1. Scientific goals and research hypotheses of the project

Cancer is one of the leading cause of deaths worldwide. Carcinogenesis is a complex process, in which normal cells acquire cancerous properties, such as sustained proliferation or escape from the immune system control. The process is driven mainly by aberrant functioning of mutated genes, however epigenetic changes, such as DNA methylation and histone modification play a pivotal role in the initiation and progression of cancer. It is well established that cancer genomes become globally hypomethylated, whereas many promoters of tumor suppressor genes are subject to hypermethylation and subsequent transcriptional repression. Nevertheless, the precise mechanisms involved in targeting the distinct tumor suppressors are not yet clarified.

Our goal is to investigate the involvement of transcriptional repressors KRAB-ZNFs (Krüppel-associated box–zinc finger proteins) in modulating epigenetic profile of cancer cells. Preliminary statistical analysis utilizing the transcriptomic data deposited in the TCGA (The Cancer Genome Atlas) databases provided evidence that expression of a distinct subgroup of KRAB-ZNFs is elevated in multiple tumor types. Since KRAB-ZNFs are potent transcriptional repressors, we hypothesize that these proteins may be involved in the epigenetic regulation of tumor suppressors genes. We aim to determine the consensus binding sites of cancer-associated KRAB-ZNFs and investigate their influence on chromatin and DNA methylation status in cancer cells. Moreover, we hypothesize that through epigenetic regulation of the expression of tumor suppressor genes, KRAB-ZNFs may have an impact on cancer cell biology. Thus, we intent to perform a series of *in vitro* functional tests evaluating an influence of KRAB-ZNFs silencing on cancer cell behavior. Moreover, we aim to correlate the expression level of selected KRAB-ZNFs with the clinical outcome in cancer patients and with the molecular profiles of their tumor samples. Obtained results will provide new, valuable insights on KRAB-ZNFs-dependent epigenetic mechanisms occurring during cancer development.

2. Significance of the project

2.1 State-of-the-art

Tumorigenesis is a complex disease driven by numerous genetic and epigenetic alterations. While genetic aberrations has been already relatively well documented, epigenomic changes in cancer cell are being currently extensively studied. It is well established that DNA methylation abnormalities, as well as chromatin structure modifications have large impact on gene expression in cancer cells (Figure 1.1). Cancer epigenome is globally hypomethylated, whereas hypermethylation occurs on the CpG islands localized within promoters of tumor suppressors and DNA repair genes (Esteller, 2008). This phenomenon, known as CpG island methylator phenotype (CIMP), was initially described in colorectal cancer. Since then, this term has been used to describe association between aberrant DNA hypermethylation pattern and tumor type or its classification (Hughes et al., 2013). Yet, it is still unclear whether CIMP is a universal feature of all human cancers caused by the same mechanisms or it is an event specific to a certain tumor type. Some evidence supports the notion of tissue-specificity of CIMP. For example, the inactivation of mismatch repair gene MLH1 is related to CIMP in colon cancer, mutation in epigenetic regulators IDH1/2 are observed in glioblastomas and mutations of TET2 are associated with CIMP in leukemia (Witte et al., 2014). On the other hand, emerging number of reports indicates that CIMP is a common feature affecting determined groups of genes in many cancers. It was shown, for instance, that targets of polycomb repressive complex (PRC) are associated with CIMP in multiple tumor types (Kalari & Pfeifer, 2010).

Changes in DNA methylation are frequently associated with modifications within chromatin state. CpG-island hypermethylation in cancer cells correlates with loss of histone active marks: histone H3 and H4 acetylation, H3K4 trimethylation and gain of repressive marks: H3K9me3 and H3K27me3 (Esteller, 2008). What is more, the expression of many chromatin modifiers is often deregulated in tumor (Dawson & Kouzarides, 2012). Yet, the exact mechanisms involved in the regulation of epigenetic landscape in cancer cells remain poorly characterized.

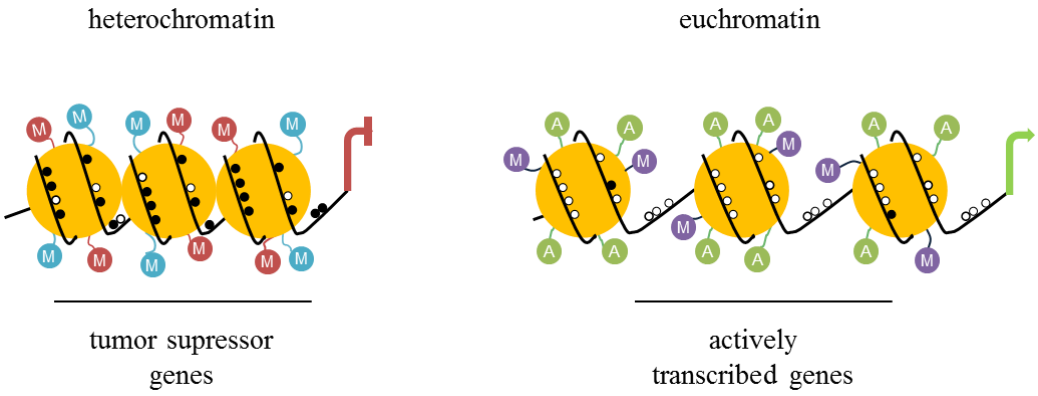

Figure 1.1. Gene promoters of tumor suppressor genes are hypermethylated (black circle) and marked by histone repressive markers like H3K9me3 (red circle) and H2K27me3 (blue circle). Global loss of DNA methylation (white circle) is correlated with open chromatin state marked by H3K4me3 (purple circle) and histone acetylation (green circle).

KRAB-ZNFs are the largest family of transcription repressors found only in tetrapod vertebrates. Human genome encodes transcripts for over 700 different proteins. A typical KRAB-ZNF contains from 4 to 30 zinc finger motifs, which permits highly specific binding to their target sequences (Lupo et al., 2013). Upon binding to DNA, KRAB-ZNFs trigger transcriptional repression through interaction with KAP1 protein that acts as a scaffold for other chromatin-remodeling factors. The repressive complex (Figure 1.2) contains histone methyltransferase SETDB1, histone deacetylase-containing complex NuRD and heterochromatin protein 1 (HP1). NuRD complex removes acetyl residues from the histones, SETDB1 deposits trimethylation mark on H3K9, whereas HP1 protein, upon recognition of H3K9me3 modification, promotes heterochromatization of affected locus (Groner et al., 2010). Moreover, we and others have shown that in a stem cell context, KRAB-ZNFs may mediate DNA methylation of the sequences adjacent to their binding motifs (Quenville et al., 2011, Gładych et al., submitted manuscript).

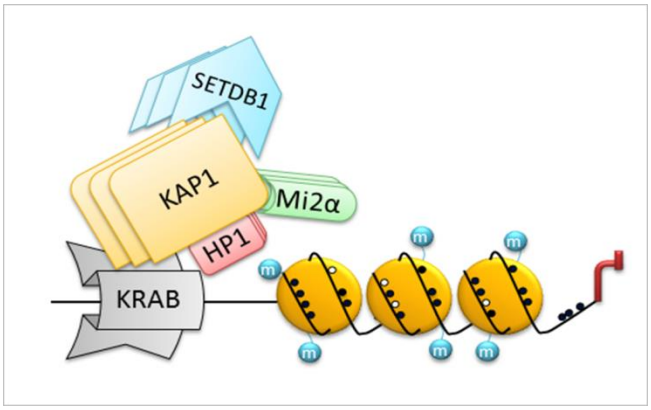

Figure 1.2. KRAB-ZFPs binds to a specific DNA sequence via zinc finger motifs. Upon binding, KRAB domain interacts with KAP1 protein. KAP1 forms a complex with chromatin-modifying enzymes and triggers formation of heterochromatin resulting in the repression of transcription.

KRAB-ZNFs expression is largely tissue-specific and their biological functions may be modulated by the cellular context and cooperation with other cofactors. Many reports indicate that KRAB-ZNFs may participate in developmental processes, apoptosis and carcinogenesis. The majority of the reports on the involvement of KRAB-ZNFs in carcinogenesis demonstrates their tumor suppressor activities, whereas only few studies focused on their tumor promotion properties. One of the reported tumor-specific KRAB-ZNFs is ZNF224, which shows both oncogene and tumor suppressor features. Harada and colleagues demonstrated that siRNA-mediated silencing of ZNF224 suppressed proliferation of bladder cancer cell line. Loss of ZNF224 disrupted its interaction with cancer-testis antigen DEPDC1, thus activating DEPDC1 downstream target genes, e.g. A20. Restoration of A20 expression promoted apoptosis through inactivation of the NF-κB pathway (Harada et al., 2010). Beside the repressive function, ZNF224 was also shown to act as a co-activator of WT1-mediated transcription (Florio et al., 2010). In human erythroleukemia cell line K562, ZNF224 cooperates with WT1 to regulate the transcription of WT1 target genes without direct binding to

DNA. This interaction leads to increased expression of pro-apoptotic genes and suppression of anti-apoptotic factors (Florio et al. 2010).

The expression of another two KRAB-ZNFs, ZNF382 and ZNF535, was shown to be ubiquitous in normal adult tissues, but reduced in many cancer cell lines due to promoter CpG methylation. Studies performed in colorectal carcinoma cell line HCT116 indicated that restoration of ZNF382 and ZNF545 expression reduced tumor cells proliferation via attenuation of NF-κB signaling pathway and subsequent induction of apoptosis (Cheng et al., 2010, Cheng et al., 2012). Tumor suppressor function was also described in the case of ZBRK1 (Furuta et al., 2005, Lin et al., 2010). ZBRK1 specifically binds to ANG1 (Angiopoietin 1) promoter and, thanks to the interaction with BRCA1 and CtIP, suppresses expression of this angiogenesis-promoting oncogene (Furuta et al., 2006). In addition, ZBRK1 expression is significantly reduced in cervical cancer, while its ectopic overexpression inhibits cancer cell growth *in vitro* and *in vivo*. Microarray analyses identified of a subset of genes negatively regulated by ZBRK1. This group included the genes responsible for cell proliferation, motility and metastasis (e.g. MMP9) (Lin et al., 2010). Additional two members of KRAB-ZNFs family were described as p53 regulators (Li et al., 2007, Tian et al., 2009, Yuan et al., 2012). One of them, ZNF307, was shown to be involved in the regulation of apoptosis by upregulation of MDM2 and EP300, which led to ubiquitination and degradation of p53 (Li et al., 2007). The second, APAK (ATM and p53-associated KZNF protein), inhibited p53-mediated apoptosis. Upon binding to p53, APAK recruited KAP1 and HDAC proteins, which impaired p53 activity via p53 deacetylation (Tian et al., 2009). Further analyses enabled the identification of a consensus binding site for APAK, which appeared to overlap with p53 binding sites. Thus, it has been proposed that APAK inhibits apoptosis through direct competition with p53 over binding to p53 target genes (Yuan et al., 2012).

While many of the above cited reports show KRAB-ZNFs involvement in carcinogenesis, little is known about the exact molecular mechanisms responsible for observed phenotypes. Majority of the studies focused on one KRAB-ZNF linked to one cancer type. Our preliminary results of TCGA pan-cancer transcriptomic profiling (point 3.4.1) have demonstrated a small, but distinct subset of KRAB-ZNFs overexpressed in multiple cancer types. We hypothesize that the observed common upregulation pattern may indicate that these cancer-associated KRAB-ZNFs act through similar molecular mechanisms. Moreover, it is tempting to speculate that some of them may play a role in the epigenetic repression of tumor suppressor genes. Nevertheless, this hypotheses need to be experimentally verified.

2.2 Justification for tackling proposed research questions

Carcinogenesis involves multiple modifications, which stem from genetic and epigenetic causes and lead to abnormal phenotype of affected cells manifesting with sustained proliferation, disrupted apoptosis, deregulated metabolism and other features (Hanahan & Weinberg, 2011). While genetic causes of cancer were the first to be thoroughly studied, epigenetic players and alterations are relatively new incomers in the field. The effect of many epigenetic factors on carcinogenesis has been already defined (Esteller, 2008). Nevertheless, many epigenetic events are still poorly characterized. For example, the mechanisms responsible for the methylation of the promoters of some of the tumor suppressor genes remain unknown.

The canonical role of KRAB-ZNFs is linked to two key epigenetic events: deposition of inactivating H3K9me3 mark and DNA methylation, both of which may lead to the heterochromatization and thus, inactivation of affected loci. These modifications are an effect of cooperation of a protein complex, in which a KRAB-ZNF is responsible for the recognition and binding to specific DNA sequences. Therefore, it may be hypothesized that some of the KRAB-ZNFs may be involved in the epigenetic repression of certain tumor suppressor genes, either via DNA methylation and/or H3K9 trimethylation.

Our team has intensively studied the role of KRAB/KAP1 complex in the reprogramming of human somatic cells to induced pluripotent stem cells (iPSCs). We found that KRAB-ZNFs and KAP1 protein play an important role during epigenetic reprogramming of somatic cells to induced pluripotent stem cells (Kulcenty et al., submitted manuscript, Gładych et al,. submitted manuscript). We also found that expression of several KRAB-ZNFs is essential in maintaining self-renewal capacity of human embryonic stem cells (Oleksiewicz et al., manuscript in preparation). It is believed that tumor initiating cells may share some common molecular characteristics with pluripotent stem cells (Kim & Orkin, 2011). Thus, it will be of interest to test whether certain KRAB-ZNFs have similar function in pluripotent and cancer stem cells.

Several tools have been developed to attempt prediction of binding motifs of KRAB-ZNFs. Yet, the main drawback to these tools is the fact that only few zinc fingers of a given KRAB-ZNF participate in DNA binding (Najafabadi et al., 2015; Persikov & Singh, 2014). This phenomenon renders any computational

modeling inaccurate. So far, only few KRAB-ZNF factors have been described in the literature in the context of human tumors (Lupo et al., 2013). The majority of the published work explored an influence of a chosen KRAB-ZNF on the phenotype and transcriptional profile of cancer cells, rather than KRAB-ZNF target genes. The current project will advance beyond previously published efforts. Using high-throughput method of chromatin immunoprecipitation sequencing (ChIP-seq) we aim to map the consensus binding sites of cancer-associated KRAB-ZFPs. By integrating molecular biology assays, global epigenomic profiling and bioinformatic analyses, this interdisciplinary project will shed a light on the KRAB-ZNFs-mediated epigenetic events and their phenotypic consequences occurring in cancer.

2.3 Pioneering nature of the project

The studies conducted within the scope of this project will combine a broad spectrum of interdisciplinary approaches that will be used for the systematic characterization of KRAB-ZNFs involvement in modulating epigenetic profile of cancer cells. To our knowledge, this will be the first study that combines molecular biology, genetic engineering, cancer cell biology, epigenetics, bioinformatics and biostatistics to explore the molecular and phenotypic function of cancer-associated KRAB-ZNFs. The proposed workflow of experiments is original and beyond the state-of-the-art. The input data for the preliminary statistical analyses were obtained from enormous TCGA datasets, which still remain largely unexplored by the TCGA non-associated scientific community. The data obtained *in silico* will be further validated by various molecular biology techniques *in vitro*, thus providing cross-validation of the generated results. What is more, consensus binding sites of selected KRAB-ZNFs will be determined by the usage of next-generation sequencing technique, namely ChIP-seq (chromatin immunoprecipitation – sequencing). Based on the results, we will be able to determine specific targets of selected KRAB-ZNFs and investigate their expression/methylation status in cancer cells. We will also correlate the expression of KRAB-ZNFs with the clinical outcome of TCGA patients and molecular profile of their tumor samples, which will further enable assessment of KRAB-ZNFs functioning in cancer development.

2.4. The expected impacts of the project

The results of this project will help explore the molecular and phenotypic role of cancer-associated KRAB-ZNFs. This will allow better understanding of the mechanisms implicated in the functioning of this large family of transcriptional repressors. We hope that the experimental and analytical strategies proposed in this project may serve as an exemplary template workflow for other scientists researching the roles of other transcription factors. What is more, the research focusing on the aberrant DNA methylation of cancer genome is currently one of the main areas of interest for the pharmaceutical companies developing epigenetic anti-cancer agents. Small molecule inhibitors of DNA methyltransferases (e.g. azacytidine, decitabine) are already used therapeutically. Preclinical studies demonstrated that administration of these molecules decreases global methylation levels and restores normal expression of tumor suppressor genes (Luebbert, 2000). Promising results in anti-cancer therapy were also observed in the case of HDAC inhibitors (Minucci & Pelicci, 2006). Nevertheless, the effect of the administration of these agents is unstable, because of the increased DNA methylation and histone acetylation observed after drug withdrawal. Secondly, and more importantly, these inhibitors are non-specific and may induce off-target side-effects (Luebbert, 2000). Determining specific proteins that are responsible for targeted epigenetic inactivation of tumor suppressor genes will greatly contribute to the development of more efficient anti-cancer therapies. New, more specific drugs are expected to have a significant role in reduction of cancer burden in our society. Such an approach is particularly important in the era of targeted, individualized treatment strategies.

3. Concept and work plan

3.1 Concept

In the current project we aim to test the hypothesis that cancer-associated KRAB-ZNFs may play a role in modulating tumor epigenetic landscape. We are specifically interested in addressing the question whether these proteins induce epigenetic repression of tumor suppressor genes. We assume that by negative regulation of tumor suppressor genes, cancer-associated KRAB-ZNFs promote progression of cancerous features. In order to test these hypotheses we plan to integrate various approaches into a multidisciplinary project. The study will involve selection of a number of KRAB-ZNFs that are overexpressed in many cancer types compared to normal cells. This will be followed by in-depth analysis of the effect that KRAB-ZNFs exert on their target genes, and thus, various signaling pathways. The bioinformatic high-throughput transcriptomic and epigenenomic profiling will be tightly interwoven with basic molecular and cell biology experiments to provide two-way cross-validation of the results obtained *in silico* and *in vitro*. Such

a complementary approach will help answer the question whether cancer-associated KRAB-ZNFs may have an influence on carcinogenesis by affecting expression of tumor suppressor genes through modulation of their epigenetic profile.

3.2 Overall and detailed work plan

The main objective of the project is to analyse the role of transcriptional repressors KRAB-ZNFs in the modulation of epigenetic profile in cancer cells. An initial list of the KRAB-ZNFs overexpressed in various tumor tissues has been already chosen based on the analysis of the data available through TCGA pan-cancer project. The preliminary results of the transcriptome analysis will be validated in several panels of cancer tissues and cell lines to ascertain appropriate selection of cancer-associated KRAB-ZNFs and cellular models for further analyses. Downstream experiments will explore two aspects. First, a set of epigenetic assays will help identify KRAB-ZNFs target genes and define the mode of their epigenetic regulation. Second, silencing of the selected KRAB-ZNFs and subsequent phenotypic assays will answer the question whether selected KRAB-ZNFs affect cancer cell behavior. To test whether obtained results correspond to *in vivo* outcome, we will cross-validate our observations with the on-line datasets harboring cancer patho-clinical profiles and molecular signatures. Finally, we plan to perform bioinformatic analyses with the data available from the ENCODE database to further explore epigenetic mechanisms exerted by the selected KRAB-ZNFs. The detailed work plan includes four tasks outlined in the table below (Table 1).

Table 1. The scheme of detailed work plan associated with the project.

Detailed work plan	
Task 1	
Validation of the bioinformatic analyses	**Goal:** to confirm KRAB-ZNFs overexpression in cancer samples and select appropriate cell line models for further downstream experiments **Materials:** mRNA and protein extracts from normal and cancer cell lines and tissues obtained from two most common cancers: breast and lung cancers **Methods:** RT-qPCR, Western blot
Task 2	
Identification and validation of KRAB-ZNF binding sites	**Goal:** to discover and characterize genome-wide distribution of KRAB-ZNF deposition sites **Materials:** four cancer cell lines with overexpression of tagged KRAB-ZNFs **Methods:** preparation of vectors and cell lines for KRAB-ZNF overexpression, chromatin immunoprecipitation followed by global sequencing (ChIP-seq) and qPCR (ChIP-qPCR), assessment of epigenetic status (DNA methylation, H3K9me3 enrichment) and transcriptional activity of target genes after KRAB-ZNF knockdown
Task 3	
The influence of KRAB-ZNF on cancer cell phenotype	**Goal:** to assess the potential of oncogenic properties of selected KRAB-ZNFs **Materials:** selected cancer cell lines with silenced KRAB-ZNF expression **Methods:** siRNA-mediated silencing followed by phenotypic assays testing cell proliferation, apoptosis, migration and invasion
Task 4	
Bioinformatic cross-validation of the epigenetic and phenotypic studies	**Goal:** to verify whether the epigenetic and phenotypic effects of KRAB-ZNFs observed *in vitro* is relevant also in the clinical setting **Materials:** datasets available from TCGA (clinical, gene expression and DNA methylation data) and ENCODE (DNA methylation, histone modification for selected cells lines) **Methods:** statistical methods (e.g. Cox model and Kaplan-Meier curves, correlation between DNA methylation/histone modifications with KRAB-ZNFs binding sites)

3.3 Preliminary results

3.3.1 Selection of cancer-associated KRAB-ZNFs

KRAB-ZNFs comprise the largest family of transcriptional repressors in the human genome. In many cases, expression of KRAB-ZNFs is largely tissue dependant. Some of the KRAB-ZNFs have been shown to play a role in metabolism, development, apoptosis and cancer (Lupo et al., 2013), data in point 3.4.2). However, these analyses were usually oriented towards one cell type and one KRAB-ZNF. Here, we propose to focus on a number of KRAB-ZNFs that are specifically overexpressed in multiple tumors when compared to normal tissues. In order to select cancer-associated KRAB-ZNF, we took advantage of the RNA-seq data available from TCGA.

TCGA is the largest project aiming at the comprehensive molecular profiling of human tumors. The scope of the TCGA efforts encompasses large cohorts of multiple cancer types (~30) and subtypes. By using a wide range of high-throughput technologies, TCGA provided enormous datasets on exom and whole-genome sequences, copy number variation, mRNA, miRNA and protein expression, splicing variants and DNA methylation. Moreover, each data point is accompanied by the clinical data, such as histology type, survival or TNM classification (Weinstein et al., 2013). These datasets are publically available through various on-line tools, e.g. The USCS Cancer Genomics Browser (Goldman et al., 2013).

The pan-cancer RNA-seq dataset downloaded from The USCS Cancer Genomic Browser included 22 tumors. We removed the data for 10 cancer types due to the low number (<20) of the corresponding normal tissues to ensure sufficient abundance of the samples in each set. Thus, the data available for the analysis comprised 5337 tumor and 620 normal tissues. The expression data were provided as log2 normalized RPKM values (RPKM: *Reads Per Kilobase per Million*). From the list of 20350 genes, we extracted the log2(RPKM) values only for the factors included in the most recently updated list of KRAB-ZNFs (Corsinotti et al., 2013). Next, we calculated the change in KRAB-ZNFs expression as follows: $\text{mean}[\log2(\text{RPKM})]_{\text{tumor}} - \text{mean}[\log2(\text{RPLM})]_{\text{normal}}$. The differential expression analysis indicated that only a fraction of KRAB-ZNF genes became deregulated in tumors. Interestingly, the majority of the KRAB-ZNFs with altered mRNA level exhibited reduced expression, while only a small, but distinct cluster of KRAB-ZNFs showed upregulation in cancer tissues (Figure 3.1).

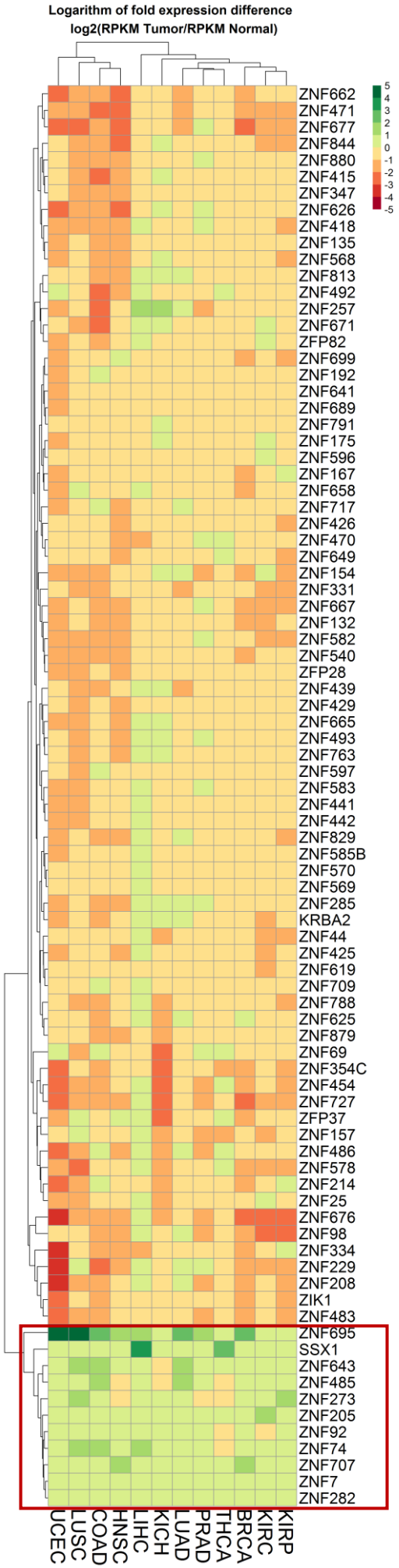

Figure 3.1 Heatmap and supervised clustering of KRAB-ZNFs with changed expression in 12 tumors compared to normal tissues [log2(fold change) > 0.4 & < -0.4]. Abbreviations: UCEC (endometriod cancer), LUSC (lung squamous cell carcinoma), COAD (colon cancer), HNSC (head and neck cancer), LIHC (liver cancer), KICH (kidney chromophobe), LUAD (lung adenocarcinoma), PRAD (prostate cancer), THCA (thyroid cancer), BRCA (breast cancer), KIRC (kidney clear cell carcinoma), KIRP (kindey papillary cell carcinoma).

The threshold value indicating gene overexpression was set at log2(fold change) > 0.4. The distinctive subgroup of KRAB-ZNFs with the highest expression in multiple tumors as pinpointed by supervised clustering analysis (Figure 3.1) was dubbed as cancer-associated KRAB-ZNFs. It includes following genes: SSX1, ZNF7, ZNF74, ZNF92, ZNF205, ZNF273, ZNF282, ZN485, ZNF643, ZNF695 and ZNF707. This pilot selection will be further validated at the mRNA and protein level using cancer and normal cell lines and tissues in Task 1.

In order to explore the biological role of cancer-associated KRAB-ZNFs, we plan to test their function in model cancer cell lines. Thus, we narrowed down further analysis only to the two most common cancer types: lung cancer (adenocarcinoma and squamous cell cancer) and breast cancer (basal and luminal). According to the global cancer statistics for 2012th year, breast cancer has the highest incidence (~1.68 mln new cases) and mortality rate (~0.52 mln cases) among all female cancer patients. Lung cancer is the first cause of cancer-related diagnosis (~1.24 mln) and death (~1.1 mln) in male population. Moreover, it is the third most frequent newly diagnosed cancer in women (~0.61 mln) and the second cancer-related reason of death (~0.49 mln) (Torre et al., 2015). Thus, given the high burden of these tumors in worldwide population, we consider breast and lung cancers as appropriate disease models for downstream experimentation. The expression profiles of cancer-associated KRAB-ZNFs in the selected tumor types and their normal counterparts is depicted in Figure 3.2.

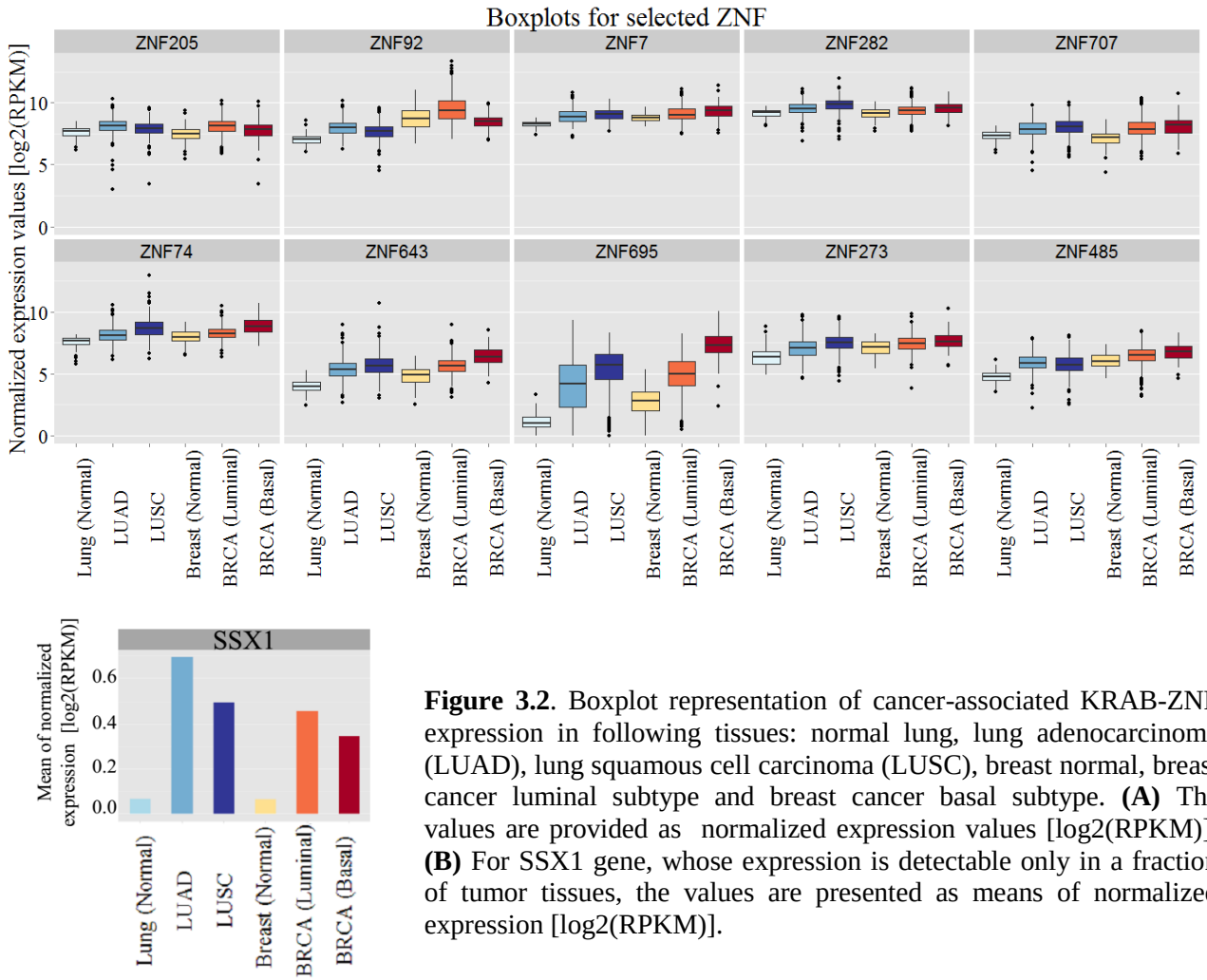

Figure 3.2. Boxplot representation of cancer-associated KRAB-ZNF expression in following tissues: normal lung, lung adenocarcinoma (LUAD), lung squamous cell carcinoma (LUSC), breast normal, breast cancer luminal subtype and breast cancer basal subtype. **(A)** The values are provided as normalized expression values [log2(RPKM)]. **(B)** For SSX1 gene, whose expression is detectable only in a fraction of tumor tissues, the values are presented as means of normalized expression [log2(RPKM)].

3.3.2 Cancer-associated ZNF695 affects stemness in pluripotent stem cells

Sustained tumor formation, progression and therapy resistance is associated with the small population of so called cancer stem cells (CSCs). These cells, known also as tumor initiating cells, may originate from somatic progenitor stem cells or dedifferentiated somatic cells. CSCs share many phenotypic and molecular characteristics with pluripotent stem cells derived from a blastocyst (ESCs, *embryonic stem cells*) and

through *in vitro* induction (iPSCs, *induced pluripotent stem cells*). For example, alike ESCs and iPSCs, tumor initiating cells can self-renew, which means they can propagate unlimitedly while maintaining an undifferentiated state. Induction of pluripotency in somatic cells with forced expression of four pluripotency factors (OCT4, SOX2, KLF4 and c-MYC) is, to a certain degree, reminiscent of the dedifferentiation suggested for some tumors. Multiple studies have shown molecular similarities in gene expression profile and epigenetic regulation between pluripotent and cancer stem cells (Kim & Orkin, 2011). Therefore, it is highly likely that certain KRAB-ZNFs overexpressed in cancers may participate in the regulation of the stemness network. Our data indicate that this indeed might be the case.

In our current project we are focusing on the epigenetic regulation of stemness mediated by KRAB-ZNFs and their associated protein KAP1/TRIM28. In order to identify KRAB-ZNFs that may be important for stem cell biology we have performed RNA-seq transcriptomic profiling using human fibroblasts (N=4) and originated from them iPSCs (N=8) generated and verified in our lab (Wroblewska et al., in preparation). Differential expression analysis allowed identification of a set of KRAB-ZNFs with high expression in iPSCs compared to their parental fibroblasts. The same repressors were found upregulated also by other research groups, as we confirmed by manual mining of the data available through the NCBI and the Gene Expression Atlas (EMBL-EBI). KRAB-ZNFs upregulated at the significance level of corrected p-value < 0.05 were selected for validation in the same panel of fibroblasts and iPSCs with RT-qPCR analysis, while human ESCs served as a control. The RT-qPCR expression data of the majority of the tested KRAB-ZNFs was concordant with the RNA-seq analysis. Interestingly, one of the pluripotency-specific KRAB-ZNFs (ZNF695, Figure 3.3) is also highly upregulated in various cancers when compared to normal tissues as observed in our pan-cancer analysis (Figure 3.1 and 3.2).

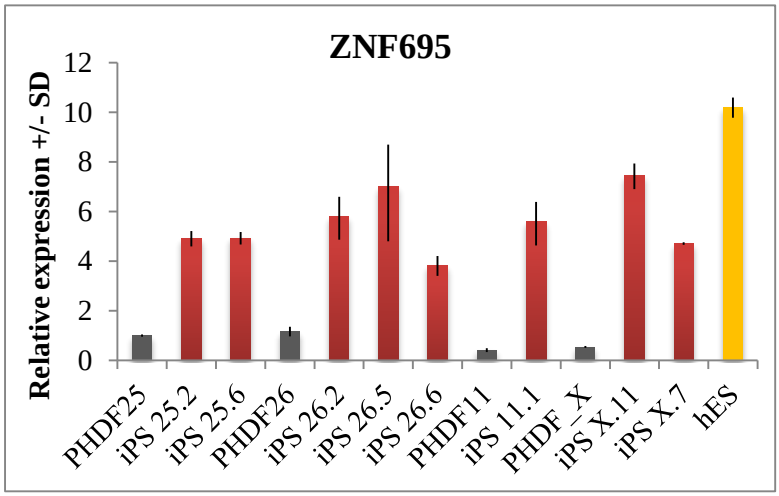

Figure 3.3. ZNF695 is overexpressed in human induced pluripotent stem cells (iPSCs, red bars) and human embryonic stem cell lines (hES, orange bars) in comparison to primary human dermal fibroblasts (PHDF, grey bars) as assessed with RT-qPCR assay. The data are provided as a mean of a triplicate experiment ± standard deviation.

Next, we wanted to test whether KRAB-ZNFs overexpressed in pluripotent stem cells may affect self-renewal and pluripotency. To this end, we used siRNA silencing method to knock-down expression of each of the selected KRAB-ZNFs. The cells were transfected every 3 days with 100nM siRNA and the outcome of the silencing was analyzed on 6th day. Here, we show the observations only for ZNF695 that were overexpressed also in human tumors. Undifferentiated hESCs are small, compacted cells growing in tight colonies that maintain well defined boundaries. Yet, the ZNF695 knock-down cells lost typical hESC morphology as the cells became bigger, acquired spindle shape and detached from the colonies (Figure 3.4a) Such features suggest the onset of differentiation. Moreover, the RT-qPCR analysis demonstrated that loss of ZNF695 led to lowered expression of both tested pluripotency markers: OCT4 and NANOG (Figure 3.4b). These results altogether indicate the involvement of ZNF695 in stemness maintenance in pluripotent stem cells. Based on these observations, it may be hypothesized that ZNF695 may participate in protection of stemness also in tumor initiating cells.

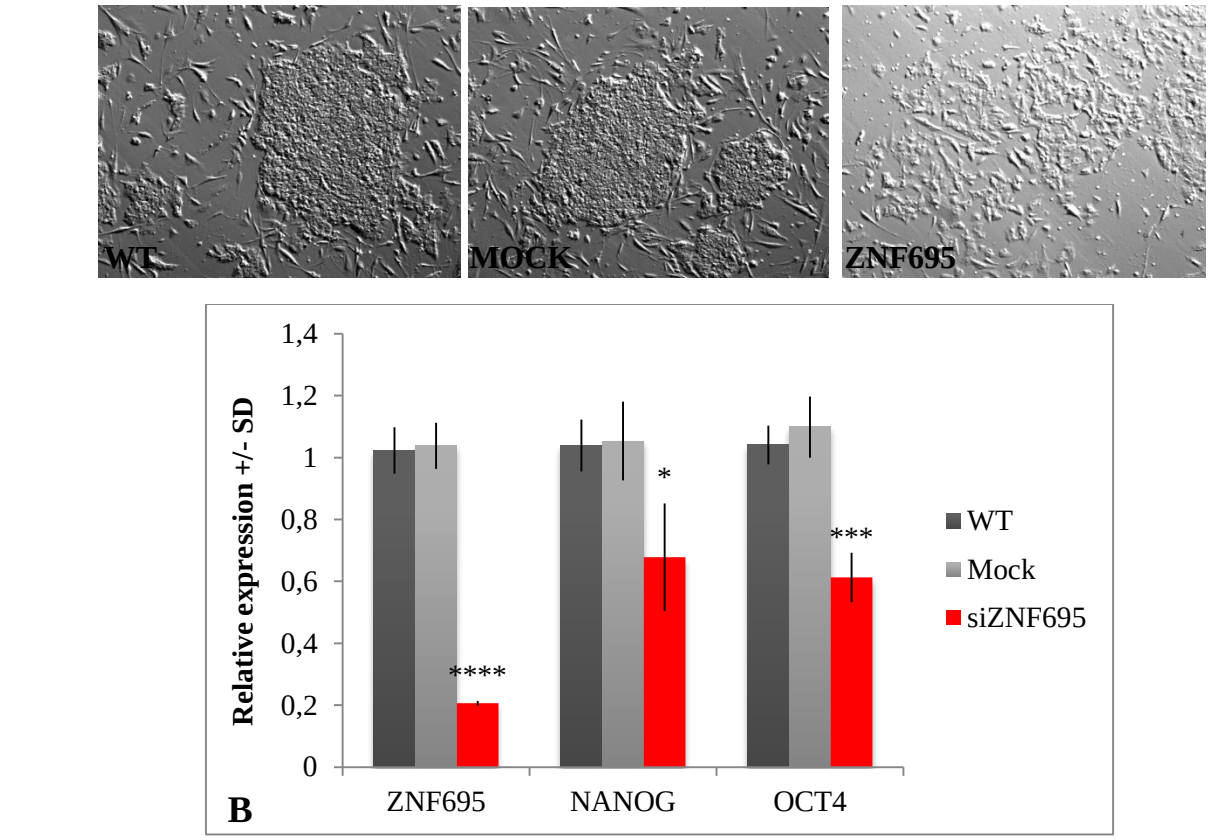

Figure 3.4. ZNF695 silencing results in reduction of pluripotency markers. **(A)** The hES colonies lost their typical morphology upon ZNF695 silencing, **(B)** while mRNA expression of pluripotency markers (NANOG and OCT4) diminished as observed in RT-qPCR assays. The data are presented as relative expression values ± standard deviation calculated from three independent biological replicates.

3.3.3 Feasibility of chromatin immunoprecipitation assay based on *in vivo* biotynylated KRAB-ZNF factors

Chromatin immunoprecipitation followed by global sequencing (ChIP-seq) is a powerful technique that allows identification of the DNA sequences bound by transcription factors, cofactors and histones. ChIP consists of multiple steps: cross-linking proteins to DNA, cell lysis, fragmentation of DNA, immunoprecipitation of the protein of interest, extensive washing of unbound material and purification of enriched DNA. While most of the steps are relatively easy to optimize to reach the quality suggested by the current guidelines (Landt et al., 2012), the biggest challenge might be associated with inadequate specificity and sensitivity of the immunoprecipitation step.

KRAB-ZNF genes are the biggest family of mammalian repressors that rapidly evolved through multiple duplication events followed by series of point and indel mutations (Lupo et al., 2013). In order to ensure high specificity of ChIP experiments, it is crucial to avoid potential cross-reactivity against multiple KRAB-ZNFs paralogues that may be co-expressed in the same cell. By modifying a previously published protocol (Kim et al., 2009), we developed an assay utilizing KRAB-ZNFs tagged with a biotin residue. The system requires two vectors: one harboring a gene of interest fused with the sequence coding for the short biotinylation peptide (Figure 3.5.a) and the other with biotin synthetase (BirA) (Figure 3.5.b). Co-expression of both genes results in biotinylation of tagged protein *in vivo*. Biotin tags have several advantages over other commonly used tags. Thanks to the high affinity of biotin to streptavidin, purification of biotinylated proteins is very efficient and specific, mainly because it is possible to apply more stringent washes during procedure. Besides, the number of endogenous biotinylated proteins is small and their localization is mainly cytoplasmic, so the background from endogenous proteins is negligible (Baubec et al., 2015; Kim et al., 2009).

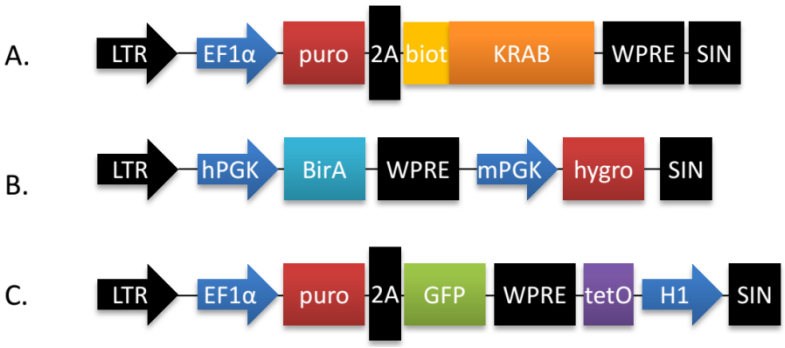

Figure 3.5. Schematic representation of vector maps used in ChIP assays: **(A)** pLVE-PB with either KRAB-ZNF-biot or TRKRAB-biot sequence, **(B)** pLenti-BirA with biotin synthetase gene, **(C)** pLVETHM-GP with *gfp* gene and 7 tandem tetO sequences.

For the optimization of the streptavidin ChIP assay, we biotin-tagged a chimeric tetracycline repressor fused to a KRAB domain (TRKRAB-biot). The cells (HEK293T) were transduced with three vectors. The first vector (Lenti-BirA) encoded biotin synthetase, the second (LVE-PB) – TRKRAB-biot sequence, while the third (LVETHM-GP) contained gfp (*green fluorescence protein*) gene and tetracycline operator (tetO) sequence (Figure 3.5). The cells were cultured in the presence or absence of doxycycline (±dox). Lack of dox resulted in the binding of TRKRAB-biot to tetO sequence on LVETHM-GP. KRAB domain recruits KAP1/TRIM28 protein complex, thus mediating heterochromatization of tetO fragment and its neighboring loci. Thus, downregulation of gfp marked by the quenching of green fluorescence indicated TRKRAB-biot binding to tetO. The cells cultured with or without dox were subjected to cross-linking and streptavidin ChIP optimization procedure. Next, purified DNA was used in the RT-qPCR assays utilizing primer pairs specific for sequences flanking tetO region and for the negative control regions (Figure 3.6). The results indicated high specificity of the optimized procedure. TRKRAB-biot isolated from dox-untreated cells was highly enriched on tetO flanking sequences, but not on the negative control regions.

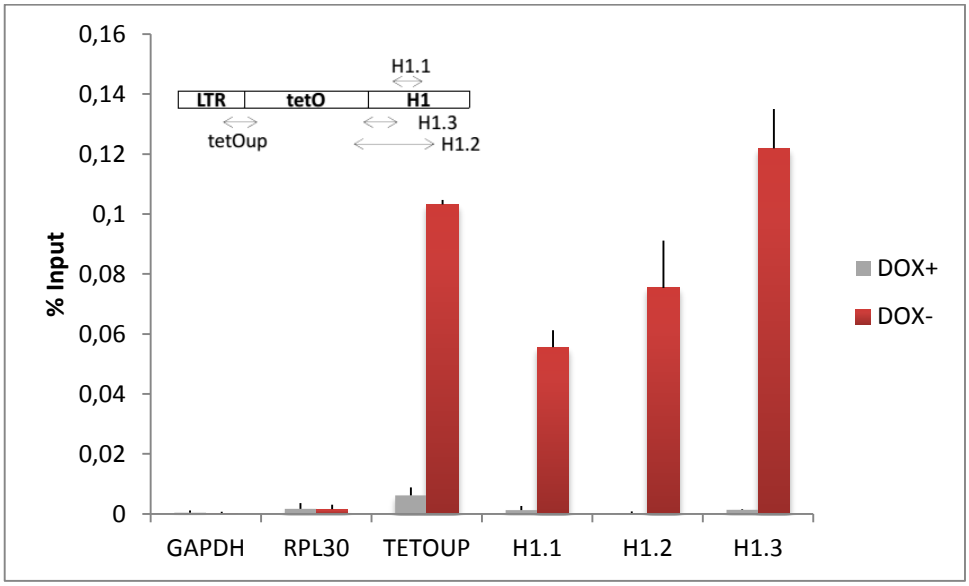

Figure 3.6. ChIP-qPCR analysis of DNA fragments bound by TRKRAB-biot. TRKRAB-biot binds to tetO and tetO neighboring sequences in HEK293T cultured without doxycycline (DOX-, red bars), but not in the cells cultured in the presence of the antibiotic (DOX+, grey bars). The upper left corner maps fragments amplified in the qPCR assays.

As a proof-of-concept study, we biotin-tagged ZFP57 protein – a KRAB-ZNF known to bind to the imprinting control regions (ICRs). Again, we observed enrichment of ZFP57-biot on its target regions (Figure 3.7). These data ensure that our streptavidin ChIP protocol might be successfully applied to the experiments aiming at the identification of KRAB-ZNF binding sites.

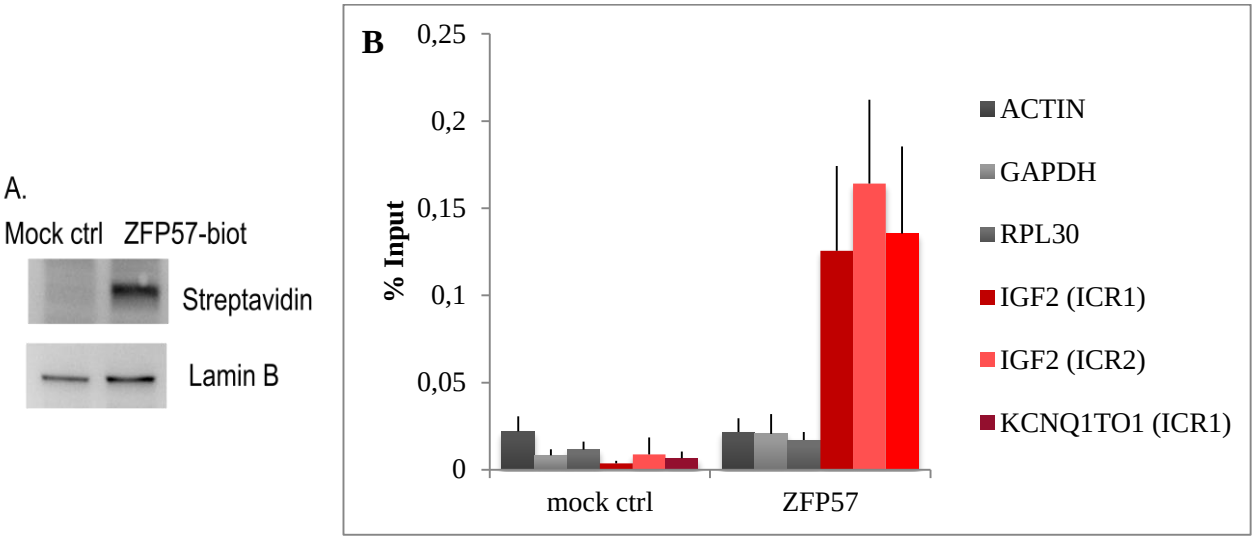

Figure 3.7 (A) Western blot analysis of *in vivo* biotinylated ZFP57 overexpressed in HEK293T cell line. Lamin B served as a loading control. **(B)** ChIP-qPCR analysis of DNA bound by ZFP57-biot. Grey bars show the enrichment over the negative control regions, while red bars present the data for Imprinting Control Regions (ICR) known to be bound by ZFP57.

4. Methodology

This research will be conducted as a collaborative, interdisciplinary project combining wide range of approaches from various fields, such as molecular biology, genetic engineering, cancer cell biology, epigenetics, bioinformatics and biostatistics. Our previous extensive experience in most of the proposed techniques renders the project highly feasible. The majority of the wet lab work will be performed in the Chair of Medical Biotechnology at the Poznan University of Medical Sciences. DNA methylation assays will be performed in collaboration with the Biomarker Group at the University of Liverpool, UK. The bioinformatic and biostatistic tasks will be performed by the experienced team of statisticians from the Interdisciplinary Centre for Mathematical and Computational Modeling in Warsaw (ICM). The co-investigators from ICM are already working with our laboratory on various projects. These projects are focused on the statistical analysis of available TCGA data, as well as on the bioinformatic analyses of data produced through next-generation sequencing that have been performed within a scope of the currently running project.

4.1. Task 1 - validation of the bioinformatic analysis

High-throughput analyses, whilst being highly efficient and informative, may be susceptible to certain intrinsic errors. These usually minor mistakes may appear for example due to the inaccurate alignment of the sequenced reads to the reference genome. The alignment errors may result from nucleotide polymorphisms, permitted mismatches and/or high homology of the analyzed sequences. This latter possibility may be especially true in the case of KRAB-ZNF family genes. Thus, in order to ascertain appropriate selection of genes identified through high-throughput analyses for *in vitro* experiments, it is crucial to validate the results using a single-gene approach. We have tested such an approach while performing our current project. In this project we measured the level of KRAB-ZNFs expression using RNA-seq data to compare induced pluripotent stem cells (N=8) with their parental fibroblasts (N=4). The bioinformatic analysis allowed selection of 16 KRAB-ZNFs, whose expression significantly increased in pluripotent stem cells. The RT-qPCR validation analysis confirmed a similar trend in the case of 11 tested KRAB-ZNFs, while two KRAB-ZNFs were overexpressed only in a fraction of the iPS cell lines compared to fibroblasts. The remaining two KRAB-ZNFs did not validate at all and one could not be validated, because its transcript fell below the detection level of RT-qPCR method. These observations emphasize the necessity to validate the results obtained from high-throughput analyses.

The candidate cancer-associated KRAB-ZNFs identified through analysis of TCGA pan-cancer data will be validated in a narrowed set of human tumors: breast cancer (basal and luminal) and lung cancer (adenocarcinoma and squamous cell carcinoma). To this end, we will use several panels of normal and cancer cells lines and tissues. Lung normal and cancer tissue samples (~48 pairs) and cell lines (~4 normal bronchial epithelial and 10 lung cancer) are available through the collaboration with the Biomarker Group (University of Liverpool, UK). Breast normal and cancer cell lines (N=9) are available within the Medical

Biotechnology lab. The paired normal and cancer tissue cDNA samples from 48 breast cancer patients will be purchased from a commercial vendor.

The expression of candidate cancer-associated KRAB-ZNF in tumor and paired normal tissues will be tested with RT-qPCR assays. The normal and cancer cell lines will be propagated, pelleted and banked for future analyses. One part of the pellets will be used for RNA extraction (using TRI reagent), reverse transcription and RT-qPCR assays. The primers for RT-qPCR analysis will be designed manually and checked for specificity using the Primer-BLAST application from the NCBI website. The second part of the pellets will be used for protein extraction, quantification and Western blot analysis using specific antibodies. The third part will be subjected to two-step cellular fractionation into cytoplasmic and nuclear compartments to assess intracellular localization of the selected KRAB-ZNFs. All analyses will be run in triplicate. For downstream analyses we will choose 5-6 KRAB-ZNFs that show nuclear localization, as well as the highest and the most frequent overexpression in cancer tissues and cell lines compared to the non-malignant samples. The screening will also allow selection of a panel of 4 cancer cell lines (breast basal and luminal, lung adenocarcinoma and squamous cell carcinoma) with high expression level of selected KRAB-ZNFs. These cell lines will be used as research model in downstream experimentations. The selected KRAB-ZNFs will be further subjected to multiple epigenetic and phenotypic assays that will address the question whether they may have an impact on tumor cell biology and gene regulation.

4.2. Task 2 - identification of KRAB-ZNFs binding sites

In the 2nd Task we would like to focus on the identification and characterization of the chromatin occupancy sites of the selected cancer-associated KRAB-ZNFs. In order to do so, we will biotin tag chosen KRAB-ZNF, overexpress them in the panel of 4 cancer cell lines and perform a ChIP assay. One part of the material will be send for sequencing and the other will serve as a validation set. Moreover, we will knock-down expression of each KRAB-ZNF to check whether their target genes change chromatin and expression status.

First of all, the selected KRAB-ZNFs will be cloned into pLVE-PB plasmid (Figure 3.4.A) that has been already constructed in our laboratory and successfully used for cloning and overexpression of the pluripotency-specific KRAB-ZNFs. The pLVE-PB vector contains EF1α promoter that drives constitutive expression with comparable strength in multiple cell lines (Qin et al., 2010). EF1α promoter regulates the expression of a bicistronic transcript that contains puromycin resistance gene fused to a biotinylation peptide sequence via 2A linking peptide. KRAB-ZNFs fragments will be amplified from cDNA with a proof-reading polymerase using specific primer pairs flanked either with a SpeI or NdeI restriction site at the 5' end. For the KRAB-ZNFs that contain SpeI and/or NdeI sites within their coding sequence, we will use alternative restriction sites (e.g. XbaI, AseI) that give compatible sticky ends. The KRAB-ZNF inserts will be ligated into NdeI/SpeI-digested pLVE-PB plasmid. Positive clones will be analyzed with enzymatic restriction and Sanger sequencing. The pLVE-PB-KRAB-ZNF and pLenti-BirA (Figure 3.4.a & b) will be used for production of lentiviral particles utilizing 2nd generation packaging system. For lentivirus production we routinely transfect HEK293T packaging cell line with the calcium phosphate method.

The cells will be transduced with Lenti-BirA vector and, after selection with hygromycin, transduced with LVE-PB-KRAB-ZNF vectors followed by selection with puromycin. After propagation, we will collect the cells to ensure overexpression and nuclear localization of the *in vivo* biotinylated KRAB factors utilizing Western blot method on fractionated cells. Next, we will cross-link the cells and perform streptavidin ChIP assay that has been already optimized in our lab (as described in point 3.4.3). As a positive control of the ChIP procedure, we will use the cells with overexpression of biotinylated ZFP57, whose target binding sites are already known. The cells transduced with an empty pLVE-PB vector will serve as a negative control. Chromatin fragmentation will be verified with the agarose gel electrophoresis, while DNA quantity will be analyzed with a dsDNA intercalating picogreen dye. Duplicate samples of input and immunoprecipitated chromatin from a selected cell line will be sequenced. According to ENCODE guidelines, duplicate samples are sufficient in ChIP-seq experiments (Landt et al., 2012). The processing of ChIP-seq data will be performed in collaboration with the team from ICM. The sequenced fragments will be aligned to the reference genome using Bowtie2 tool, and further analysed with chipseq and ChIPpeakAnno packages for R (Sakar et al., Zhu et al., 2010) For the verification of the sequencing results we will use three replicates of ChIP and input samples produced from all 4 cell lines. We will select 5-8 most interesting binding sites of each of the tested KRAB-ZNF to perform ChIP-qPCR validation assays.

Finally, we will explore the mechanisms of the KRAB-ZNF-mediated gene regulation. We will perform ChIP assays using H3K9me3 antibody according to our working protocol (Gladych et al, submitted manuscript) followed by the same qPCR assays as in the case of ChIP-seq validation. What is more, we will knock-down KRAB-ZNFs expression using RNAi technology in order to assess potential changes in DNA methylation of the binding sites and expression of the nearest genes. DNA isolated from the control and knock-down cells will be subjected to bisulfate conversion, PCR amplification and pyrosequencing. Pyrosequencing will be performed in collaboration with the Biomarker Group from the University of Liverpool in UK. The Biomarker Group has remarkable expertise in using the pyrosequencing method and designing successful pyrosequencing assays.

4.3. Task 3 - the influence of KRAB-ZNFs on cancer cell phenotype

In order to assess the cancer-associated KRAB-ZNF function we will perform a series of phenotypic assays measuring various parameters related to tumor biology. Functional tests will be performed upon silencing of KRAB-ZNFs expression using RNAi technology in all four selected cancer cell lines. First, we will design 4-6 siRNAs targeting the transcript of each selected KRAB-ZNFs. Their silencing efficiency and duration will be tested at the mRNA level with the RT-qPCR assays and at the protein level with the Western blot analysis. We will use two siRNAs with the highest silencing efficiency (>75% of repression) and one FITC-stained, non-specific siRNA serving as a negative control and for the monitoring of transfection efficiency. To ensure low protein expression during all functional assays, siRNA will be administered for at least 2 days prior to the measurements.

The cells with reduced expression of KRAB-ZNFs and control cells will be subjected to various functional *in vitro* assays measuring cell proliferation, apoptosis, migration and invasion. Proliferation assay will be performed using [H^3]-thymidine incorporation method and MTT test. Cell death will be analyzed via annexin V and propidium iodide staining (on flow cytometer) and via caspase 3 activation assay using ELISA-based kit. Migration will be measured using wound healing assay and Boyden chamber test, whereby the cells migrate through porous membrane into a chemoattractant (i.e. FBS). To the latter setup we will also add Matrigel-coated membranes in order to assess the invasive potential of siRNA-transfected cells. Cell proliferation and death assays will be performed in the 96-well format, while migration and invasion tests in the 24-well format. These experiments will address the question, whether cancer-associated KRAB-ZNF genes may affect tumor biology.

4.4. Task 4 - bioinformatic cross-validation of epigenetic and phenotypic studies

In order to cross-validate the results obtained through molecular and phenotypic studies we intent to perform a series of biostatistical analyses utilizing the clinical and molecular data available from TCGA and ENCODE portals. First, we would like to assess whether level of the mRNA expression of the cancer-associated KRAB-ZNFs may have an impact on the outcome of the disease. To this end, we will use transcriptomic profiling TCGA pan-cancer data and correlate KRAB-ZNFs high and low expression groups with various clinic-pathological and molecular parameters, such as TNM classification, survival, recurrence free survival, age, gender, smoking history, receptor status, mutator or methylator pheontype. In order to estimate differences in clinical outcomes we are going to use the Cox model and Kaplan-Meier curves available from the survival package for R (Therneau, 2015). To quantify similarities and differences between KRAB factors and various patient groups and molecular signatures in cancer we are going to use correspondence analysis implemented in ca package for R (Nenadic & Greenacre, 2007).
Secondly we will take advantage of data that are deposited in ENCODE (Encyclopedia of DNA Elements). Histone and DNA modifications within human genome will be correlated with consensus binding sites of selected KRAB-ZNFs through statistical analyzes in R environment.

4.5. Equipment

The Chair of Medical Biotechnology is fully equipped with all specialized, high tech instruments required for the execution of this project. There are two cell culture laboratories equipped with six laminar hoods, nine CO_2 incubators (including two with nitrogen connection for hypoxic conditions), three inverted fluorescent microscopes with associated cameras and documentation software. The flow cytometry lab possesses flow cytometer and cell sorter machines. The molecular biology lab is equipped with centrifuges, ultracentrifuge, rotors, shakers with temperature control, thermoblocks, 37^0C incubator for bacteria growth, vertical and horizontal electrophoresis units, three transfer units for wet, semi-dry and dry blotting, systems for gel/membrane documentations, Bioruptor sonicator with cooling system and sound-proof box, magnetic stands, Bioanalyzer, Nanodrop and cuvette spectrophotometers, plate reader for the luminescence,

colorimetry and fluorescence measurements, four thermocyclers, two real-time PCR thermocyclers, cell harvester, scintillation beta-counter, autoclave and water purification system. The pyroseuqencing assays will be performed in close collaboration with the Biomarker Group for the University of Liverpool that possesses a pyrosequencer in their lab.

5. Literature

Baubec T, Colombo DF, Wirbelauer C, Schmidt J, Burger L, Krebs AR, Akalin A, Schubeler D. (2015) Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. *Nature* **520:** 243-247.

Cheng Y, Geng H, Cheng SH, Liang P, Bai Y, Li J, Srivastava G, Ng MH, Fukagawa T, Wu X, Chan AT, Tao Q. (2010) KRAB zinc finger protein ZNF382 is a proapoptotic tumor suppressor that represses multiple oncogenes and is commonly silenced in multiple carcinomas. *Cancer Res* **70:** 6516-6526.

Cheng Y, Liang P, Geng H, Wang Z, Li L, Cheng SH, Ying J, Su X, Ng KM, Ng MH, Mok TS, Chan AT, Tao Q. (2012) A novel 19q13 nucleolar zinc finger protein suppresses tumor cell growth through inhibiting ribosome biogenesis and inducing apoptosis but is frequently silenced in multiple carcinomas. *Mol Cancer Res* **10:** 925-936.

Corsinotti A, Kapopoulou A, Gubelmann C, Imbeault M, Santoni de Sio FR, Rowe HM, Mouscaz Y, Deplancke B, Trono D. (2013) Global and stage specific patterns of Kruppel-associated-box zinc finger protein gene expression in murine early embryonic cells. *PLoS One* **8:** e56721.

Dawson MA, Kouzarides T. (2012) Cancer epigenetics: from mechanism to therapy. *Cell* **150**: 12-27.

Esteller M. (2008) Epigenetics in cancer. *N Engl J Med* **358**: 1148–59.

Furuta S, Wang JM, Wei S, Jeng YM, Jiang X, Gu B, Chen PL, Lee EY, Lee WH. (2005) Removal of BRCA1/CtIP/ZBRK1 repressor complex on ANG1 promoter leads to accelerated mammary tumor growth contributed by prominent vasculature. *Cancer Cell* **10:** 13–24.

Groner, A. C., Meylan, S., Ciuffi, A., Zangger, N., et al. (2010). KRAB-zinc finger proteins and KAP1 can mediate long-range transcriptional repression through heterochromatin spreading. *PLoS Genetics* **6**: e1000869.

Florio F, Cesaro E, Montano G, Izzo P, Miles C, Costanzo P. (2010) Biochemical and functional interaction between ZNF224 and ZNF255, two members of the Kruppel-like zinc-finger protein family and WT1 protein isoforms. *Hum Mol Genet* **19**: 3544-3556.

Hanahan D, Weinberg RA. (2011) Hallmarks of cancer: the next generation. *Cell* **144:** 646-674.

Harada Y, Kanehira M, Fujisawa Y, Takata R, Shuin T, Miki T, Fujioka T, Nakamura Y, Katagiri T. (2010) Cell-permeable peptide DEPDC1-ZNF224 interferes with transcriptional repression and oncogenicity in bladder cancer cells. *Cancer Res* **70:** 5829-39.

Hughes LA, Melotte V, de Schrijver J, de Maat M, Smit VT, Bovée JV, French PJ, van den Brandt PA, Schouten LJ, de Meyer T, van Criekinge W, Ahuja N, Herman JG, Weijenberg MP, van Engeland M. (2013) The CpG island methylator phenotype: what's in a name? *Cancer Res* **273**: 5858-5868.

Kalari S, Pfeifer GP. (2010) Identification of driver and passenger DNA methylation in cancer by epigenomic analysis. *Adv Genet* **70**: 277–308.

Kim J, Cantor AB, Orkin SH, Wang J. (2009) Use of in vivo biotinylation to study protein-protein and protein-DNA interactions in mouse embryonic stem cells. *Nat Protoc* **4:** 506-517.

Kim J, Orkin SH. (2011) Embryonic stem cell-specific signatures in cancer: insights into genomic regulatory networks and implications for medicine. *Genome medicine* **3**: 75.

Landt SG, Marinov GK, Kundaje A, Kheradpour P, Pauli F, Batzoglou S, Bernstein BE, Bickel P, Brown JB, Cayting P, Chen Y, DeSalvo G, Epstein C, Fisher-Aylor KI, Euskirchen G, Gerstein M, Gertz J, Hartemink AJ, Hoffman MM, Iyer VR, Jung YL, Karmakar S, Kellis M, Kharchenko PV, Li Q, Liu T, Liu XS, Ma L, Milosavljevic A, Myers RM, Park PJ, Pazin MJ, Perry MD, Raha D, Reddy TE, Rozowsky J, Shoresh N, Sidow A, Slattery M, Stamatoyannopoulos JA, Tolstorukov MY, White KP, Xi S, Farnham PJ, Lieb JD, Wold BJ, Snyder M. (2012) ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. *Genome Res* **22:** 1813-1831.

Li J, Wang Y, Fan X, Mo X, Wang Z, Li Y, Yin Z, Deng Y, Luo N, Zhu C, Liu M, Ma Q, Ocorr K, Yuan W, Wu X. (2007) ZNF307, a novel zinc finger gene suppresses p53 and p21 pathway. *Biochem Biophys Res Commun* **363:** 895-900.

Lin LF, Chuang CH, Li CF, Liao CC, Cheng CP, Cheng TL, Shen MR, Tseng JT, Chang WC, Lee WH, Wang JM. (2010) ZBRK1 acts as a metastatic suppressor by directly regulating MMP9 in cervical cancer. *Cancer Res* **70:** 192-201.

Luebbert M. (2000) DNA methylation inhibitors in the treatment of leukemias, myelodysplastic syndromes and hemoglobinopathies: clinical results and possible mechanisms of action. *Curr Top Microbiol Immunol* **249:** 135–164.

Lupo A, Cesaro E, Montano G, Zurlo D, Izzo P, Costanzo P (2013) KRAB-Zinc Finger Proteins: A Repressor Family Displaying Multiple Biological Functions. *Current genomics* **14:** 268-278.

Minucci S, Pelicci PG. (2006). Histone deacetylase inhibitors and the promise of epigenetic (and more) treatments for cancer. *Nat Rev Cancer* **6:** 38–51.

Najafabadi HS, Mnaimneh S, Schmitges FW, Garton M, Lam KN, Yang A, Albu M, Weirauch MT, Radovani E, Kim PM, Greenblatt J, Frey BJ, Hughes TR. (2015) C2H2 zinc finger proteins greatly expand the human regulatory lexicon. *Nat Biotechnol* **33:** 555-562.

Nenadic O, Greenacre M. (2007) Correspondence Analysis in R, with two- and three-dimensional graphics: The ca package. *Journal of Statistical Software* **20:** 1-13.

Persikov AV, Singh M. (2014) De novo prediction of DNA-binding specificities for Cys2His2 zinc finger proteins. *Nucleic Acids Res* **42:** 97-108.

Qin JY, Zhang L, Clift KL, Hulur I, Xiang AP, Ren BZ, Lahn BT. (2010) Systematic comparison of constitutive promoters and the doxycycline-inducible promoter. *PLoS One* **5:** e10611

Quenneville S, Verde G, Corsinotti A, Kapopoulou A, Jakobsson J, Offner S, Baglivo I, Pedone PV, Grimaldi G, Riccio A, Trono D. (2011) In embryonic stem cells, ZFP57/KAP1 recognize a methylated hexanucleotide to affect chromatin and DNA methylation of imprinting control regions. *Mol Cell* **44:** 361-372.

Sarkar D, Gentleman R, Lawrence M, Yao Z. chipseq: chipseq: A package for analyzing chipseq data. R package version 1.18.0.

Therneau T. (2015). A Package for Survival Analysis in S. version 2.38. http://CRAN.R-project.org/package=survival

Tian C, Xing G, Xie P, Lu K, Nie J, Wang J, Li L, Gao M, Zhang L, He F. (2009) KRAB-type zinc-finger protein Apak specifically regulates p53-dependent apoptosis. *Nat Cell Biol* **11:** 580-591.

Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. (2015) Global cancer statistics, 2012. *CA Cancer J Clin* **65:** 87-108.

Weinstein JN, Collisson EA, Mills GB, Shaw KR, Ozenberger BA, Ellrott K, Shmulevich I, Sander C, Stuart JM. (2013) The Cancer Genome Atlas Pan-Cancer analysis project. *Nat Genet* **45:** 1113-1120.

Witte T, Plass C, Gerhauser C. (2014) Pan-cancer patterns of DNA methylation. *Genome Med* **6**: 66.

Yuan L, Tian C, Wang H, Song S, Li D, Xing G, Yin Y, He F, Zhang L. (2012) Apak competes with p53 for direct binding to intron 1 of p53AIP1 to regulate apoptosis. *EMBO Rep* **13:** 363-370.

Zhu L, Gazin C, Lawson N, Pages H, Lin S, Lapointe D, Green M. (2010). "ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data." *BMC Bioinformatics*, **11:** pp. 237.