein interaction interfaces for oligomerization. While the DUF3669 of ZNF746 forms homo- and hetero-association, DUF3669 of ZNF777 forms hetero-association but not homo-association. Just recently, during the writing of this thesis, Helleboid and coauthors published their work in which they reported that the DUF3669 of ZNF282 and ZNF398 co-associate in the same protein complex, however, without showing direct physical interaction (Helleboid et al., 2019). DUF3669 domains of ZNF746 and ZNF777 have a very similar amino acid sequence but they behave differently in terms of their ability to form homo-oligomers. This suggests that only a few amino acids in DUF3669 of ZNF777 might be responsible for the missing of self-association property. Two versions of DUF3369-containing fragments of ZNF777 were tested; one matches Pfam12417 and the other is encoded by the entire exon1 (Fig. 24).

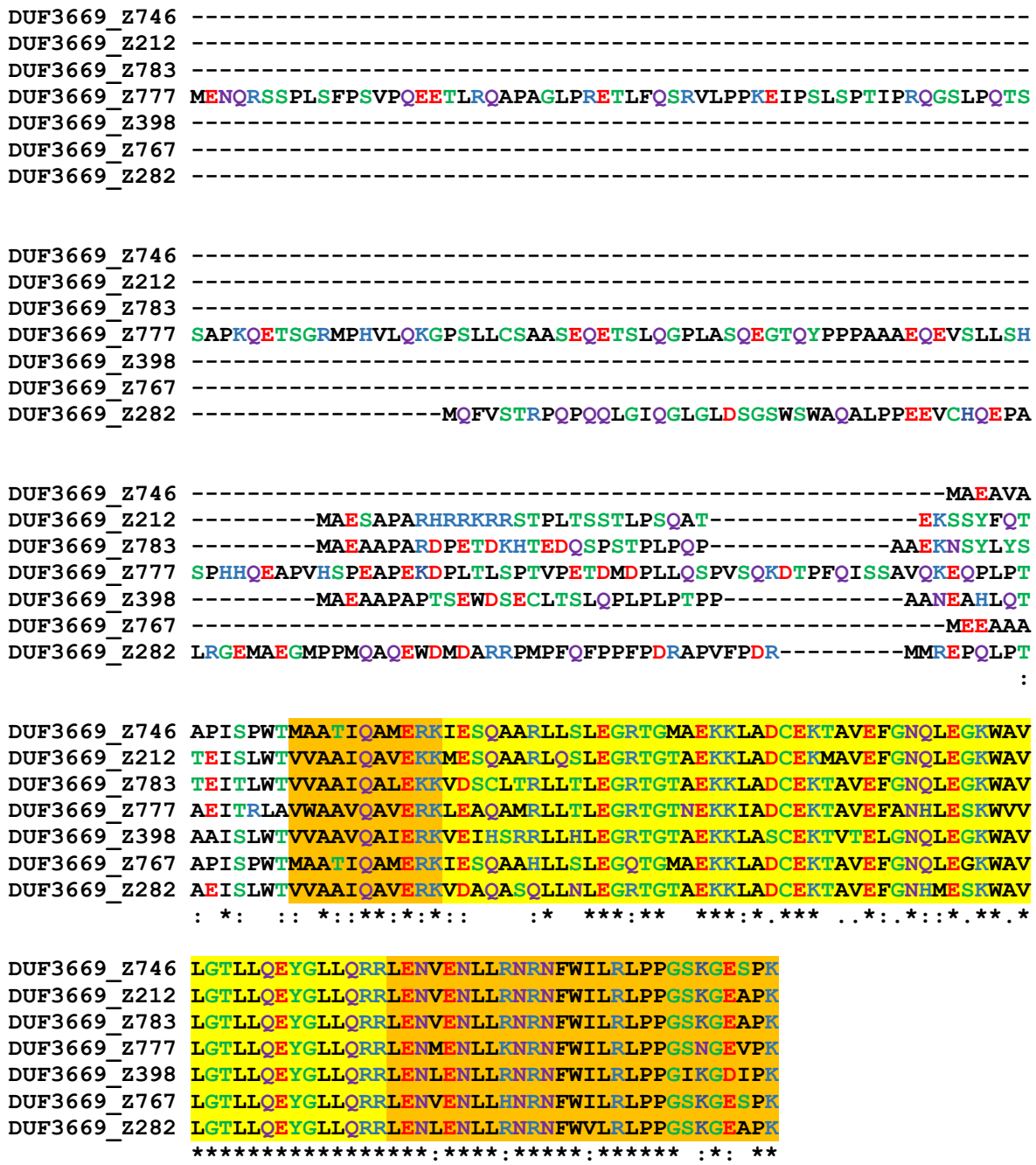

Fig. 24: Multiple alignment of human DUF3669-containing fragments encoded by single exons. *Alignment was performed and the color code as in the Fig.16. Amino acids highlighted in yellow represent the amino acids that match Pfam12417. Amino acids highlighted in orange represent the highly conserved fragment that might be an extension for DUF3669.*

Reporter assay results revealed significant differences in reporter gene repression, this suggests that PF12417 lacks some functional residues. Sequence analysis of DUF3669 containing orthologous proteins revealed an extremely conserved 95aa consensus of DUF3669 was identified that is longer than Pfam12417. Therefore, the disparities in

repression activities between two DUF3669-containing segments of ZNF777 would suggest that this missing conserved sequence (aa255-282) contributes to the DUF3669 repression function.

The oligomerization feature of DUF3669 can explain the increased repression activity of the amino-terminal segments of two ZNF746 isoforms and ZNF777 as compared with their respective KRAB domain. The repression increment was observed especially when the DUF3669 was followed by the non-truncated repressor KRAB domain of ZNF746b. It is noteworthy, that a heterologous transfer of DUF3669 failed to further augment transcriptional silencing mediated by potent canonical KRAB-AB domains like the one from ZNF10. The potent canonical KRAB-AB domain might be too dominant to observe any effects. Based on this, the DUF3669 can form higher-order molecular assemblies. Such complexes might be instrumental for function and open possibilities for context-dependent regulation if different DUF3669 members interact at distinct scenarios. Oligomerization thus may offer increased stability or provide combinatorial binding site signatures with respect to DNA binding. In addition, the high evolutionary conservation of DUF3669 sequences within an ortholog group and even between paralogs underscores the importance of DUF3669 for the ZNF746 and ZNF777 biological functions implying the presence of comparable synergistic functionalities.

In addition to the contribution of DUF3669 in the transcriptional repression activity, a recent study highlighted that DUF3669 of ZNF398 is critical for its function. The isoform of ZNF398 that lacks DUF3669 enhances the ubiquitination and proteasomal degradation of P53 through its KRAB domain, while the other isoform that has the amino-terminal DUF3669 does not affect P53 ubiquitination (Huang et al., 2019).

A comparison of the SCAN domain and DUF3669 identifies several astonishing common characteristics. Both domains are evolutionarily conserved domains and found in the amino termini of several KRAB-ZNFs (Huntley et al., 2006; Imbeault et al., 2017), The SCAN domain is an oligomerizing domain (Schumacher et al., 2000; Williams et al., 1999; Wu et al., 2003) The contribution of the SCAN domain to the repression activity has been investigated in few studies. In one example, the oligomerization state of the SCAN domain of ZNF202 was shown to be required for the interaction with TRIM28 and the transcriptional repression (Porsch-Ozcurumez et al., 2001). It was also reported that some SCAN-KRAB-ZNF proteins were transcriptional repressors although their respective isolated KRAB domain failed to downregulate the reporter gene and they did not interact with TRIM28 (Itokawa et al., 2009). Similarly, the DUF3669-KRAB configuration of ZNF746a and ZNF777 downregulated the

reporter gene although their KRAB domains did not repress the reporter gene and did not interact with TRIM28. One explanation might be that the DUF3669 domain oligomerizes with endogenous DUF3669 containing proteins by making use of their cognate TRIM28 protein mediated repression. Alternatively, the DUF3669 might confer residual repression activity independent of TRIM28 and SETDB1.

Little is known about the role of posttranslational modifications of KRAB-ZNFs in their transcriptional repression and biological function. While ubiquitination usually targets proteins for degradation (Glickman and Ciechanover, 2002), SUMOylation often regulates protein-protein interaction and alters the protein's biological function (Matunis et al., 2006). A handful of studies addressing the impact of SUMOylation on KRAB-ZNFs functionalities are available. For example, the SUMOylation of ZFP282 stabilizes the multiprotein complex with coiled-coil co-activator (CoCoA) and Estrogen receptor alpha (ERα) (Yu et al., 2013), and SUMOylation of ZNF451 is important for its nuclear localization (Karvonen et al., 2008).

The protein region of ZNF746 encoded by exon 5 offers a SUMO-receptor lysine (Nishida and Yamada, 2016) and the removal of this residue in the ZNF746a abolished the repressive activity of the whole N-terminal DUF3669-KRAB-exons 5/6 protein fragment. The SUMOylation events thus might contribute to the repression activities through stabilizing the interaction with TRIM28 or conferring a docking surface for other proteins, e.g. histone modulating cofactors.

It was previously reported that PIASy (also known as PIAS4) interacts with and SUMOylates ZNF746 at K189 and K286. PIASy can independently downregulate the PGC-1α promoter (Nishida and Yamada, 2016). On the other hand, it was shown that PIASy interacts with HDAC to confer transcriptional repression *in vitro* (Long et al., 2003). Furthermore, by searching for SETDB1 interaction partners archived in BioGRID 3.5 database (https://thebiogrid.org/), PIASy was identified as an interacting partner of SETDB1. In a related context, it was shown that the SUMOylation of ZNF133 enhances its transcriptional repressive activity. It was reported that zinc finger motifs of ZNF133 downregulates the reporter gene in a dose- and SUMOylation- dependent manner and the increasing amounts of PIAS1 (an E3 SUMO ligase) enhanced its repressive activity. It was suggested that PIAS1 interacts with zinc finger fingers and recruits other co-repressors such as HDAC (Lee et al., 2007).

Taken together, PIASy might function as an adaptor for SETDB1 and HDACs as based on our K189 substitution experiments as well. This could explain why the amino halves of

ZNF746 showed higher repression activity in TRIM28ko cells than SETDB1ko cells. Therefore, both amino halves of ZNF746 function through two distinct mechanisms; TRIM28-dependent and SETDB1-dependent mechanisms. Nevertheless, further investigations are required to confirm the assumption of molecular mechanisms underlying the PIASy-mediated repression activity.

Of note, the ZFs-containing fragment of ZNF746 also possessed transcriptional repressive activity that is dependent on SETDB1 and independent of TRIM28. This part is SUMOylated at K289 (Nishida and Yamada, 2016) suggesting that the ZFs-containing fragment of ZNF746 functions as a transcriptional repressor by recruiting PIASy in SETDB1-dependent and TRIM28-independent mechanism.

SUMOylation might be a common mechanistic theme for the biological functions of DUF-KRAB-ZNF proteins. Global SUMOylation analysis showed that ZNF777 is SUMOylated (Schimmel et al., 2014; Xiao et al., 2015). Sequence analysis revealed three SUMO-acceptor lysines are located within perfectly matching consensus sequences for SUMOylation modification (ψKXE, where ψ represents a large hydrophobic amino acid, K is the lysine conjugated to SUMO, X is any amino acid, and E is a glutamic acid) (Sampson et al., 2001); K370 is located within exon 4 encoded fragment; K370 and K689 are located in the exon 5 that encodes zinc finger motifs region. It is tempting to speculate that SUMOylation of ZNF777 can contribute to its repressive functionality. SUMOylation can explain why the protein region of ZNF777 corresponding to DUF3669-KRAB that does not harbor any SUMO-receptor lysine showed a poor repression activity, whereas full-length ZNF777 caused significant repression that is not completely dependent on TRIM28.

Another comprehensive SUMOylation analysis revealed that the DUF3669-containing KRAB-ZNFs ZNF282, ZNF398, and ZNF783 are also SUMOylated (Hendriks and Vertegaal, 2016) but it has not yet been identified at which lysine residue. Interestingly, the “IKTE” SUMOylation consensus is highly conserved in the zinc finger regions within all ZNF746, ZNF777, and ZNF783 orthologous proteins. The conservation of this motif implies that this motif might be involved in the biological function of these proteins including their transcriptional repression activities.

CoIP results showed that cellular TRIM28 co-associates with the protein regions of ZNF746b that correspond to DUF3669-KRAB and KRAB-exons 5/6. Conversely, the amino half of ZNF746b (aa 1-294) could not form a stable complex with TRIM28 although its shorter versions showed significant association with TRIM28. It is more plausible that this part is

able to recruit TRIM28 since all relevant protein regions could recruit TRIM28 and for some reason it was not detected in the same complex. The oligomerizing property of GST can lead to the formation of a higher complex that might mask the TRIM28 interacting surface, or this truncated tagged protein may not fold properly.

Non-of the tested ZNF746a- and ZNF777-derived protein fragments could stably associate with TRIM28 although most of them were partially inhibited in their moderate repressor activities by TRIM28 depletion. This might imply that complex formation with TRIM28 in cells might be very transient and could therefore not be shown by co-immunoprecipitation. Alternatively, the dependency might not be due to direct protein-protein interaction but to mechanistic links mediated through other molecules such as PIASy.

The weak interaction between ZNF746a and ZNF777 and other two DUF3669-containing KRAB-ZNFs; ZNF212, ZNF398 have been recently reported even despite the use of crosslinking before affinity enrichment from cellular extracts (Helleboid et al., 2019; Kang and Shin, 2015). Current knowledge regarding previous studies and our results on interactions of DUF3669 containing KRAB–ZNF proteins illustrate that ZNF746 as bait interacts with ZNF783, ZNF777, ZNF 282 and ZNF212, whereas ZNF777 only interacts as bait with ZNF746 (Fig. 25).

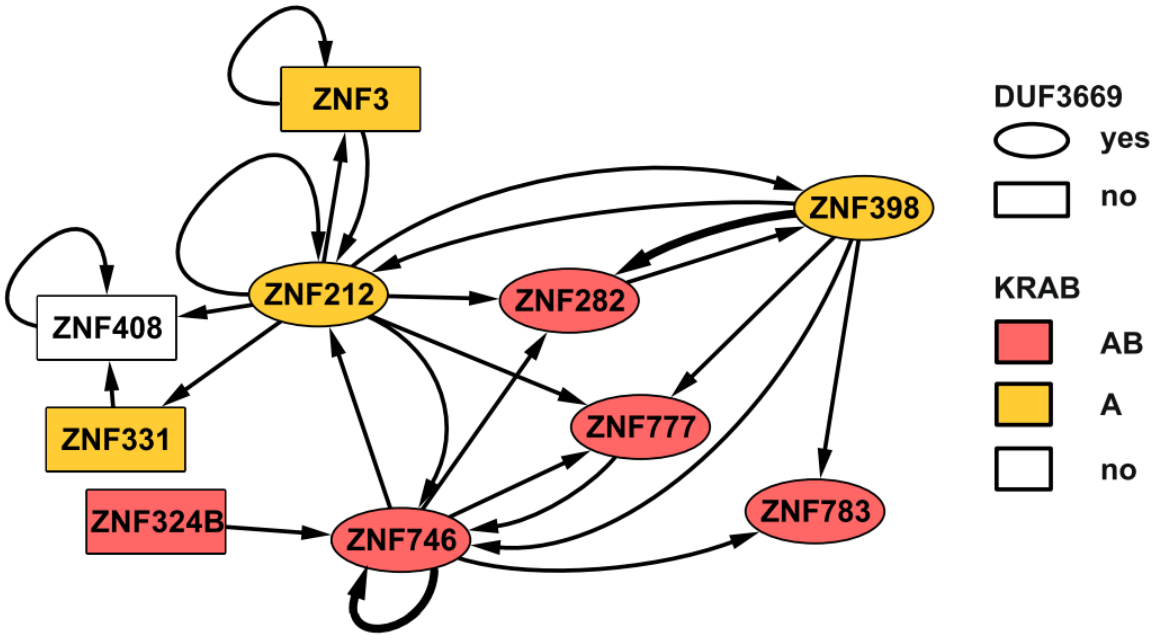

Fig. 25: Summary of the current knowledge about shared complexes of DUF3669-containing KRAB-ZNF proteins based on our results and previous studies. *Arrows point from bait to prey. Line width reflects detection in one or two (thicker line) studies. Derived from data that were based on affinity-purification/mass spectrometry using tagged DUF3669-KRAB-ZNF proteins (Bioplex 2.0) (Kang et al., 2015; Huttlin et al., 2017), a yeast two-hybrid system (Gao et al., 2008), co-*

immunoprecipitation (Helleboid et al., 2019) and including the data of this study (Al Chiblak et al., 2019).

For the first time, this study demonstrates that ZNF746 forms direct homodimer and heterodimer with ZNF777 whereas ZNF777 failed to form homodimers. The issue lies in which single amino acids do contribute to this divergent behavior.

Given the substantial differences in their principle capacity to silence gene expression, it will be important to consider which isoform happens to be involved in a certain biological context when studying and interpreting the cellular functions of ZNF746. The isoform-dependent biological function has been recently shown for ZNF398 (Huang et al., 2019). It is conceivable that isoform ZNF746b with full-length KRAB-A does not only show increased TRIM28 interaction and a boosted repressor activity like we observed here but might display a different overall interactome as well. ZNF746 could, therefore, be subject to a scenario in which differential splicing leads to the expression of protein isoforms with differential transcriptional/translational regulatory potentials.

Using mutational analysis in ZNF746-A, the results of this study pinpointed the main reason for this difference in repression activity due to the lack of a glutamic acid residue at a conserved “MLE” motif of canonical KRAB-A subdomains. The results showed that glutamic acid alone caused an extreme increase in repressive activity even more than the KRAB domain of ZNF10.

In addition, the interaction with TRIM28 was remarkably improved. Leucine did not make any differences. This result highlights the importance of the negative charge of glutamic acid to stabilize the interaction between the KRAB domain and TRIM28. The very recent studies published on the structural details of the TRIM28/KRAB interaction (Stoll et al., 2019; Sun et al., 2019) might enable the molecular modeling of the difference between glycine vs. glutamic acid in the future.

The KRAB-A of all human DUF3669-containing zinc finger proteins and their orthologs show a common difference in comparison to canonical KRAB A domains. Interestingly, ZNF282 orthologs are characterized by “VKE”, i.e. there is a glutamic acid residue. Nevertheless, the KRAB domain of human ZNF282/HUB1 was unable to repress a reporter gene (Okumura et al., 1997), indicating that other amino acids besides the glutamic acid are also contributing to this property. The lack of LE conserved residues in KRAB-A box in almost all DUF3669-containing KRAB-ZNFs suggests that they have an impaired interaction

with TRIM28 leading to reduced repression, but opening up the possibility that additional DUF3669 mediated functions might be executed not directly linked to TRIM28 binding.

Altogether, the canonical KRAB/TRIM28 pathway cannot alone explain the repressive activities of DUF3669-containing KRAB-ZNF proteins. Other mechanisms that rely on other structural components are intrinsic and important for this protein subfamily.

The gene regulation ability of the KRAB-ZNFs has been researched mainly using artificial systems because of the limited number of endogenous targets of KRAB-ZNFs identified up to now. In this study here, classical heterologous luciferase reporter assays in which the protein of interest is artificially targeted to DNA via the fused Gal4 DNA-binding domain was used.

The luciferase reporter assay is a golden standard that is widely used in gene regulation studies to gain insight into nuclear cell biology. In this study, we relied on the dual-luciferase assay to evaluate the transcriptional repressive capacities of several protein regions of ZNF746 and ZNF777. The major concern is that this exogenous reporter plasmid may not give insight into the native epigenetic gene regulation machinery. The presence of proper chromatin configuration on such plasmid DNA is still controversial. It is not clear whether plasmid DNA owns a similar chromatin dynamics and response to chromatin modifiers as a cellular chromosome.

Transiently transfected plasmid DNA has been shown to be associated with nucleosomes of comparable stoichiometry but positioned randomly with intermediate nucleosomal assembly levels and fewer H1 molecules compared to stably integrated promoters. The decreasing level of H1 suggests that the nucleosomes become less stable as compared to a stable template (Hebbar and Archer, 2008). Nuclease digestion method in HEK293 cells confirmed that the transiently-transfected plasmid DNA is packed into histones to form an abnormal chromatin-like structure (Mladenova et al., 2009). In addition, it was shown that the KRAB/TRIM28 pathway is completely efficient within the context of episomal DNA (Barde et al., 2009).

7- Conclusion and outlook

In this study, two members of the DUF3669-containing KRAB-ZNF subfamily have been investigated for their repressive activities in an attempt to unveil the role of their unconventional amino termini in their overall repression capacity. This study demonstrated transcriptional repression and the oligomerization properties of DUF3669. The amino terminal configuration of DUF3669-KRAB and the adjacent sequence constitutes an integral structural unit to exert sufficient repression activities even if the KRAB domain is a weak or non-functional repressor. The DUF3669 domain and the SUMOylation events add novel functionalities to their amino-terminal configuration. A TRIM28-independent residual repressor mechanism is considered to be partially involved as seen in HAP1 cells deleted of TRIM28. The insufficient repression activities of the tested KRAB domains are mainly due to their impaired binding to TRIM28 because of the absence of a conserved MLE motif in their KRAB-A subdomain.

The conservation of some SUMO-receptor lysines in some DUF3669-containing zinc finger protein draws the attention to their importance in cellular physiology including transcriptional gene regulation which deserves further investigations.

The underlying mechanism by which DUF3669 downregulates the reporter gene is unknown. Repression could occur directly e.g. by steric hindrance in this artificial reporter system of the RNA polymerase machinery, e.g. by recruitment of unknown factors to the promoter or by forming complexes with other DUF3669-containing KRAB-ZNFs or histone modifiers.

Mutants of DUF3669 that fail to oligomerize are considered to be useful in the future to evaluate whether oligomerization properties are necessary for providing repression activities of DUF3669-containing KRAB-ZNF proteins.

Identification of the endogenous target genes is a crucial step to unveil the biological function of individual KRAB-ZNF proteins. In addition, the use of such endogenous promoters gives more insight into the regulatory capacity of KRAB-ZNF proteins. For example, investigation of the transcriptional activities of ZNF746-derived fragment by using PGC-1α or TKT promotor can be more meaningful than an artificial luciferase system.

Definitely, the HAP1 cell system has been proven to be an essential method to study gene modifications on the single gene level.

8- Acronyms and Abbreviations

APS	Ammonium persulfate
ATF7IP	Activating Transcription Factor 7 Interacting Protein
BR	Bromodomain
BSA	Bovine Serum Albumin
Co-IP	Co-Immunoprecipitation
DUF3669	Domain of Unknown Function 3669
DTT	Dithiothreitol
FAM129A	Family With Sequence Similarity 129 Member A
GAPDH	Glyceraldehyde-3-Phosphate Dehydrogenase
GFP	Green Fluorescent Protein
GST	Glutathione S Transferase
H3K9me	Methyl mark at lysine 9 of histone 3 amino-terminal tail
HDAC	Histone deacetylase
HMTases	Histone methyltransferase
HP1	Heterochromatin protein 1
HP1-BD	Heterochromatin protein 1-binding domain
KAP1	KRAB associated protein 1
KOX	Krüppel associated box
KRAB	Krüppel associated box
MBP	Maltose binding protein
Ni-NTA	Nickel-Nitrilotriacetic acid
NuRD	Nucleosome Remodeling Deacetylase
PARIS	Parkin Interacting Substrate
PBS	Phosphate Buffered Saline
PGC-1α	Peroxisome Proliferator-Activated Receptor gamma (PPARγ) coactivator-1a
PHD	Plant Homeodomain
PINK1	PTEN-induced putative kinase 1
PPARγ	Peroxisome Proliferator-Activated Receptor gamma
PTMs	Post-transitional Modifications
RBCC	RING, B1box, B2box, and Coiled-Coil
SAM	S-adenosyl-L-methionine
SCAN	SRE-ZBP, Ctfin51, AW-1, and Number 18
SDS	Sodium Dodecyl Sulfate
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
SETDB1	SET Domain Bifurcated Histone Lysine Methyltransferase 1
SIM	SUMO Interacting Motif
SUMO	Small Ubiquitin-like Modifier
TE	Transposable Element
TIF1	Transcription intermediary factor 1
TKT	transketolase
TAE	Tris-acetate-EDTA
TBE	Tris-Borate-EDTA
TRIM28	Tripartite Motif Containing 28
TRIS	Tris(hydroxymethyl)aminomethane
TSS	TIF1 Signature Sequence

9- References

Al Chiblak, M., Steinbeck, F., Thiesen, H.-J., Lorenz, P., 2019. DUF3669, a “domain of unknown function” within ZNF746 and ZNF777, oligomerizes and contributes to transcriptional repression. BMC Mol. Cell Biol. 20, 60. https://doi.org/10.1186/s12860-019-0243-y

Allfrey, V.G., Faulkner, R., Mirsky, A.E., 1964. ACETYLATION AND METHYLATION OF HISTONES AND THEIR POSSIBLE ROLE IN THE REGULATION OF RNA SYNTHESIS*. Proc. Natl. Acad. Sci. U. S. A. 51, 786–794.

Bannister, A.J., Kouzarides, T., 2011. Regulation of chromatin by histone modifications. Cell Res. 21, 381–395. https://doi.org/10.1038/cr.2011.22

Bannister, A.J., Zegerman, P., Partridge, J.F., Miska, E.A., Thomas, J.O., Allshire, R.C., Kouzarides, T., 2001. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. Nature 410, 120. https://doi.org/10.1038/35065138

Barde, I., Laurenti, E., Verp, S., Groner, A.C., Towne, C., Padrun, V., Aebischer, P., Trumpp, A., Trono, D., 2009. Regulation of Episomal Gene Expression by KRAB/KAP1-Mediated Histone Modifications. J. Virol. 83, 5574–5580. https://doi.org/10.1128/JVI.00001-09

Bellefroid, E.J., Marine, J.C., Ried, T., Lecocq, P.J., Rivière, M., Amemiya, C., Poncelet, D.A., Coulie, P.G., de Jong, P., Szpirer, C., 1993. Clustered organization of homologous KRAB zinc-finger genes with enhanced expression in human T lymphoid cells. EMBO J. 12, 1363–1374.

Bellefroid, E.J., Poncelet, D.A., Lecocq, P.J., Revelant, O., Martial, J.A., 1991. The evolutionarily conserved Krüppel-associated box domain defines a subfamily of eukaryotic multifingered proteins. Proc. Natl. Acad. Sci. U. S. A. 88, 3608–3612.

Birnboim, H.C., Doly, J., 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. 7, 1513–1523.

Born, N., Thiesen, H.-J., Lorenz, P., 2014. The B-Subdomain of the Xenopus laevis XFIN KRAB-AB Domain Is Responsible for Its Weaker Transcriptional Repressor Activity Compared to Human ZNF10/Kox1. PLoS ONE 9. https://doi.org/10.1371/journal.pone.0087609

Campos, E.I., Reinberg, D., 2009. Histones: Annotating Chromatin. Annu. Rev. Genet. 43, 559–599. https://doi.org/10.1146/annurev.genet.032608.103928

Canzio, D., Chang, E.Y., Shankar, S., Kuchenbecker, K.M., Simon, M.D., Madhani, H.D., Narlikar, G.J., Al-Sady, B., 2011. Chromodomain-mediated oligomerization of HP1 suggests a nucleosome-bridging mechanism for heterochromatin assembly. Mol. Cell 41, 67–81. https://doi.org/10.1016/j.molcel.2010.12.016

Carette, J.E., Guimaraes, C.P., Varadarajan, M., Park, A.S., Wuethrich, I., Godarova, A., Kotecki, M., Cochran, B.H., Spooner, E., Ploegh, H.L., Brummelkamp, T.R., 2009. Haploid Genetic Screens in Human Cells Identify Host Factors Used by Pathogens. Science 326, 1231–1235. https://doi.org/10.1126/science.1178955

Carette, J.E., Raaben, M., Wong, A.C., Herbert, A.S., Obernosterer, G., Mulherkar, N., Kuehne, A.I., Kranzusch, P.J., Griffin, A.M., Ruthel, G., Cin, P.D., Dye, J.M., Whelan, S.P., Chandran, K., Brummelkamp, T.R., 2011. Ebola virus entry requires the cholesterol transporter Niemann-Pick C1. Nature 477, 340–343. https://doi.org/10.1038/nature10348

Cesaro, E., De Cegli, R., Medugno, L., Florio, F., Grosso, M., Lupo, A., Izzo, P., Costanzo, P., 2009. The Kruppel-like zinc finger protein ZNF224 recruits the arginine methyltransferase PRMT5 on the transcriptional repressor complex of the aldolase A gene. J. Biol. Chem. 284, 32321–32330. https://doi.org/10.1074/jbc.M109.043349

Chatterjee, S., Ahituv, N., 2017. Gene Regulatory Elements, Major Drivers of Human Disease. Annu. Rev. Genomics Hum. Genet. 18, 45–63. https://doi.org/10.1146/annurev-genom-091416-035537

Chen, W., Schwalie, P.C., Pankevich, E.V., Gubelmann, C., Raghav, S.K., Dainese, R., Cassano, M., Imbeault, M., Jang, S.M., Russeil, J., Delessa, T., Duc, J., Trono, D., Wolfrum, C., Deplancke, B., 2019. ZFP30 promotes adipogenesis through the KAP1-mediated activation of a retrotransposon-derived Pparg2 enhancer. Nat. Commun. 10, 1–16. https://doi.org/10.1038/s41467-019-09803-9

Chen, Y.-T., Yang, C.-C., Shao, P.-L., Huang, C.-R., Yip, H.-K., 2019. Melatonin-mediated downregulation of ZNF746 suppresses bladder tumorigenesis mainly through inhibiting the AKT-MMP-9 signaling pathway. J. Pineal Res. 66, e12536. https://doi.org/10.1111/jpi.12536

Crooks, G.E., Hon, G., Chandonia, J.-M., Brenner, S.E., 2004. WebLogo: a sequence logo generator. Genome Res. 14, 1188–1190. https://doi.org/10.1101/gr.849004

Czerwińska, P., Mazurek, S., Wiznerowicz, M., 2017. The complexity of TRIM28 contribution to cancer. J. Biomed. Sci. 24. https://doi.org/10.1186/s12929-017-0374-4

Deng, N., Zhou, H., Fan, H., Yuan, Y., 2017. Single nucleotide polymorphisms and cancer susceptibility. Oncotarget 8, 110635–110649. https://doi.org/10.18632/oncotarget.22372

Deuschle, U., Meyer, W.K., Thiesen, H.J., 1995. Tetracycline-reversible silencing of eukaryotic promoters. Mol. Cell. Biol. 15, 1907–1914. https://doi.org/10.1128/mcb.15.4.1907

Di Lorenzo, A., Bedford, M.T., 2011. Histone arginine methylation. FEBS Lett. 585, 2024–2031. https://doi.org/10.1016/j.febslet.2010.11.010

Dillon, S.C., Zhang, X., Trievel, R.C., Cheng, X., 2005. The SET-domain protein superfamily: protein lysine methyltransferases. Genome Biol. 6, 227. https://doi.org/10.1186/gb-2005-6-8-227

Ecco, G., Imbeault, M., Trono, D., 2017. KRAB zinc finger proteins. Development 144, 2719–2729. https://doi.org/10.1242/dev.132605

Eom, K.S., Cheong, J.S., Lee, S.J., 2016. Structural Analyses of Zinc Finger Domains for Specific Interactions with DNA. J. Microbiol. Biotechnol. 26, 2019–2029. https://doi.org/10.4014/jmb.1609.09021

Essletzbichler, P., Konopka, T., Santoro, F., Chen, D., Gapp, B.V., Kralovics, R., Brummelkamp, T.R., Nijman, S.M.B., Bürckstümmer, T., 2014. Megabase-scale deletion using CRISPR/Cas9 to generate a fully haploid human cell line. Genome Res. 24, 2059–2065. https://doi.org/10.1101/gr.177220.114

Fedotova, A.A., Bonchuk, A.N., Mogila, V.A., Georgiev, P.G., 2017. C2H2 Zinc Finger Proteins: The Largest but Poorly Explored Family of Higher Eukaryotic Transcription Factors. Acta Naturae 9, 47–58.

Finkbeiner, S., 2001. New roles for introns: sites of combinatorial regulation of Ca2+- and cyclic AMP-dependent gene transcription. Sci. STKE Signal Transduct. Knowl. Environ. 2001, pe1. https://doi.org/10.1126/stke.2001.94.pe1

Friedman, J.R., Fredericks, W.J., Jensen, D.E., Speicher, D.W., Huang, X.P., Neilson, E.G., Rauscher, F.J., 1996. KAP-1, a novel corepressor for the highly conserved KRAB repression domain. Genes Dev. 10, 2067–2078. https://doi.org/10.1101/gad.10.16.2067

Gao, J., Li, W.-X., Feng, S.-Q., Yuan, Y.-S., Wan, D.-F., Han, W., Yu, Y., 2008. A protein–protein interaction network of transcription factors acting during liver cell proliferation. Genomics 91, 347–355. https://doi.org/10.1016/j.ygeno.2007.12.007

Gerhards, N.M., Blomen, V.A., Mutlu, M., Nieuwenhuis, J., Howald, D., Guyader, C., Jonkers, J., Brummelkamp, T.R., Rottenberg, S., 2018. Haploid genetic screens identify genetic vulnerabilities to microtubule-targeting agents. Mol. Oncol. 12, 953–971. https://doi.org/10.1002/1878-0261.12307

Glickman, M.H., Ciechanover, A., 2002. The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction. Physiol. Rev. 82, 373–428. https://doi.org/10.1152/physrev.00027.2001

Graham, F.L., Smiley, J., Russell, W.C., Nairn, R., 1977. Characteristics of a human cell line transformed by DNA from human adenovirus type 5. J. Gen. Virol. 36, 59–74. https://doi.org/10.1099/0022-1317-36-1-59

Groner, A.C., Meylan, S., Ciuffi, A., Zangger, N., Ambrosini, G., Dénervaud, N., Bucher, P., Trono, D., 2010. KRAB–Zinc Finger Proteins and KAP1 Can Mediate Long-Range Transcriptional Repression through Heterochromatin Spreading. PLOS Genet. 6, e1000869. https://doi.org/10.1371/journal.pgen.1000869

Gutierrez-Guerrero, A., Sanchez-Hernandez, S., Galvani, G., Pinedo-Gomez, J., Martin-Guerra, R., Sanchez-Gilabert, A., Aguilar-González, A., Cobo, M., Gregory, P., Holmes, M., Benabdellah, K., Martin, F., 2018. Comparison of Zinc Finger Nucleases Versus CRISPR-Specific Nucleases for Genome Editing of the Wiskott-Aldrich Syndrome Locus. Hum. Gene Ther. 29, 366–380. https://doi.org/10.1089/hum.2017.047

Hallen, L., Klein, H., Stoschek, C., Wehrmeyer, S., Nonhoff, U., Ralser, M., Wilde, J., Röhr, C., Schweiger, M.R., Zatloukal, K., Vingron, M., Lehrach, H., Konthur, Z., Krobitsch, S., 2011. The KRAB-containing zinc-finger transcriptional regulator ZBRK1 activates SCA2 gene transcription through direct interaction with its gene product, ataxin-2. Hum. Mol. Genet. 20, 104–114. https://doi.org/10.1093/hmg/ddq436

Hanahan, D., 1983. Studies on transformation of Escherichia coli with plasmids. J. Mol. Biol. 166, 557–580. https://doi.org/10.1016/S0022-2836(83)80284-8

Harte, P.J., Wu, W., Carrasquillo, M.M., Matera, A.G., 1999. Assignment of a novel bifurcated SET domain gene, SETDB1, to human chromosome band 1q21 by in situ hybridization and radiation hybrids. Cytogenet. Cell Genet. 84, 83–86. https://doi.org/10.1159/000015220

Hebbar, P.B., Archer, T.K., 2008. Altered histone H1 stoichiometry and an absence of nucleosome positioning on transfected DNA. J. Biol. Chem. 283, 4595–4601. https://doi.org/10.1074/jbc.M709121200

Hebbes, T.R., Thorne, A.W., Crane-Robinson, C., 1988. A direct link between core histone acetylation and transcriptionally active chromatin. EMBO J. 7, 1395–1402.

Helleboid, P.-Y., Heusel, M., Duc, J., Piot, C., Thorball, C.W., Coluccio, A., Pontis, J., Imbeault, M., Turelli, P., Aebersold, R., Trono, D., 2019. The interactome of KRAB zinc finger proteins reveals the evolutionary history of their functional diversification. EMBO J. 0, e101220. https://doi.org/10.15252/embj.2018101220

Hendriks, I.A., Vertegaal, A.C.O., 2016. A comprehensive compilation of SUMO proteomics. Nat. Rev. Mol. Cell Biol. 17, 581–595. https://doi.org/10.1038/nrm.2016.81

Huang, C., Wu, S., Li, W., Herkilini, A., Miyagishi, M., Zhao, H., Kasim, V., 2019. Zinc-finger protein p52-ZER6 accelerates colorectal cancer cell proliferation and tumour progression through promoting p53 ubiquitination. EBioMedicine. https://doi.org/10.1016/j.ebiom.2019.08.070

Huntley, S., Baggott, D.M., Hamilton, A.T., Tran-Gyamfi, M., Yang, S., Kim, J., Gordon, L., Branscomb, E., Stubbs, L., 2006. A comprehensive catalog of human KRAB-associated zinc finger genes: insights into the evolutionary history of a large family of transcriptional repressors. Genome Res. 16, 669–677. https://doi.org/10.1101/gr.4842106

Huttlin, E.L., Bruckner, R.J., Paulo, J.A., Cannon, J.R., Ting, L., Baltier, K., Colby, G., Gebreab, F., Gygi, M.P., Parzen, H., Szpyt, J., Tam, S., Zarraga, G., Pontano-Vaites, L., Swarup, S., White, A.E., Schweppe, D.K., Rad, R., Erickson, B.K., Obar, R.A., Guruharsha, K.G., Li, K., Artavanis-Tsakonas, S., Gygi, S.P., Harper, J.W., 2017. Architecture of the human interactome defines protein communities and disease networks. Nature 545, 505–509. https://doi.org/10.1038/nature22366

Imbeault, M., Helleboid, P.-Y., Trono, D., 2017. KRAB zinc-finger proteins contribute to the evolution of gene regulatory networks. Nature 543, 550–554. https://doi.org/10.1038/nature21683

Itokawa, Y., Yanagawa, T., Yamakawa, H., Watanabe, N., Koga, H., Nagase, T., 2009. KAP1-independent transcriptional repression of SCAN-KRAB-containing zinc finger proteins. Biochem. Biophys. Res. Commun. 388, 689–694. https://doi.org/10.1016/j.bbrc.2009.08.065

Ivanov, A.V., Peng, H., Yurchenko, V., Yap, K.L., Negorev, D.G., Schultz, D.C., Psulkowski, E., Fredericks, W.J., White, D.E., Maul, G.G., Sadofsky, M.J., Zhou, M.-M., Rauscher, F.J., 2007. PHD domain-mediated E3 ligase activity directs intramolecular sumoylation of an adjacent bromodomain required for gene silencing. Mol. Cell 28, 823–837. https://doi.org/10.1016/j.molcel.2007.11.012

Jung, J.H., Jung, D.-B., Kim, H., Lee, H., Kang, S.-E., Srivastava, S.K., Yun, M., Kim, S.-H., 2018. Zinc finger protein 746 promotes colorectal cancer progression via c-Myc stability mediated by glycogen synthase kinase 3β and F-box and WD repeat domain-containing 7. Oncogene 37, 3715–3728. https://doi.org/10.1038/s41388-018-0225-0

Kang, H., Jo, A., Kim, Hyein, Khang, R., Lee, J.-Y., Kim, Hanna, Park, C.-H., Choi, J.-Y., Lee, Y., Shin, J.-H., 2018. PARIS reprograms glucose metabolism by HIF-1α induction in dopaminergic neurodegeneration. Biochem. Biophys. Res. Commun. 495, 2498–2504. https://doi.org/10.1016/j.bbrc.2017.12.147

Kang, H., Shin, J.-H., 2015. Repression of rRNA transcription by PARIS contributes to Parkinson's disease. Neurobiol. Dis. 73, 220–228. https://doi.org/10.1016/j.nbd.2014.10.003

Karimi, M., Goldie, L.C., Ulgiati, D., Abraham, L.J., 2007. Integration site-specific transcriptional reporter gene analysis using Flp recombinase targeted cell lines. BioTechniques 42, 217–224. https://doi.org/10.2144/000112317

Karvonen, U., Jääskeläinen, T., Rytinki, M., Kaikkonen, S., Palvimo, J.J., 2008. ZNF451 is a novel PML body- and SUMO-associated transcriptional coregulator. J. Mol. Biol. 382, 585–600. https://doi.org/10.1016/j.jmb.2008.07.016

Keskitalo, S., Haapaniemi, E.M., Glumoff, V., Liu, X., Lehtinen, V., Fogarty, C., Rajala, H., Chiang, S.C., Mustjoki, S., Kovanen, P., Lohi, J., Bryceson, Y.T., Seppänen, M., Kere, J., Heiskanen, K., Varjosalo, M., 2019. Dominant TOM1 mutation associated with combined immunodeficiency and autoimmune disease. Npj Genomic Med. 4, 1–7. https://doi.org/10.1038/s41525-019-0088-5

Kim, B., Sohn, E.J., Jung, J.H., Shin, E.A., You, O.H., Im, J., Kim, S.-H., 2014. Inhibition of ZNF746 suppresses invasion and epithelial to mesenchymal transition in H460 non-small cell lung cancer cells. Oncol. Rep. 31, 73–78. https://doi.org/10.3892/or.2013.2801

Kim, H., Kang, H., Lee, Y., Park, C.-H., Jo, A., Khang, R., Shin, J.-H., 2017. Identification of transketolase as a target of PARIS in substantia nigra. Biochem. Biophys. Res. Commun. 493, 1050–1056. https://doi.org/10.1016/j.bbrc.2017.09.090

Kotecki, M., Reddy, P.S., Cochran, B.H., 1999. Isolation and Characterization of a Near-Haploid Human Cell Line. Exp. Cell Res. 252, 273–280. https://doi.org/10.1006/excr.1999.4656

Kyhse-Andersen, J., 1984. Electroblotting of multiple gels: a simple apparatus without buffer tank for rapid transfer of proteins from polyacrylamide to nitrocellulose. J. Biochem. Biophys. Methods 10, 203–209.

Lechner, M.S., Begg, G.E., Speicher, D.W., Rauscher, F.J., 2000. Molecular Determinants for Targeting Heterochromatin Protein 1-Mediated Gene Silencing: Direct Chromoshadow Domain–KAP-1 Corepressor Interaction Is Essential. Mol. Cell. Biol. 20, 6449–6465.

Lee, S.-J., Lee, J.-R., Hah, H.-S., Kim, Y.-H., Ahn, J.-H., Bae, C.-D., Yang, J.-M., Hahn, M.-J., 2007. PIAS1 interacts with the KRAB zinc finger protein, ZNF133, via zinc finger motifs and regulates its transcriptional activity. Exp. Mol. Med. 39, 450–457. https://doi.org/10.1038/emm.2007.49

Lee, Y., Stevens, D.A., Kang, S.-U., Jiang, H., Lee, Y.-I., Ko, H.S., Scarffe, L.A., Umanah, G.E., Kang, H., Ham, S., Kam, T.-I., Allen, K., Brahmachari, S., Kim, J.W., Neifert, S., Yun, S.P., Fiesel, F.C., Springer, W., Dawson, V.L., Shin, J.-H., Dawson, T.M., 2017. PINK1 Primes Parkin-Mediated Ubiquitination of PARIS in Dopaminergic Neuronal Survival. Cell Rep. 18, 918–932. https://doi.org/10.1016/j.celrep.2016.12.090

Liu, H., Chang, L.-H., Sun, Y., Lu, X., Stubbs, L., 2014. Deep vertebrate roots for mammalian zinc finger transcription factor subfamilies. Genome Biol. Evol. 6, 510–525. https://doi.org/10.1093/gbe/evu030

Long, J., Matsuura, I., He, D., Wang, G., Shuai, K., Liu, F., 2003. Repression of Smad transcriptional activity by PIASy, an inhibitor of activated STAT. Proc. Natl. Acad. Sci. 100, 9791–9796. https://doi.org/10.1073/pnas.1733973100

Looman, C., Åbrink, M., Mark, C., Hellman, L., 2002. KRAB Zinc Finger Proteins: An Analysis of the Molecular Mechanisms Governing Their Increase in Numbers and Complexity During Evolution. Mol. Biol. Evol. 19, 2118–2130. https://doi.org/10.1093/oxfordjournals.molbev.a004037

Looman, C., Hellman, L., Abrink, M., 2004. A novel Krüppel-Associated Box identified in a panel of mammalian zinc finger proteins. Mamm. Genome Off. J. Int. Mamm. Genome Soc. 15, 35–40. https://doi.org/10.1007/s00335-003-3022-0

Losson, R., Nielsen, A.L., 2010. The NIZP1 KRAB and C2HR domains cross-talk for transcriptional regulation. Biochim. Biophys. Acta 1799, 463–468. https://doi.org/10.1016/j.bbagrm.2010.02.003

Lupo, A., Cesaro, E., Montano, G., Zurlo, D., Izzo, P., Costanzo, P., 2013. KRAB-Zinc Finger Proteins: A Repressor Family Displaying Multiple Biological Functions. Curr. Genomics 14, 268–278. https://doi.org/10.2174/13892029113149990002

Machnik, M., Cylwa, R., Kiełczewski, K., Biecek, P., Liloglou, T., Mackiewicz, A., Oleksiewicz, U., 2019. The expression signature of cancer-associated KRAB-ZNF factors identified in TCGA

pan-cancer transcriptomic data. Mol. Oncol. 13, 701–724. https://doi.org/10.1002/1878-0261.12407

Margolin, J.F., Friedman, J.R., Meyer, W.K., Vissing, H., Thiesen, H.J., Rauscher, F.J., 1994. Krüppel-associated boxes are potent transcriptional repression domains. Proc. Natl. Acad. Sci. U. S. A. 91, 4509–4513.

Mark, C., Abrink, M., Hellman, L., 1999. Comparative Analysis of KRAB Zinc Finger Proteins in Rodents and Man: Evidence for Several Evolutionarily Distinct Subfamilies of KRAB Zinc Finger Genes. DNA Cell Biol. 18, 381–396. https://doi.org/10.1089/104454999315277

Martinez, E., 2002. Multi-protein complexes in eukaryotic gene transcription. Plant Mol. Biol. 50, 925–947. https://doi.org/10.1023/A:1021258713850

Mascle, X.H., Germain-Desprez, D., Huynh, P., Estephan, P., Aubry, M., 2007. Sumoylation of the Transcriptional Intermediary Factor 1β (TIF1β), the Co-repressor of the KRAB Multifinger Proteins, Is Required for Its Transcriptional Activity and Is Modulated by the KRAB Domain. J. Biol. Chem. 282, 10190–10202. https://doi.org/10.1074/jbc.M611429200

Maston, G.A., Evans, S.K., Green, M.R., 2006. Transcriptional Regulatory Elements in the Human Genome. Annu. Rev. Genomics Hum. Genet. 7, 29–59. https://doi.org/10.1146/annurev.genom.7.080505.115623

Matunis, M.J., Zhang, X.-D., Ellis, N.A., 2006. SUMO: The Glue that Binds. Dev. Cell 11, 596–597. https://doi.org/10.1016/j.devcel.2006.10.011

Meylan, S., Groner, A.C., Ambrosini, G., Malani, N., Quenneville, S., Zangger, N., Kapopoulou, A., Kauzlaric, A., Rougemont, J., Ciuffi, A., Bushman, F.D., Bucher, P., Trono, D., 2011. A gene-rich, transcriptionally active environment and the pre-deposition of repressive marks are predictive of susceptibility to KRAB/KAP1-mediated silencing. BMC Genomics 12, 378. https://doi.org/10.1186/1471-2164-12-378

Mladenova, V., Mladenov, E., Russev, G., 2009. Organization of Plasmid DNA into Nucleosome-Like Structures after Transfection in Eukaryotic Cells. Biotechnol. Biotechnol. Equip. 23, 1044–1047. https://doi.org/10.1080/13102818.2009.10817609

Mudgal, R., Sandhya, S., Chandra, N., Srinivasan, N., 2015. De-DUFing the DUFs: Deciphering distant evolutionary relationships of Domains of Unknown Function using sensitive homology detection methods. Biol. Direct 10, 38. https://doi.org/10.1186/s13062-015-0069-2

Murphy, K.E., Shylo, N.A., Alexander, K.A., Churchill, A.J., Copperman, C., García-García, M.J., 2016. The Transcriptional Repressive Activity of KRAB Zinc Finger Proteins Does Not Correlate with Their Ability to Recruit TRIM28. PLOS ONE 11, e0163555. https://doi.org/10.1371/journal.pone.0163555

Mysliwiec, M.R., Kim, T., Lee, Y., 2007. Characterization of Zinc finger protein 496 that interacts with Jumonji/Jarid2. FEBS Lett. 581, 2633–2640. https://doi.org/10.1016/j.febslet.2007.05.006

Nishida, T., Yamada, Y., 2016. SUMOylation of the KRAB zinc-finger transcription factor PARIS/ZNF746 regulates its transcriptional activity. Biochem. Biophys. Res. Commun. 473, 1261–1267. https://doi.org/10.1016/j.bbrc.2016.04.051

O'Geen, H., Squazzo, S.L., Iyengar, S., Blahnik, K., Rinn, J.L., Chang, H.Y., Green, R., Farnham, P.J., 2007. Genome-Wide Analysis of KAP1 Binding Suggests Autoregulation of KRAB-ZNFs. PLOS Genet. 3, e89. https://doi.org/10.1371/journal.pgen.0030089

Okumura, K., Sakaguchi, G., Naito, K., Tamura, T., Igarashi, H., 1997. HUB1, a novel Krüppel type zinc finger protein, represses the human T cell leukemia virus type I long terminal repeat-mediated expression. Nucleic Acids Res. 25, 5025–5032.

Olbrich, T., Mayor-Ruiz, C., Vega-Sendino, M., Gomez, C., Ortega, S., Ruiz, S., Fernandez-Capetillo, O., 2017. A p53-dependent response limits the viability of mammalian haploid cells. Proc. Natl. Acad. Sci. U. S. A. 114, 9367–9372. https://doi.org/10.1073/pnas.1705133114

Pao, P.-C., Huang, N.-K., Liu, Y.-W., Yeh, S.-H., Lin, S.-T., Hsieh, C.-P., Huang, A.-M., Huang, H.-S., Tseng, J.T., Chang, W.-C., Lee, Y.-C., 2011. A novel RING finger protein, Znf179, modulates cell cycle exit and neuronal differentiation of P19 embryonal carcinoma cells. Cell Death Differ. 18, 1791–1804. https://doi.org/10.1038/cdd.2011.52

Peng, Hongzhuang, Begg, G.E., Harper, S.L., Friedman, J.R., Speicher, D.W., Rauscher, F.J., 2000. Biochemical Analysis of the Kruppel-associated Box (KRAB) Transcriptional Repression Domain SPECTRAL, KINETIC, AND STOICHIOMETRIC PROPERTIES OF THE

KRAB·KAP-1 COMPLEX. J. Biol. Chem. 275, 18000–18010. https://doi.org/10.1074/jbc.M001499200

Peng, H., Begg, G.E., Schultz, D.C., Friedman, J.R., Jensen, D.E., Speicher, D.W., Rauscher, F.J., 2000. Reconstitution of the KRAB-KAP-1 repressor complex: a model system for defining the molecular anatomy of RING-B box-coiled-coil domain-mediated protein-protein interactions. J. Mol. Biol. 295, 1139–1162. https://doi.org/10.1006/jmbi.1999.3402

Peng, H., Ivanov, A.V., Oh, H.J., Lau, Y.-F.C., Rauscher, F.J., 2009. Epigenetic gene silencing by the SRY protein is mediated by a KRAB-O protein that recruits the KAP1 co-repressor machinery. J. Biol. Chem. 284, 35670–35680. https://doi.org/10.1074/jbc.M109.032086

Pengue, G., Calabrò, V., Bartoli, P.C., Pagliuca, A., Lania, L., 1994. Repression of transcriptional activity at a distance by the evolutionarily conserved KRAB domain present in a subfamily of zinc finger proteins. Nucleic Acids Res. 22, 2908–2914.

Pengue, G., Lania, L., 1996. Krüppel-associated box-mediated repression of RNA polymerase II promoters is influenced by the arrangement of basal promoter elements. Proc. Natl. Acad. Sci. U. S. A. 93, 1015–1020.

Porsch-Ozcurumez, M., Langmann, T., Heimerl, S., Borsukova, H., Kaminski, W.E., Drobnik, W., Honer, C., Schumacher, C., Schmitz, G., 2001. The zinc finger protein 202 (ZNF202) is a transcriptional repressor of ATP binding cassette transporter A1 (ABCA1) and ABCG1 gene expression and a modulator of cellular lipid efflux. J. Biol. Chem. 276, 12427–12433. https://doi.org/10.1074/jbc.M100218200

Quenneville, S., Turelli, P., Bojkowska, K., Raclot, C., Offner, S., Kapopoulou, A., Trono, D., 2012. The KRAB-ZFP/KAP1 System Contributes to the Early Embryonic Establishment of Site-Specific DNA Methylation Patterns Maintained during Development. Cell Rep. 2, 766–773. https://doi.org/10.1016/j.celrep.2012.08.043

Remboutsika, E., Lutz, Y., Gansmuller, A., Vonesch, J.L., Losson, R., Chambon, P., 1999. The putative nuclear receptor mediator TIF1alpha is tightly associated with euchromatin. J. Cell Sci. 112, 1671–1683.

Rousseau-Merck, M.F., Koczan, D., Legrand, I., Möller, S., Autran, S., Thiesen, H.-J., 2002. The KOX zinc finger genes: genome wide mapping of 368 ZNF PAC clones with zinc finger gene clusters predominantly in 23 chromosomal loci are confirmed by human sequences annotated in EnsEMBL. Cytogenet. Genome Res. 98, 147–153. https://doi.org/10.1159/000069802

Rowe, H.M., Friedli, M., Offner, S., Verp, S., Mesnard, D., Marquis, J., Aktas, T., Trono, D., 2013. De novo DNA methylation of endogenous retroviruses is shaped by KRAB-ZFPs/KAP1 and ESET. Development 140, 519–529. https://doi.org/10.1242/dev.087585

Sadowski, I., Bell, B., Broad, P., Hollis, M., 1992. GAL4 fusion vectors for expression in yeast or mammalian cells. Gene 118, 137–141. https://doi.org/10.1016/0378-1119(92)90261-M

Sampson, D.A., Wang, M., Matunis, M.J., 2001. The small ubiquitin-like modifier-1 (SUMO-1) consensus sequence mediates Ubc9 binding and is essential for SUMO-1 modification. J. Biol. Chem. 276, 21664–21669. https://doi.org/10.1074/jbc.M100006200

Santoni de Sio, F.R., 2014. Kruppel-associated box (KRAB) proteins in the adaptive immune system. Nucleus 5, 138–148. https://doi.org/10.4161/nucl.28738

Schimmel, J., Eifler, K., Sigurðsson, J.O., Cuijpers, S.A.G., Hendriks, I.A., Verlaan-de Vries, M., Kelstrup, C.D., Francavilla, C., Medema, R.H., Olsen, J.V., Vertegaal, A.C.O., 2014. Uncovering SUMOylation Dynamics during Cell-Cycle Progression Reveals FoxM1 as a Key Mitotic SUMO Target Protein. Mol. Cell 53, 1053–1066. https://doi.org/10.1016/j.molcel.2014.02.001

Schultz, D.C., Ayyanathan, K., Negorev, D., Maul, G.G., Rauscher, F.J., 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. https://doi.org/10.1101/gad.973302

Schultz, D.C., Friedman, J.R., Rauscher, F.J., 2001. Targeting histone deacetylase complexes via KRAB-zinc finger proteins: the PHD and bromodomains of KAP-1 form a cooperative unit that recruits a novel isoform of the Mi-2α subunit of NuRD. Genes Dev. 15, 428–443. https://doi.org/10.1101/gad.869501

Schumacher, C., Wang, H., Honer, C., Ding, W., Koehn, J., Lawrence, Q., Coulis, C.M., Wang, L.L., Ballinger, D., Bowen, B.R., Wagner, S., 2000. The SCAN domain mediates selective oligomerization. J. Biol. Chem. 275, 17173–17179. https://doi.org/10.1074/jbc.M000119200

Shin, J.-H., Ko, H.S., Kang, H., Lee, Y., Lee, Y.-I., Pletinkova, O., Troconso, J.C., Dawson, V.L., Dawson, T.M., 2011. PARIS (ZNF746) Repression of PGC-1α Contributes to Neurodegeneration in Parkinson's Disease. Cell 144, 689–702. https://doi.org/10.1016/j.cell.2011.02.010

Shulman, J.M., Chibnik, L.B., Aubin, C., Schneider, J.A., Bennett, D.A., Jager, P.L.D., 2010. Intermediate Phenotypes Identify Divergent Pathways to Alzheimer's Disease. PLOS ONE 5, e11244. https://doi.org/10.1371/journal.pone.0011244

Stevens, D.A., Lee, Y., Kang, H.C., Lee, B.D., Lee, Y.-I., Bower, A., Jiang, H., Kang, S.-U., Andrabi, S.A., Dawson, V.L., Shin, J.-H., Dawson, T.M., 2015. Parkin loss leads to PARIS-dependent declines in mitochondrial mass and respiration. Proc. Natl. Acad. Sci. U. S. A. 112, 11696–11701. https://doi.org/10.1073/pnas.1500624112

Stoll, G.A., Oda, S., Chong, Z.-S., Yu, M., McLaughlin, S.H., Modis, Y., 2019. Structure of KAP1 tripartite motif identifies molecular interfaces required for retroelement silencing. Proc. Natl. Acad. Sci. 116, 15042–15051. https://doi.org/10.1073/pnas.1901318116

Sun, Y., Keown, J.R., Black, M.M., Raclot, C., Demarais, N., Trono, D., Turelli, P., Goldstone, D.C., 2019. A Dissection of Oligomerization by the TRIM28 Tripartite Motif and the Interaction with Members of the Krab-ZFP Family. J. Mol. Biol. https://doi.org/10.1016/j.jmb.2019.05.002

Thiesen, H.J., 1990. Multiple genes encoding zinc finger domains are expressed in human T cells. New Biol. 2, 363–374.

Thiesen, H.J., Bach, C., 1991a. Transition metals modulate DNA-protein interactions of SP1 zinc finger domains with its cognate target site. Biochem. Biophys. Res. Commun. 176, 551–557. https://doi.org/10.1016/s0006-291x(05)80219-0

Thiesen, H.J., Bach, C., 1991b. Determination of DNA binding specificities of mutated zinc finger domains. FEBS Lett. 283, 23–26. https://doi.org/10.1016/0014-5793(91)80545-e

Thiesen, H.J., Bach, C., 1990. Target Detection Assay (TDA): a versatile procedure to determine DNA binding sites as demonstrated on SP1 protein. Nucleic Acids Res. 18, 3203–3209. https://doi.org/10.1093/nar/18.11.3203

Thiesen, H.J., Bellefroid, E., Revelant, O., Martial, J.A., 1991. Conserved KRAB protein domain identified upstream from the zinc finger region of Kox 8. Nucleic Acids Res. 19, 3996.

Thiesen, H.J., Schröder, B., 1991. Amino acid substitutions in the SP1 zinc finger domain alter the DNA binding affinity to cognate SP1 target site. Biochem. Biophys. Res. Commun. 175, 333–338. https://doi.org/10.1016/s0006-291x(05)81239-2

Thomas, J.H., Schneider, S., 2011. Coevolution of retroelements and tandem zinc finger genes. Genome Res. 21, 1800–1812. https://doi.org/10.1101/gr.121749.111

Urnov, F.D., Rebar, E.J., Holmes, M.C., Zhang, H.S., Gregory, P.D., 2010. Genome editing with engineered zinc finger nucleases. Nat. Rev. Genet. 11, 636–646. https://doi.org/10.1038/nrg2842

Vaquerizas, J.M., Kummerfeld, S.K., Teichmann, S.A., Luscombe, N.M., 2009. A census of human transcription factors: function, expression and evolution. Nat. Rev. Genet. 10, 252–263. https://doi.org/10.1038/nrg2538

Vissing, H., Meyer, W.K.-H., Aagaard, L., Tommerup, N., Thiesen, H.-J., 1995. Repression of transcriptional activity by heterologous KRAB domains present in zinc finger proteins. FEBS Lett. 369, 153–157. https://doi.org/10.1016/0014-5793(95)00728-R

Wang, W., Cai, J., Wu, Y., Hu, L., Chen, Zongyun, Hu, J., Chen, Ze, Li, W., Guo, M., Huang, Z., 2013. Novel activity of KRAB domain that functions to reinforce nuclear localization of KRAB-containing zinc finger proteins by interacting with KAP1. Cell. Mol. Life Sci. 70, 3947–3958. https://doi.org/10.1007/s00018-013-1359-4

Williams, A.J., Blacklow, S.C., Collins, T., 1999. The Zinc Finger-Associated SCAN Box Is a Conserved Oligomerization Domain. Mol. Cell. Biol. 19, 8526–8535.

Witzgall, R., O'Leary, E., Leaf, A., Onaldi, D., Bonventre, J.V., 1994. The Krüppel-associated box-A (KRAB-A) domain of zinc finger proteins mediates transcriptional repression. Proc. Natl. Acad. Sci. U. S. A. 91, 4514–4518. https://doi.org/10.1073/pnas.91.10.4514

Wolfe, S.A., Nekludova, L., Pabo, C.O., 2000. DNA recognition by Cys2His2 zinc finger proteins. Annu. Rev. Biophys. Biomol. Struct. 29, 183–212. https://doi.org/10.1146/annurev.biophys.29.1.183

Wu, Y., Yu, L., Bi, G., Luo, K., Zhou, G., Zhao, S., 2003. Identification and characterization of two novel human SCAN domain-containing zinc finger genes ZNF396 and ZNF397. Gene 310, 193–201. https://doi.org/10.1016/S0378-1119(03)00551-1

Xiao, Z., Chang, J.-G., Hendriks, I.A., Sigurðsson, J.O., Olsen, J.V., Vertegaal, A.C.O., 2015. System-wide Analysis of SUMOylation Dynamics in Response to Replication Stress Reveals Novel Small Ubiquitin-like Modified Target Proteins and Acceptor Lysines Relevant for Genome Stability. Mol. Cell. Proteomics MCP 14, 1419–1434. https://doi.org/10.1074/mcp.O114.044792

Xing, M., Oksenych, V., 2019. Genetic interaction between DNA repair factors PAXX, XLF, XRCC4 and DNA-PKcs in human cells. FEBS Open Bio 9, 1315–1326. https://doi.org/10.1002/2211-5463.12681

Yang, L., Xia, L., Wu, D.Y., Wang, H., Chansky, H.A., Schubach, W.H., Hickstein, D.D., Zhang, Y., 2002. Molecular cloning of ESET, a novel histone H3-specific methyltransferase that interacts with ERG transcription factor. Oncogene 21, 148. https://doi.org/10.1038/sj.onc.1204998

Yu, E.J., Kim, S.-H., Kim, M.J., Seo, W.-Y., Song, K.-A., Kang, M.-S., Yang, C.K., Stallcup, M.R., Kim, J.H., 2013. SUMOylation of ZFP282 potentiates its positive effect on estrogen signaling in breast tumorigenesis. Oncogene 32, 4160–4168. https://doi.org/10.1038/onc.2012.420

Yuki, R., Aoyama, K., Kubota, Sho, Yamaguchi, Noritaka, Kubota, Shoichi, Hasegawa, H., Morii, M., Huang, X., Liu, K., Williams, R., Fukuda, M.N., Yamaguchi, Naoto, 2015. Overexpression of Zinc-Finger Protein 777 (ZNF777) Inhibits Proliferation at Low Cell Density Through Down-Regulation of FAM129A. J. Cell. Biochem. 116, 954–968. https://doi.org/10.1002/jcb.25046

10- Supplements

S1: *Protein equences of ZNF746 isoform a, ZNF746 isoform b, and ZNF777.*

>ZNF746a (UniProtKB:Q6NUN9)
MAEAVAAPIS**PWTMAATIQAMERKIESQAARLLSLEGRTGMAEKKLADCEKTAVEFGNQLEG**
KWAVLGTLLQEYGLLQRRLENVENLLRNRNFWILRLPPGSKGESPKEWGKLEDWQKELYKHV
MRGNYETLVSLDYAISKPEVLSQIEQGKEPCNWRRPGPKIPDVPVDPSPGSGPPVPAPDLLM
QIKQEGELQLQEQQALGVEAWAAGQPDIGEEPWGLSQLDSGAGDISTDATSGVHSNFSTTIP
PTSWQTDLPPHHPSSACSDGTLKLNTAASTEDVKIVIKTEVQEEEVVATPVHPTDLEAHGTL
FGPGQATRFFPSPAQEGAWESQGSSFPSQDPVLGLREPARPERDMGELSPAVAQEETPPGDW
LFGGVRWGWNFRCKPPVGLNPRTGPEGLPYSSPDNGEAILDPSQAPRPFNEPCKYPGRTKGF
GHKPGLKKHPAAPPGGRPFTCATCGKSFQLQVSLSAHQRSCGAPDGSGPGTGGGGSGSGGGG
GGSGGGSARDGSALRCGECGRCFTRPAHLIRHRMLHTGERPFPCTECEKRFTERSKLIDHYR
THTGVRPFTCTVCGKSFIRKDHLRKHQRNHAAGAKTPARGQPLPTPPAPPDPFKSPASKGPL
ASTDLVTDWTCGLSVLGPTDGGDM

EXON1- DUF3669(amino acids in bald are PF12417(EXON2)-KRAB-A(EXON3)-KRAB-B(EXON4)-EXON5-EXON6-Zinc Finger Domain (highlighted motifs are Zinc fingers) (EXON7)

>ZNF746b (UniProtKB:A0A2R8YDQ5)
MAEAVAAPIS**PWTMAATIQAMERKIESQAARLLSLEGRTGMAEKKLADCEKTAVEFGNQLEG**
KWAVLGTLLQEYGLLQRRLENVENLLRNRNFWILRLPPGSKGESPKVPVTFDDVAVYFSEQE
WGKLEDWQKELYKHVMRGNYETLVSLDYAISKPEVLSQIEQGKEPCNWRRPGPKIPDVPVDP
SPGSGPPVPAPDLLMQIKQEGELQLQEQQALGVEAWAAGQPDIGEEPWGLSQLDSGAGDIST
DATSGVHSNFSTTIPPTSWQTDLPPHHPSSACSDGTLKLNTAASTEDVKIVIKTEVQEEEVV
ATPVHPTDLEAHGTLFGPGQATRFFPSPAQEGAWESQGSSFPSQDPVLGLREPARPERDMGE
LSPAVAQEETPPGDWLFGGVRWGWNFRCKPPVGLNPRTGPEGLPYSSPDNGEAILDPSQAPR
PFNEPCKYPGRTKGFGHKPGLKKHPAAPPGGRPFTCATCGKSFQLQVSLSAHQRSCGAPDGS
GPGTGGGGSGSGGGGGGSGGGSARDGSALRCGECGRCFTRPAHLIRHRMLHTGERPFPCTEC
EKRFTERSKLIDHYRTHTGVRPFTCTVCGKSFIRKDHLRKHQRNHAAGAKTPARGQPLPTPP
APPDPFKSPASKGPLASTDLVTDWTCGLSVLGPTDGGDM

EXON1- DUF3669(amino acids in bald are PF12417(EXON2)-KRAB-A(EXON3)-KRAB-B(EXON4)-EXON5-EXON6-Zinc Finger Domain (highlighted motifs are Zinc fingers) (EXON7)

>ZNF777 (UniProtKB:Q9ULD5-2)
MENQRSSPLSFPSVPQEETLRQAPAGLPRETLFQSRVLPPKEIPSLSPTIPRQGSLPQTSSA
PKQETSGRMPHVLQKGPSLLCSAASEQETSLQGPLASQEGTQYPPPAAAEQEVSLLSHSPHH
QEAPVHSPEAPEKDPLTLSPTVPETDMDPLLQSPVSQKDTPFQISSAVQKEQPLPTAEITRL
AV**WAAVQAVERKLEAQAMRLLTLEGRTGTNEKKIADCEKTAVEFANHLESKWVVLGTLLQEY**
GLLQRRLENMENLLKNRNFWILRLPPGSNGEVPKVPVTFDDVAVHFSEQEWGNLSEWQKELY
KNVMRGNYESLVSMDYAISKPDLMSQMERGERPTMQEQEDSEEGETPTDPSAAHDGIVIKIE
VQTNDEGSESLETPEPLMGQVEEHGFQDSELGDPCGEQPDLDMQEPENTLEESTEGSSEFSE
LKQMLVQQRNCTEGIVIKTEEQDEEEEEEEEDELPQHLQSLGQLSGRYEASMYQTPLPGEMS
PEGEESPPPLQLGNPAVKRLAPSVHGERHLSENRGASSQQQRNRRGERPFTCMECGKSFRLK
INLIIHQRNHIKEGPYECAECEISFRHKQQLTLHQRIHRVRGGCVSPERGPTFNPKHALKPR
PKSPSSGSGGGGPKPYKCPECDSSFSHKSSLTKHQITHTGERPYTCPECKKSFRLHISLVIH
QRVHAGKHEVSFICSLCGKSFSRPSHLLRHQRTHTGERPFKCPECEKSFSEKSKLTNHCRVH

SRERPHACPECGKSFIRKHHLLEHRRIHTGERPYHCAECGKRFTQKHHLLEHQRAHTGERPY
PCTHCAKCFRYKQSLKYHLRTHTGE

DUF3669(Amino acids in bald are PF12417) (EXON1)-KRAB-A (EXON2)-
KRAB-B (EXON 3)-Exon4-ZFs-containing region (highlighted motifs are
zinc fingers) (EXON5)

S2: Zinc finger recognition code of human DUF3669-containing zinc finger proteins

For each individual zinc finger, residues -1, 3 and 6 with respect to the start of the α-helix (that were proposed to be especially important for DNA-binding specificity) were extracted and consecutively written down from the N- to the C-terminal end of the whole array.

```
ZNF398 HRY-DRS-LSL-GAR-RHN-RHK-YTD
ZNF282 VSI-CGR-RHN-RNK-YSD
ZNF212 YQT-HDR-QHQ
ZNF777 LNI-HQL-HSK-LSI-RSR-EKN-RHE-QHE-YSY
ZNF783 LNI-RDR-RHV
ZNF746 LSA-RHR-EKD-RHK

-1,3,6 amino acids:
>ZNF398_HS
HRY-DRS-LSL-GAR-RHN-RHK-YTD
                            1234567
ZNF398HS_ZF_01  PTCPHCARTFTHPSRLTYHLRVH
ZNF398HS_ZF_02  FPCPDCPKRFADQARLTSHRRAH
ZNF398HS_ZF_03  FRCAQCGRSFSLKISLLLHQRGH
ZNF398HS_ZF_04   SCPQCGIDFNGHSALIRHQMIH
ZNF398HS_ZF_05  YPCTDCSKSFMRKEHLLNHRRLH
ZNF398HS_ZF_06  FSCPHCGKSFIRKHHLMKHQRIH
ZNF398HS_ZF_07  YPCSYCGRSFRYKQTLKDHLRSGH

>ZNF282_HS
VSI-CGR-RHN-RNK-YSD
ZNF282HS_ZF_01  YSCPECGKSFGVRKSLIIHHRSH
ZNF282HS_ZF_02  YECAECEKSFNCHSGLIRHQMTH
ZNF282HS_ZF_03  YKCSECEKTYSRKEHLQNHQRLH
ZNF282HS_ZF_04  FQCALCGKSFIRKQNLLKHQRIH
ZNF282HS_ZF_05  YTCGECGKSFRYKESLKDHLRVH

>ZNF212_HS
YQT-HDR-QHQ
ZNF212HS_ZF_01  YECSECEITFRYKQQLATHLRSH
ZNF212HS_ZF_02  LICGYCGKSFSHPSDLVRHQRIH
ZNF212HS_ZF_03  YSCTECEKSFVQKQHLLQHQKIH

>ZNF777_HS
LNI-HQL-HSK-LSI-RSR-EKN-RHE-QHE-YSY
ZNF777HS_ZF_01  FTCMECGKSFRLKINLIIHQRNH
ZNF777HS_ZF_02  YECAECEISFRHKQQLTLHQRIH
ZNF777HS_ZF_03  YKCPECDSSFSHKSSLTKHQITH
ZNF777HS_ZF_04  YTCPECKKSFRLHISLVIHQRVH
ZNF777HS_ZF_05  FICSLCGKSFSRPSHLLRHQRTH
ZNF777HS_ZF_06  FKCPECEKSFSEKSKLTNHCRVH
ZNF777HS_ZF_07  HACPECGKSFIRKHHLLEHRRIH
ZNF777HS_ZF_08  YHCAECGKRFTQKHHLLEHQRAH
ZNF777HS_ZF_09  YPCTHCAKCFRYKQSLKYHLRTH

>ZNF783_HS
LNI-RDR-RHV
ZNF783HS_ZF_01  FPCPDCGQSFRLKINLTIHQRTH
ZNF783HS_ZF_02  PACPYCGKAFRRPSDLFRHQRIH
ZNF783HS_ZF_03  YQCPQCGRTFNRNHHLAVHMQTH

>ZNF746_HS
LSA-RHR-EKD-RHK
ZNF746HS_ZF_01  FTCATCGKSFQLQVSLSAHQRSC
ZNF746HS_ZF_02  LRCGECGRCFTRPAHLIRHRMLH
ZNF746HS_ZF_03  FPCTECEKRFTERSKLIDHYRTH
ZNF746HS_ZF_04  FTCTVCGKSFIRKDHLRKHQRNH
```

***S3**: List of repression factors achieved by ZNF10 and GAL4 in HeLa cells used in Fig.15B, 15C, 15D and repression factors achieved by ZNF10 in HAP1 wildtype and knockout cells used in Fig. 17A, 17B, 17C.*

HeLa		**HAP1-Wildtype**	**TRIM28KO**	**SETDB1KO**
Gal4	**ZNF10AB**	**ZNF10AB**	**ZNF10AB**	**ZNF10AB**
0,97366808	12,50520914	50,45913359	1,478939356	3,656835502
1,02633192	13,09945775	45,00753292	1,437191973	3,296664887
0,96934805	11,62159531	47,56336486	1,403518083	3,843987534
1,03065195	14,32881065	46,6781602	1,357914386	3,69032766
1,10912148	10,470916	40,28486762	1,566484636	3,674279257
0,89087852	10,42398343	43,6563235	1,407605408	2,969034727
1,0098002	21,26456711	36,33583383	1,329019285	3,706977365
0,9901998	22,1467391	42,49818855	1,242886801	3,960148513
1,01331501	18,93671614	29,60964175	1,008224623	3,752404825
0,98668499	19,5204579	35,00268232	1,130681765	4,183090042
0,98508653	10,26297016	25,86091198	1,059044327	3,928262396
1,01491347	12,22212254	26,32035632	1,086609543	4,127453484
1,01841572	16,18807922	31,91513024	2,004563839	4,921075901
0,98158428	14,96601217	30,46001873	1,99393062	4,987182001
1,0776672	15,6077092	27,20521046	1,83283882	3,893163155
0,9223328	13,71760544	27,0455012	1,732456343	4,129676385
0,96165551	18,57950018	33,32609557	1,493861902	4,136550046
1,03834449	18,44051569	32,80160352	1,49940949	4,439783162
0,9553938	10,98045607	21,18490541	1,18282444	5,650764199
1,0446062	11,21397974	20,6841286	1,134214783	5,66300918
1,04342659	12,72144182	16,19680341	1,702390109	4,877554107
0,95657341	11,74156068	15,77813952	1,662287344	5,182000782
1,10723202	15,3453358	38,477684	1,46085738	4,101300758
0,89276798	15,30151655	36,25738881	1,19177303	4,768663935
0,95762903	7,238729593	29,23969602	1,051731141	4,837995265
1,04237097	6,922383943	25,66391791	1,194935126	4,977716808
1,01171551	25,22110714	39,28697104	1,25410485	3,719907047
0,98828449	27,05648653	45,01829826	1,178347509	3,960696313
0,87462603	22,66681748	41,52768757	1,187254095	3,625022204
1,12537397	21,70332869	40,95601601	1,40089945	4,372124336
1,00656545	24,55139298	22,09838463	1,131931905	3,318306561
0,99343455	21,10101902	20,29681716	1,177357748	3,453545482
1,05037037	24,04568076	43,24086058	1,186830631	
0,94962963	22,95987061	41,14441158	1,267656263	
1,0042104	21,30135439	40,76443291	1,093512516	
0,9957896	24,98981635	40,12305434	1,147535178	
0,94701254	22,55963076	40,48324117	1,124774677	
1,05298746	20,34271703	41,73334809	1,109052704	
0,88584095		37,97758611		
1,11415905		37,66221718		
		34,09019865		

		35,16603165		
N				
40	38	42	38	32
Average				
1	**16,95441034**	**34,69244709**	**1,33961716**	**4,181421994**
SD				
0,061327714	5,518666892	8,87533874	0,260468963	0,675726577
unpaired two-tailed T-test				
	ZNF10 vs GAL4	**HAP1 wildtype vs TRIM28 ko**	**TRIM28ko vs SETDB1ko**	
	9,493E-20	5,0263E-26	8,52894E-24	
		HAP1 wildtype vs SETDB1ko		
		1,09408E-24		

***S4**: List of repression factors achieved by ZNF746a fragments in HeLa cells used in Fig.15B*

Z746a / 1-644	**Z746a/ 1-279**	**Z746a/ 1-173**	**Z746a/ 1-108**	**Z746a/ 94-279**	**Z746a/ 94-173**	**Z746a/ 174-279**	**Z746a/ 280-644**
2,100006478	4,18975412	1,38474845	2,94150750	2,57850254	1,107183431	0,967865185	1,683788228
1,911794348	3,80785741	1,35096665	2,7364057	3,22043387	0,81952296	1,02829344	1,50510219
4,728283484	4,63197353	1,47739652	2,40339932	2,70408124	0,996753248	1,779295482	2,230220305
3,674090738	4,07964799	1,42579461	2,33516077	2,96951539	1,046916862	1,885170547	2,474123064
3,423215021	4,56529251	1,41362184	3,22430345	2,52859120	0,992911939	1,640693366	2,370686141
3,939516304	4,16729932	1,40444861	2,75339911	2,30878616	1,060777419	1,51749689	2,385510311
2,134253979						1,626957983	1,492198234
2,677473714						1,691690602	1,532118913
N							
8	6	6	6	6	6	8	8
Average							
3,0736	**4,2403**	**1,4095**	**2,7324**	**2,7183**	**1,0040**	**1,5172**	**1,9592**
SD							
1,02	0,31	0,04	0,33	0,33	0,10	0,34	0,44
unpaired two-tailed T-test of repression factors of tested evector plasmids against GAL4 alone							
0,000703333	1,52181E-06	1,37172E-08	4,95088E-05	4,84164E-05	0,927131896	0,003435123	0,000469123

***S5**: List of repression factors achieved by ZNF746b fragments in HeLa cells used in Fig.15C*

Z746b/1-294	**Z746b/1-188**	**Z746b/109-294**	**Z746b/109-188**
12,2450472	4,985895063	5,461520518	3,054794033
12,02701075	4,379831785	5,078700287	2,977315711
13,76963434	5,715670565	9,818317823	3,240247619
14,46929188	5,206681633	9,746710717	2,905634087
10,65289426	5,67474417	9,456061966	3,128137536
11,89422973	5,396785243	11,11068062	3,011387956
		6,740849355	

		6,755229155	
N			
6	6	8	6
Average			
12,5097	**5,2266**	**8,0210**	**3,0529**
SD			
1,382533625	0,498866365	2,27605493	0,118225958
unpaired two-tailed T-test of repression factors of tested evector plasmids against GAL4 alone			
5,22224E-06	4,6407E-06	5,21123E-05	5,36865E-08

***S6**: List of repression factors achieved by ZNF777 fragments in HeLa cells used in Fig.15D*

Z777/1-831	**Z777/1-362**	**Z777/1-282**	**Z777/189-254**	**Z777/283-362**
2,725625453	1,316557756	2,50703925	2,033601603	1,402093462
2,821105671	1,092337521	2,60998511	2,062353052	1,514022989
5,902834278	0,973358559	2,22350114	2,107847722	1,613482031
5,783553427	1,012607804	1,75480319	1,983337124	1,583279111
5,283019912	1,216146524	2,26481408	2,11057548	1,599818781
5,504130765	1,222582295	2,04933742	1,813110739	1,602521947
2,993613453	1,101349869	2,66063126		1,332493982
2,970268961	1,192821338	2,87887998		1,512445599
	1,511270623	3,02679772		1,454642166
	1,160825488	3,33750473		1,341400558
	1,177028502	2,74489694		1,362275913
	1,047104469	2,41029877		1,296219345
N				
8	12	12	6	12
Average				
4,2480	**1,1687**	**2,5390**	**2,0185**	**1,4679**
SD				
1,478670515	0,145464776	0,43696665	0,111385124	0,118174603
unpaired two-tailed T-test of repression factors of tested evector plasmids against GAL4 alone				
0,000438964	0,001998875	8,9807E-08	1,55687E-06	7,79018E-09

***S7**: List of repression factors achieved by ZNF746a fragments in HAP1 Wildtype cells used in Fig.17A*

			HAP1 Wildtype			
Z746a	**Z746a/ 1-279**	**Z746a/ 1-173**	**Z746/ 1-108**	**Z746a/ 94-279**	**Z746a/ 94-173**	**Z746/ 280-644**
3,217	8,819	2,361	4,639	1,795	1,059	2,405617962
2,286	9,253	2,087	4,103	1,691	1,070	2,516815717
3,633	6,506	2,685	4,884	1,540	0,955	2,45769486

3,323	6,964	2,335	4,788	1,633	1,030	2,508287577
3,668	6,107	2,857	4,570	1,967	1,120	2,23599923
3,900	5,725	2,702	4,614	2,033	0,916	2,427438718
3,896		2,565		1,989		3,874032155
4,002		2,335		1,793		3,962742237
5,565		3,102		2,613		
5,854		3,153		2,645		
N						
10	6	10	6	10	6	8
Average						
3,93	**7,23**	**2,62**	**4,60**	**1,97**	**1,03**	**2,80**
1,060153	1,465627093	0,349673721	0,270432801	0,381719175	0,075958498	0,696972706
unpaired two-tailed T-test of repression factors of tested construct in Wildtype against TRIM28 KO						
0,477831	0,001278878	0,012502185	7,68607E-09	0,441149443	0,509680715	0,85887289
unpaired two-tailed T-test of repression factors of tested construct in Wildtype against SETDB1 KO						
0,220344	0,000385263	9,55625E-07	5,2071E-07	0,000241891	0,243939532	0,001053836

S8*: List of repression factors achieved by ZNF746a fragments in HAP1 –TRIM28ko cells used in Fig.17A*

			HAP1-TRIM28ko			
Z746a	**Z746a/ 1-279**	**Z746a/ 1-173**	**Z746/1-108**	**Z746a/ 94-279**	**Z746a/ 94-173**	**Z746/ 280-644**
2,140917	5,22851807	1,573043041	2,328215696	1,603509731	1,029611985	1,740754369
1,862196	5,34470525	1,61107775	2,52584088	1,661894119	0,983229237	1,813613617
2,421793	3,124443719	1,496783047	2,167244467	1,372529359	1,031997709	3,526382529
2,366033	3,010068445	1,551057551	2,535678056	1,36427399	0,969565574	4,139023202
5,602894	2,905786972	2,60731011	1,969470625	2,428298174	1,008737411	2,986941136
5,626056	3,061689723	2,418919441	1,88974283	2,302719168	0,989763058	3,088701285
			2,556909129			
			2,230588088			
N						
6	6	6	8	6	6	6
Average						
3,34	**3,78**	**1,88**	**2,28**	**1,79**	**1,00**	**2,88**
SD						
1,775393	1,170407622	0,498183702	0,258714437	0,464114633	0,025540622	0,947941279
unpaired two-tailed T-test of repression factors of tested construct in TRIM28 KO against SETDB1 KO						
0,720026	0,02520176	0,101283438	0,014020124	0,045142674	0,365357247	0,01527356

***S9**: List of repression factors achieved by ZNF746a fragments in HAP1 –SETDB1ko cells used in Fig.17A*

HAP1-SETDB1ko						
Z746a	**Z746a/ 1-279**	**Z746a/ 1-173**	**Z746/1-108**	**Z746a/ 94-279**	**Z746a/ 94-173**	**Z746/ 280-644**
4,829447	2,410847118	1,477466295	1,855678411	1,222510288	0,961477388	1,431417948
5,124717	2,404680499	1,351970008	1,912608062	1,139309824	1,088231372	1,543740103
1,981404	2,350473142	1,425638674	1,932080757	1,273256381	0,98502512	1,546305805
1,996832	2,363271852	1,537435725	1,890987667	1,352424593	0,925453839	1,511556813
1,994105	2,064393759	1,595404393	1,998427397	1,352066141	0,984393727	1,421390664
1,964666	2,076154748	1,417235207	2,122879841	1,38545575	0,902867031	1,455146165
			2,089536203			
			2,022051125			
N						
6	6	6	8	6	6	6
Average						
2,9819	**2,2783**	**1,4675**	**1,9780**	**1,2875**	**0,9746**	**1,4849**
SD						
1,548351	0,16284258	0,088290387	0,096144731	0,094197866	0,064575308	0,056075785

***S10**: List of repression factors achieved by ZNF746b fragments in HAP1 –Wildtype cells used in Fig.17B.*

HAP1-Wildtype			
Z746b/1-294	**Z746b/1-188**	**Z746b/109-294**	**Z746b/109-188**
24,463	5,002	3,835	2,973
24,467	5,081	3,881	3,085
14,945	4,565	3,480	3,185
12,982	5,042	3,427	3,020
16,078	13,817	10,245	3,702
16,209	13,350	11,639	3,560
	13,811	9,950	
	12,530	8,159	
	19,586	17,047	
	21,825	17,445	
N			
6	10	10	6
Average			
18,19	**11,46**	**8,91**	**3,25**
SD			
4,995739825	6,328005801	5,380171995	0,303934596
unpaired two-tailed T-test of repression factors of tested construct in Wildtype against TRIM28 KO			
0,000625484	0,000698578	0,001432716	3,04962E-06
unpaired two-tailed T-test of repression factors of tested construct in Wildtype against SETDB1 KO			
0,000544251	0,000692745	0,002376444	3,45731E-06

***S11**: List of repression factors achieved by ZNF746b fragments in HAP1 –TRIM28ko cells used in Fig.17B*

HAP1-TRIM28ko			
Z746b/1-294	**Z746b/1-188**	**Z746b/109-294**	**Z746b/109-188**
4,68260052	1,30036185	1,22159191	0,45764483
4,78291593	1,32818494	1,246936	0,47996225
2,61088643	1,16696447	0,90845322	0,47891294
2,81015789	1,1695091	0,9443227	0,46582325
2,83973427	1,65192982	1,50769023	0,44344997
2,95912829	1,67124519	1,47178542	0,43965541
N			
6	6	6	6
Average			
3,447570555	**1,381365895**	**1,21679658**	**0,460908108**
SD			
1,002279633	0,226925532	0,252937718	0,017195963
unpaired two-tailed T-test of repression factors of tested construct in TRIM28 KO against SETDB1 KO			
0,030051347	0,843204721	0,236205995	1,46934E-05

***S12**: List of repression factors achieved by ZNF746b fragments in HAP1 –SETDB1ko cells used in Fig.17B*

HAP1-SETDB1ko			
Z746b/1-294	**Z746b/1-188**	**Z746b/109-294**	**Z746b/109-188**
2,17816671	1,37609171	3,40988972	0,74709723
2,31387584	1,46274573	3,25785031	0,64448322
2,28182742	1,29463642	1,07495617	0,78695764
2,21092339	1,34322137	1,05655326	0,7619407
2,11695834	1,40605412	1,1746402	0,78425047
2,22321976	1,285522	1,15473781	0,73517808
N			
6	6	6	6
Average			
2,2208	**1,3614**	**1,8548**	**0,7433**
SD			
0,070840022	0,067861771	1,147598613	0,052489132

***S13**: List of repression factors achieved by ZNF777 fragments in HAP1 –Wildtype cells used in Fig.17C.*

HAP1-Wildtype				
Z777/1-831	**Z777/1-362**	**Z777/1-282**	**Z777/189-254**	**Z777/283-362**
4,292	2,616	5,187	1,415	1,061
3,958	2,735	5,636	1,696	1,034
6,459	1,722	3,486	1,307	1,146
6,468	1,845	3,728	1,400	1,219

5,683	1,371	3,495	1,461	1,211
5,786	1,507	2,882	1,368	1,202
6,483				
5,898				
8,620				
8,987				
N				
10	6	6	6	6
Average				
6,26335	**1,96580**	**4,06885**	**1,44124**	**1,14559**
SD				
1,598562017	0,575066732	1,086354077	0,134964166	0,08039312
unpaired two-tailed T-test of repression factors of tested construct in Wildtype against TRIM28 KO				
0,020933551	0,084972165	0,006474081	0,043354969	0,004639678
unpaired two-tailed T-test of repression factors of tested construct in Wildtype against SETDB1 KO				
0,000533537	0,02457104	0,00276001	0,001412723	7,29727E-06

***S14**: List of repression factors achieved by ZNF777 fragments in HAP1 –TRIM28ko cells used in Fig.17C.*

		HAP1-TRIM28ko		
Z777/1-831	Z777/1-362	Z777/1-282	Z777/189-254	Z777/283-362
2,3340	1,5587	2,0747	1,2328	0,9952
2,2659	1,6320	2,0508	1,1966	0,9728
2,9531	1,4500	2,3180	1,3501	0,8730
2,9053	1,6877	2,1788	1,3590	0,8010
5,8630	1,2777	2,0730	1,3465	0,8362
6,3312	1,4415	1,9396	1,2585	1,1073
	1,3480	2,0358		
	1,2932	2,0414		
N				
10	12	12	10	10
Average				
3,78	**2,32**	**2,60**	**2,25**	**1,79**
SD				
1,826463473	0,153649147	0,113309621	0,070063252	0,11508773
unpaired two-tailed T-test of repression factors of tested construct in TRIM28 KO against SETDB1 KO				
0,503934494	0,002693851	1,37196E-06	0,007694671	0,022142931

***S15**: List of repression factors achieved by ZNF777 fragments in HAP1 –SETDB1ko cells used in Fig.17C.*

		HAP1-SETDB1		
Z777/1-831	**Z777/1-362**	**Z777/1-282**	**Z777/189-254**	**Z777/283-362**
4,8063	1,2833	1,6760	1,0574	0,7395
4,3748	1,3285	1,5109	1,0377	0,8890
2,2089	1,2910	1,6366	1,2795	0,7515

2,3950	1,2259	1,6819	1,1309	0,7841
2,3280	1,1654	1,6399	1,1679	0,7948
2,8600	1,0998	1,6655	1,1519	0,7115
	1,2438	1,6630		
	1,1412	1,6708		
N				
6	8	8	6	6
Average				
3,1622	**1,2224**	**1,6431**	**1,1375**	**0,7784**
SD				
1,136503908	0,080168478	0,055767311	0,086868051	0,062055537

***S16**: List of repression factors achieved by ZNF10, KRAB domain of ZNF746b and mutants in HAP1 –Wildtype cells used in Fig. 20B.*

		HAP-Wildtype		
Z10AB	**Z746b/109-188**	**Z746b/109-188_R141L**	**Z746b/109-188_G142E**	**Z746b/109-188_RG/LE**
43,241	2,973	2,835	40,855	39,871
41,144	3,084	2,865	42,572	40,906
40,764	3,185	2,586	37,504	37,471
40,123	3,020	2,677	36,149	39,176
40,483	3,702	4,065	46,385	48,553
41,733	3,560	4,485	42,423	48,276
31,915				
30,460				
27,205				
27,046				
33,326				
32,802				
21,185				
20,684				
41,528				
40,956				
22,098				
20,297				
N				
18	6	6	6	6
Average				
33,1662	**3,2539**	**3,2522**	**40,9813**	**42,3755**
SD				
8,39582163	0,303948225	0,810048581	3,72206508	4,81026154
unpaired two-tailed T-test of repression factors				
			vs WT 109-188	
		0,996296001	1,7594E-06	5,5383E-06
			GE vs RG/LE	

				0,58759067

***S17**: List of repression factors achieved by ZNF10, KRAB domain of ZNF746b and mutants in HAP1 –TRIM28ko cells used in Fig. 20B.*

		TRIM28KO		
Z10AB	**Z746b/109-188**	**Z746b/109-188_R141L**	**Z746b/109-188_G142E**	**Z746b/109-188_RG/LE**
1,494	0,458	0,494	0,452	0,390
1,499	0,480	0,390	0,449	0,452
1,183	0,479	0,422	0,419	0,428
1,134	0,466	0,420	0,401	0,422
1,702	0,443	0,476	0,445	0,542
1,662	0,440	0,477	0,473	0,570
1,187				
1,268				
1,094				
1,148				
1,125				
1,109				
N				
12	6	6	6	6
Average				
1,3004	**0,4609**	**0,4465**	**0,4397**	**0,4674**
SD				
0,22535438	0,01719596	0,0415898	0,02560241	0,07218706
unpaired two-tailed T-test of repression factors				
			vs WT 109-188	
		0,4585975	0,12808083	0,83886963
			GE vs RG/LE	
			0,40986848	

***S18**: List of repression factors achieved by ZNF10, KRAB domain of ZNF746b and mutants in HAP1 –SETDB1ko cells used in Fig. 20B*

		SETDB1KO		
Z10AB	**Z746b/109-188**	**Z746b/ 109-188_R141L**	**Z746b/ 109-188_G142E**	**Z746b/ 109-188_RG/LE**
3,625	0,787	0,770	2,960	3,009
4,372	0,762	0,735	3,151	2,972
3,318	0,784	0,758	3,284	3,693
3,454	0,735	0,745	3,314	3,236
3,752	0,747	0,901	4,589	3,939
4,183	0,644	0,886	4,452	3,824
3,893				
4,130				

4,137				
4,440				
5,651				
5,663				
N				
12	6	6	6	6
Average				
4,2181	**0,7433**	**0,7991**	**3,6250**	**3,4455**
SD				
0,75662048	0,05248913	0,07420585	0,70605679	0,42594089
unpaired two-tailed T-test of repression factors				
			vs WT 109-188	
		0,16727129	0,00016299	1,6485E-05
			GE vs RG/LE	
			0,60812323	

***S19**: List of repression factors achieved by ZNF10, ZNF10-DUF3669 chimera in HAP1-wildtype, HAP1-TRIM28ko and HAP1-SETDB1ko cells used in Fig. 21A.*

HAP1-Wildtype			**TRIM28KO**			**SETDB1KO**		
Z10AB	**Z746a/1-108_Z10AB**	**Z777/189-254_Z10AB**	**Z10AB**	**Z746a/1-108_Z10AB**	**Z777/189-254_Z10AB**	**Z10AB**	**Z746a/1-108_Z10AB**	**Z777/189-254_Z10AB**
36,3358	33,8379	32,8450	1,4789	2,7177	1,4688	4,8776	4,3225	3,9139
42,4982	35,9857	38,0455	1,4372	2,7294	1,4738	5,1820	3,9371	3,9431
29,6096	31,9580	28,6842	1,4035	2,8425	1,6236	4,1013	4,1416	3,5864
35,0027	30,8477	28,9311	1,3579	2,9286	1,6971	4,7687	3,4052	4,2003
25,8609	27,1385	25,7084	1,5665	2,7532	1,4775	4,8380	4,1196	4,3554
26,3204	24,6391	28,0109	1,4076	2,6698	1,4698	4,9777	3,8377	4,6752
N								
6	6	6	6	6	6	6	6	6
Average								
32,60460	**30,73449**	**30,37084**	**1,44194**	**2,77352**	**1,53508**	**4,79087**	**3,96062**	**4,11238**
SD								
6,50376	4,214269	4,410159	0,07296	0,09492	0,0998	0,36705	0,3203676	0,3816353
unpaired two-tailed T-test of repression factors "with DUF" vs "without DUF"								
	0,5697198	0,5042210		3,0075E-10	0,09750518		0,00198126	0,01055199

***S20**: List of repression factors achieved by the amino half wildtype of ZNF746a and the SUMO-receptor lysine mutant in HeLa and HAP1 cells used in Fig. 21A.*

HeLa		**HAP1-Wildtype**	
Z746a/1-279	**Z746a/1-279K189R**	**Z746a/1-279**	**Z746a/1-279K189R**
4,189754125	1,243734892	8,81861139	0,849927679
3,807857417	0,93061514	9,25319955	0,886695303

4,631973538	1,371008119	6,50572295	0,674472076
4,079647999	1,433332102	6,96357528	0,63169645
4,565292519	1,296789279	6,10672447	0,867370571
4,167299328	1,235841755	5,72467831	0,904583389
	1,400966778		
	1,527053951		
N			
6	8	6	6
Average			
4,2403	**1,3049**	**7,2288**	**0,8025**
SD			
0,309750846	0,180519663	1,46562709	0,117925769
unpaired two-tailed T-test of repression factors mutant vs wildtype			
6,55799E-08		0,000114022	

11- Acknowledgment

It would not have been possible to accomplish this Ph.D. without the support and guidance that I received from many people.

Foremost, I would like to express my sincere gratitude to my supervisor **Prof. Dr. Hans-Jürgen Thiesen** for giving me the opportunity to pursue my Ph.D. studies and for his continuous guidance, encouragement, and advice.

I am deeply indebted to **Dr. Peter Lorenz** not only for providing me with a fantastic lab training, but also for his follow up, patience, motivation, enthusiasm, and immense knowledge. I have been extremely lucky to have such a person who cared so much about my work, and who responded to my questions and queries so promptly.

My sincere thanks also go to **Ms. Ildiko Toth** for assistance and technical support with DNA preparations and cloning.

I would also extend my profound gratitude to **Mr. Felix Steinbeck** who has been so helpful and cooperative during the bioinformatic analysis of protein sequences.

I am sincerely thankful to **the German academic exchange service** (DAAD), Leadership for Syria program (LFS). This dissertation would not have been possible without their funding support.

My acknowledgment would be incomplete without thanking the biggest source of my strength, my family. My parents who prayed for me throughout the time of my research. My wife **Asmaa**, and my children **Mohammad** and **Lama** for unwavering and unselfish love and support given to me at all times.

12- Declaration of Originality

I hereby declare that this Ph.D. Thesis entitled "**Transcription Factor Properties of ZNF746 and ZNF777 with Focus on their Domains of Unknown Function 3669 (DUF3669)**" is the result of my original work and was carried out under the supervision of Prof. Dr. H-J Thiesen, Institute of Immunology, University Medicine Rostock, University of Rostock.

All sources used in this work are clearly referenced and the assistance I received in this work is fully acknowledged.

Furthermore, I confirm that no degree has been granted to me for this Ph.D. thesis or any part of it before, either at the University of Rostock or any other university or scientific institution.

Mohannad Al Chiblak Rostock, 06.10.2020