Proteomic analysis identifies ZMYM2 as endogenous binding partner of TBX18 protein in 293 and A549 cells

Timo H.-W. Lüdtke¹*, Marc-Jens Kleppa¹*, Reginaldo Rivera-Reyes¹*, Dervla M. Connaughton²,³, Shirlee Shril⁴, Friedhelm Hildebrandt² and Andreas Kispert¹

* These authors contributed equally to this work.

¹Institut für Molekularbiologie, Medizinische Hochschule Hannover, Carl-Neuberg-Str. 1, 30625 Hannover, Germany.
²Schulich School of Medicine and Dentistry, University of Western Ontario, 1151 Richmond St, London, ON N6A 5C1, Canada.
³Department of Medicine, Division of Nephrology, London Health Sciences Centre, 339 Windermere Road, London, ON N6A 5A5, Canada
⁴Department of Medicine, Boston Children’s Hospital, Harvard Medical School, 300 Longwood Ave, Boston, MA 02115, USA.

Correspondence: Andreas Kispert (kispert.andreas@mh-hannover.de)
Abstract

The TBX18 transcription factor regulates patterning and differentiation programs in the primordia of many organs yet the molecular complexes in which TBX18 resides to exert its crucial transcriptional function in these embryonic contexts have remained elusive. Here, we used 293 and A549 cells as an accessible cell source to search for endogenous protein interaction partners of TBX18 by an unbiased proteomic approach. We tagged endogenous TBX18 by CRISPR/Cas9 targeted genome editing with a triple FLAG peptide, and identified by anti-FLAG affinity purification and subsequent LC-MS analysis the ZMYM2 protein to be statistically enriched together with TBX18 in both 293 and A549 nuclear extracts. Using a variety of assays, we confirmed binding of TBX18 to ZMYM2, a component of the CoREST transcriptional corepressor complex. Tbx18 is coexpressed with ZmyM2 in the mesenchymal compartment of the developing ureter of the mouse, and mutations in TBX18 and in ZMYM2 were recently linked to congenital anomalies in the kidney and urinary tract (CAKUT) in line with a possible in vivo relevance of TBX18-ZMYM2 protein interaction in ureter development.
Introduction

T-box (Tbx) genes encode a large family of proteins that are characterized by a conserved DNA-binding domain, the T-box [1]. This 180 amino acid long region recognizes a short conserved stretch of DNA, the T-box binding element (TBE), which can occur in repeats of variable spacing and orientation [2,3]. Binding to these elements in the genome results in transcriptional changes of adjacent genes, the nature of which, i.e. activation or repression, depends on both the composition of the DNA binding site as well as on protein motifs outside the T-box domain to which transcriptional cofactors can bind [3-6]. Besides DNA-binding, the T-box domain mediates interaction with other transcription factors, which is likely to increase target gene specificity. Functional analysis characterized T-box transcription factors as essential regulators of various cellular processes in diverse tissue contexts during metazoan development [4,7]. Not surprisingly, mutations in T-box genes have been identified to cause a large set of human congenital diseases [8].

T-box 18 (TBX18) is a member of a vertebrate-specific subgroup of T-box genes. Mutations in TBX18 were recently recognized to underlie congenital forms of urinary tract malformations in human patients [9]. These phenotypic changes relate to a function of TBX18 in the early development of the ureteric mesenchyme as shown by analysis of mice carrying a null allele of the gene [10,11]. Besides dilatations of the ureter and the renal pelvis (hydroureter nephrosis) [10], mutant mice feature defects in the otic capsule and the otic fibrocyte compartment [12], in smooth muscles of the prostate [13], and in the pericardium and epicardium [14-16]. They also have a hypoplasia of the sinoatrial node and the caval vein myocardium [17,18], and malformation of the vertebral column and the rib cage, the latter of which causes postnatal lethality [19]. Defects in these organ systems were traced to changes in patterning and differentiation of the respective tissue primordia during early organogenesis [10-12,17,19]. To date, both the molecular complexes in which TBX18 acts as well as the transcriptional targets of its activity have remained unexplored in all of its expression domains.

Murine Tbx18 encodes a protein of 613 amino acid residues (aa) that can be divided in an N-terminal region (154 aa), a middle T-box domain (182 aa) and a large C-terminal region (277 aa) [20]. A couple of years ago, we started the biochemical characterization of the TBX18 protein. We found that the T-box domain of TBX18
preferentially binds to a palindromic set of two TBEs in vitro, and that the protein represses transcription upon binding to these elements in cellular transactivation assays. Transcriptional repression depends on a conserved EH1-motif in the N-terminal region that acts as a binding site for TLE3/GROUCHO corepressors [20]. GROUCHO proteins attenuate transcription by facilitating pausing of RNA polymerase II and/or by recruiting histone deacetylases to condensate chromatin [21-24]. The N-terminal region of TBX18 also harbors a short stretch of basic amino acid residues that is necessary and sufficient to mediate nuclear localization in cells [20]. No conserved protein motifs have been identified in the large C-terminal region. However, TBX18 proteins lacking the C-terminal region lose part of their repression function indicating that additional cofactors for transcription modulation bind to this domain [9].

In vitro, TBX18 forms homodimers and heterodimerizes with related T-box proteins. TBX18 also binds to members of the PAX, SIX and NKX homeobox families and to the Zn-finger transcription factor GATA4 [6,20,25,26]. In vivo relevance has been shown for the interaction with PAX3 in somite patterning and with SIX1 in ureter development [25,26].

More recently, we embarked on a proteomic approach to search for TBX18 protein interaction partners in an unbiased fashion. Using overexpressed dual tagged TBX18 as bait, we identified by tandem purification and subsequent liquid chromatography-mass spectrometry (LC-MS) analysis several transcriptional cofactors and homeobox transcription factors as preferred binding partners in 293 cells. The transcriptional cofactors CBFB, GAR1, IKZF2, NCOA5, SBNO2 and CHD7 and all identified homeobox proteins (members of the DUX, GSC, PAX, PBX1 and PRRX families) bound to the T-box of TBX18 in vitro and interacted with TBX18 upon coexpression in 293 cell nuclei [27]. However, overexpression in cells may force unnatural protein interactions, particularly when a relatively promiscuous protein interaction motif such as the T-box is involved. Therefore, we aimed to identify and characterize proteins with which TBX18 forms functionally relevant endogenous complexes in cells. Here, we present a proteomic screen for interaction partners of endogenously tagged TBX18 protein in 293 and A549 cells obtained through CRISPR/Cas9 gene editing. We identified the zinc finger protein ZMYM2 as a binding partner of TBX18 in both cell lines, validated
its binding, provide data for its function as a transcriptional corepressor of TBX18 and discuss the relevance of TBX18 interaction with ZMYM2 in congenital anomalies of the kidney and the urinary tract (CAKUT).
Materials and methods

Mice
NMRI wildtype mice were housed with ad libitum access to food and water under conditions of regulated temperature (22°C) and humidity (50%) and a 12 h light/dark cycle in the central animal facility (ZTL) of the Medizinische Hochschule Hannover. For timing of pregnancies, vaginal plugs were checked in the morning after mating and noon was designated as embryonic day (E) 0.5. Pregnant mice were sacrificed by cervical dislocation, embryos were dissected in PBS, decapitated, fixed in 4% paraformaldehyde (PFA) overnight and stored in 100% methanol at -20°C before further use. All experiments were in accordance with the German Animal Welfare Legislation (§4, TierSchG) and the ARRIVE guidelines, were approved by the local Institutional Animal Care and Research Advisory Committee of the Medizinische Hochschule Hannover and permitted by the Niedersächsisches Landesamt für Verbraucherschutz und Lebensmittelsicherheit (LAVES, Lower Saxony State Office for Consumer Protection and Food Safety; reference number 42500/1H).

Cell culture and transient transfections
293 cells were purchased (ACC 305, DSMZ, Braunschweig, Germany), all other cell lines were kindly provided by Thomas von Hahn (Medizinische Hochschule Hannover, Germany). Cells were cultured in standard DMEM medium (#61965026, ThermoFisher Scientific, Waltham, MA, USA) containing 10% fetal bovine serum (#S0115, Biochrom, Berlin, Germany) and kept in a Sanyo MCO-19ALC incubator (Sanyo, Gunma, Japan) at 37°C with 5% CO₂ under strict sterile conditions. For 293 cells, transient transfections were performed using the calcium phosphate method as previously described [28]; for A549 cells, we used Lipofectamine 3000 (#L3000008, ThermoFisher). We plated 6x10E5 cells in one well of 6-well plates and transfected them with 2.5 µg of DNA and 10 µl of Lipofectamine diluted in 1000 µl of Opti-MEM (#31985062, ThermoFisher). The pd2EGFP-N1 plasmid (Clontech, Mountain View, CA, USA) was independently transfected to check for efficiency of transfection, which was verified by epifluorescence microscopy.

Semi-quantitative reverse transcription PCR
Analysis of TBX18 mRNA expression in human cell lines was done by semi-quantitative RT-PCR with normalization against GAPDH exactly as described before [27].

Generation of A549 and 293 cell lines with FLAG-tagged TBX18

We used a CRISPR/Cas9-mediated knock-in approach based on a protocol from the Medenhall and Myers laboratory [29] to tag endogenous TBX18 with a triple FLAG peptide. Briefly, the homology arms of the TBX18 locus were PCR-amplified from genomic DNA of A549 cells and cloned into the pFETCh-Donor (EMM0021) plasmid (deposited by Eric Medenhall & Richard M. Myers, Addgene, Watertown, Massachusetts, USA) with Gibson Assembly Master Mix (#E2611S, NEB, Ipswich, Massachusetts, USA). Two variants were generated. For Homology Directed Repair (donor) plasmid 1, the 5'-arm was amplified using the primer pair 5'-TCCCCGACCTGGAGCCAGCTACAGGTAACTACCCCTCTC-3' and 5'-CAAGTATCTGCACATATGGTCGGAGCGGAGGAGGTTCCGG-3' (homology arm length of 994 bp, complementary to TBX18, complementary to pFETCh-Donor); the 3'-arm was amplified with primers 5'-AGTTCTTCTGATTCCACATCGTCTAACATATTTTCTTCTTC-3' and 5'-CTTATTATATCTTTGTAAATCTCCCTGGAAAGTCCCTCTCCA-3' (homology arm length of 1023 bp, complementary to TBX18, complementary to pFETCh-Donor).

For donor plasmid 2, the 3'-arm was amplified with 5'-AGTTCTTCTGATTCCACATCAAGAAACTCTCTCTGTGGTTTA-3' and 5'-CTTATTATATCTTGTAAATCTCCCTGGAAAGTCCCTCTCCA-3' which generates a 19 bp deletion compared to donor plasmid 1 in the 3'-UTR of TBX18. In all donor plasmids the endogenous TBX18 stop codons (two consecutive stop codons) were deleted to generate a TBX18 triple FLAG (TBX18.3xFLAG) fusion protein. For details see Supplementary Table S1.

Three guide RNAs, located around the stop codon of TBX18 were designed using the CRISPR design platform CrispRGold [30]. Guide1: 5'-CCAAATGTCTATTTAACCTTAAG-3', location chr6:84736657-84736679, guide2: 5'-CCTTTAAGTTAAATGACATT-3', location chr6:84736657-84736679, guide3: 5'-CCACATAGCCTTTAAAAAGAAA-3', location chr6:84736613-84736635. Guide RNAs were ligated into the CRISPR/Cas9 expression plasmid pX330 (pX330-U6-Chimeric_BB-CBh-hSpCas9) (deposited by Feng Zhang at Addgene, plasmid
#42230) following the published protocol of the Orkin laboratory [31]. For TBX18 tagging, 293 or A549 cells were transfected with one or a combination of two guide RNAs (guide1, 2, 3, 1+3 or 2+3), and selected with G418 (#A291-25, Biochrom) (300 µg/ml for 293 cells, 450 µg/ml for A549 cells) as described [29]. Surviving colonies were expanded and subsequently screened for correct integration by RFLP analysis.

**Southern Blot**

Genomic DNA from single cell derived cultures was screened for correct integration of the 3’-homology arm, the 5’-integration and neomycin by Southern blot analysis of *PstI*-digested genomic DNA. The 5’-probe was amplified from the genomic region adjacent but outside the targeting vector using the primer pair 5’-GCATGGATCCGAGGTCCAACAAATCAACAG-3’ and 5’-CGATGAATTCCCTAGAAACTCAGGCCTAAAAG-3’ and subcloned as 828 bp *BamH*I/*EcoRI* fragment. This probe recognizes a 5.5 kbp *PstI* fragment in the wildtype and a 2.9 kbp *PstI* fragment in the mutant owing to the insertion of a new *PstI* site inside the FLAG-neomycin cassette. The 3’-probe was amplified using the primer pair 5’-GCATGGATCCGGGGCCACCCAAGGATTCTCTG-3’ and 5’-CGATGAATTCCGGTATATAATTCAAGAAGGAG-3’ and subcloned as an 876 bp *BamH*I/*EcoRI* fragment. This probe detects a 5.5 kbp *PstI* fragment in the wildtype and a 3.7 kbp *PstI* fragment in the targeted allele. The *neomycin* (*neo*) probe is a 613 bp fragment, PCR-amplified out of the donor vector using the primers 5’-GGACGAGGCCGCGGTCTAT-3’ and 5’-TCAGAAGAAGTCGCGCAGAGGC-3’. It detects a 3.7 kbp *PstI* fragment in the targeted allele. Probe synthesis was performed with the Rediprime II DNA Labeling System (#RPN1633, Amersham, Little Chalfont, UK) from 50 ng DNA fragment, and labeled with [*α-32P*]dCTP (#SCP-205, Hartmann Analytic, Braunschweig, Germany). Probe synthesis reactions were column purified with NucleoSpin gel and PCR Clean-up kit (#740609.250, Macherey-Nagel, Düren, Germany). Probe hybridization was performed in Church buffer according to standard protocols [32] before exposing the membranes to Fujifilm imaging plates (#BAS-IP MS 2025, Fujifilm, Tokio, Japan) for 3 days. Detection of the DNA fragments was done on a FLA-7000 Laser photodocumentation system (Fujifilm).

**Identification of TBX18 interacting proteins in 293 and A549 cells by LC-MS**
To immunoprecipitate TBX18.3xFLAG containing complexes, we followed a modified batch purification strategy [33]. 293- and A549-TBX18.3xFLAG knock-in cells along with parental control cells were used for nuclear extraction. Twenty dishes of 293 cells at 90% confluency were used to obtain 3x10E9 cells and yielded 90 mg per sample. Ten 15 cm dishes of A549 cells at 90% confluency were used to obtain 3-4x10E8 cells and 40-45 mg of protein per sample. The resulting extracts were used to immunoprecipitate TBX18.3xFLAG containing complexes by batch purification with anti-FLAG M2 affinity gel (#A2220, Sigma Aldrich, St. Louis, Missouri, USA) according to the manufacturer’s instructions. Three independent repetitions for each cell line along with parental cell controls were run in one-well 10% SDS polyacrylamide gels, and the proteins were silver stained. The gels were delivered to the MHH Proteomics facility for fractionation and liquid chromatography-mass spectrometry (LC-MS) analysis. There, proteins were mixed, alkylated by acrylamide and further processed as described [34]. Peptide samples were analyzed with a shotgun approach and data dependent analysis in a LC-MS system (RSLC, LTQ Orbitrap Velos, both Thermo Fisher) as described [34]. Raw MS data were processed using Proteome discoverer 1.4 (ThermoFisher) and Max Quant software (version 1.5) [35], and a database containing human and viral proteins and common contaminants. Proteins were stated identified by a false discovery rate of 0.01 on protein and peptide level. A two tailed Student’s t-test was performed to evaluate the enrichment of proteins identified in the immunoprecipitates of 3xFLAG 293 and A549 knock-in cell lines versus the parental controls. Protein classification was done with the Uniprot database; ontological classification was performed with the database for Annotation, Visualization and Integrated Discovery (DAVID) web program but no functional clustering from the annotations was found [36].

Expression constructs
An expression plasmid for ZMYM2.MYC-FLAG was purchased (#RC205350, OriGene Technologies, Rockville, MD, USA); the KDM1A plasmid was a kind gift from Prof. Yang Shi from Harvard Medical School. Additional seamless cloning of the donor plasmid was performed with the Gibson assembly master mix from NEB (#E2611S, NEB), while ligations of the guide RNA plasmid were performed with T4 DNA ligase (#M0202, NEB). For details see Supplementary Table S1. Expression
constructs for mutant forms of ZMYM2 from CAKUT patients were previously described in detail [37].

**Antibodies**

For Western blot analysis, we used mouse anti-FLAG (1:10000, #F3165, Sigma Aldrich/Merck, Darmstadt, Germany), rabbit anti-MYC (1:10000, # C3956, Sigma Aldrich/Merck), rabbit anti-ZMYM2 (1:5000, #PA031765, Sigma-Aldrich/Merck), goat anti-HA-HRP (1:10000, #ab1265, Abcam), sheep anti-mouse HRP (1:10000, #515-035-003, Dianova, Hamburg, Germany) and goat anti-rabbit-HRP (1:10000, #111-035-045, Dianova). For Immunofluorescent stainings, we used mouse anti-FLAG (1:200, #F3165, Sigma-Aldrich/Merck), rabbit anti-MYC (1:200, # C3956, Sigma Aldrich/Merck), rat anti-HA (1:300, #11867423001, Roche, Mannheim, Germany), rabbit anti-ZMYM2 (1:125, #PA031765, Sigma-Aldrich/Merck), goat anti-rat-Alexa Fluor 555 (1:300, #A21434, #A21202, Invitrogen, Carlsbad, California, USA), donkey anti-rabbit Dylight 488 FAB fragment (1:500, #711-487-003, Dianova) and donkey anti-mouse-Alexa Fluor 488 (1:500, #A21202, Invitrogen). For *in situ* hybridization experiments we used sheep anti-Digoxigenin-AP Fab fragments (#11093274910, Roche). For proximity ligation assays of endogenous protein we used mouse anti-FLAG (1:200, #F3165, Sigma Aldrich/Merck) and rabbit anti-ZMYM2 (1:200, #PA031765, Sigma-Aldrich/Merck), for overexpressed proteins we used rabbit anti-MYC (1:200, # C3956, Sigma Aldrich/Merck) and mouse anti-HA (1:200, #H3663, Sigma-Aldrich/Merck).

**Western blot**

Cell extracts from 90% confluent cells were obtained with RIPA buffer (25 mM Tris-HCl, 150 mM NaCl, 1% NP-40, 1% sodium deoxycholate, 0.1% SDS, pH 7.6) containing 1x complete protease inhibitor (#4693132001, Roche). After homogenization and centrifugation, supernatants were kept at -20°C until further analysis. 20 to 40 µg of protein extract were denatured with 4x Laemmli buffer and loaded onto 5-10% SDS polyacrylamide gels. After running the gels, the proteins were transferred to PVDF membranes (#T830.1, Roth, Karlsruhe, Germany) in semidy blot transfer chambers (#1703848, Bio-Rad, Hercules, California, USA). The membranes were then blocked with 5% non-fat dry milk (#A0830, Applichem, Darmstadt, Germany) in PBS-T (137 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, 1.8
mM KH₂PO₄, 0.1% Tween-20) for 1 h and then incubated an additional hour with manufacturers’ recommended antibody dilutions in blocking solution. After 3 PBS-T washes for 15 min, the membranes were incubated with secondary antibodies coupled to HRP in blocking solution for 1 h. After 3 additional PBS-T washes, the membranes were developed with the CheLuminate-HRP FemtoDetect kit (#A7807, Applichem) and a LAS-4000 photodocumentation system (Fujifilm) with increment exposure and super resolution detection modes unless stated otherwise. PageRuler™ Plus Prestained Protein Ladder, 10 to 250 kDa (#26620, ThermoFisher) was used for size determination.

Immunofluorescent stainings
Immunofluorescent staining was performed on 50-60% confluent cells in 24-well plates. Cells were washed once with PBS and then fixed in either 4% PFA for 10 min or cold methanol for 15 min at 4°C and later permeabilized with 0.5% Triton X-100 for 10 min. Epitopes were blocked with TNB blocking solution (#FP1012, PerkinElmer, Waltham, Massachusetts, USA) for 1 h and later stained with the antibody dilutions recommended by the manufacturers in TNB blocking solution for an hour or overnight at 4°C. After washing 3 times with PBS-T for 15 min they were incubated with secondary antibody dilutions in TNB blocking solution for an hour. After counterstaining with DAPI (4’,6-Diamidino-2-phenylindole dihydrochloride, #6843.2, Roth) and 3 last washing steps, the stainings were documented in a DM6000 inverted microscope (Leica, Wetzlar, Germany).

Nuclear recruitment assays
To check for interactions inside the nucleus, we cotransfected 293 cells with plasmids encoding a tagged form of full-length ZMYM2 with a variant of TBX18 which lacked the nuclear localization signal (TBX18ΔNLS) [20]. Mutant forms of ZMYM2 from CAKUT patients were cotransfected with either full-length TBX18 or TBX18ΔNLS. One day after transfection, cells were fixed with PFA and immunofluorescent stainings against the different protein tags were carried out and documented with a DM6000 microscope.

GST pull-down assays
To check for physical interaction, we performed Glutathione S-Transferase (GST) fusion protein pull-downs as described previously [20]. GST-fusion proteins of the different TBX18 domains were produced in the BL21 *E. coli* strain after IPTG induction. The GST fusion proteins were batch purified from the bacterial lysates using glutathione sepharose 4B beads (#17075601, GE, Chicago, Illinois, USA). To produce radioactively tagged candidate proteins we used a reticulocyte lysate based TNT Quick Coupled Transcription/Translation system (#L2080, #L1170, Promega, Madison, WI, USA) with $^{35}$S-labeled Methionine/Cysteine 75/25% (#SRIS-103, Hartmann Analytic) on plasmids with a suitable promoter (see Supplementary Table S1 for details on plasmids used). To perform the pull-down assays, the bead solutions containing the GST-TBX18 fusion proteins were incubated with the *in vitro* produced radioactively-labeled candidate proteins and washed several times with HEPES buffer (20 mM HEPES, 10 mM KCl, 5 mM MgCl$_2$, 0.1 mM EDTA, 5% Glycerol, 0.5% Triton X-100, 80 mM NaCl, pH 7.5). Analysis of the pulled-down proteins was achieved by separating the proteins by denaturing SDS-PAGE, heat-drying the resulting gels on filter paper and then exposing them to phosphorescent imaging plates for 1 day. Detection of the proteins was done on a FLA-7000 Laser photo documentation system (Fujifilm) with a sensitivity of S10000 and a pixel size of 50 µm. HiMark™ Prestained Protein Ladder, 10 to 250 kDa (#LC5699, ThermoFisher) was used for size determination.

**Immunoprecipitation assays**

To perform immunoprecipitations of endogenous proteins, we used whole cell extracts from 293 and 293-TBX18.3xFLAG knock-in cells. Cells were cultured to 90% confluency on 10 cm dishes and then washed with PBS before harvesting them twice with PBS and a plastic scraper. The cells were lysed with 1 ml RIPA buffer, and the lysate was homogenized and centrifuged. 100 or 200 µl of the supernatants were used to detect the interaction in either direction either with 10 µl of direct anti-FLAG M2 affinity gel batch immunoprecipitation (#A2220, Sigma Aldrich) or immunoprecipitation with 2 µl of the primary antibodies against ZMYM2 together with 20 µl of protein A-agarose beads (#sc-2001, Santa Cruz, Santa Cruz, California, USA) according to the manufacturer's protocol. The FLAG beads were washed four times with TBS-T buffer, while the protein A beads with or without antibody were
washed thrice with HEPES Buffer 2 (25 mM HEPES, 1 mM MgCl₂, 10% Glycerol, 0.1% NP40, 75 mM NaCl, 1 mM DTT, 10 µg/ml BSA, pH 7.4) before boiling the beads with Laemmli buffer to release the pulled proteins. Western Blot analysis was used to detect the presence of the proteins in either direction of interaction.

Transactivation assays
Transactivation assays were done on lysates of 293 cells with the Dual-Luciferase Reporter Assay System (#E1910, Promega) as previously reported [27]. We seeded 50-60% confluent 293 cells on 6-well plates and transfected a total of 4 µg of the plasmid combinations, including pCMV-EGFP and pRL-tk as transfection controls, into the cells with the calcium phosphate method. One day after transfection, we lysed the cells and followed the manufacturer’s protocol to automatically measure the firefly and renilla luciferase activities in duplicates with a Glomax luminometer (#9100-002, Turner Biosystems, Sunnyvale, California, USA). Relative luciferase units were calculated by dividing the firefly to renilla luciferase activity ratio from each well and later normalizing this ratio to the firefly luciferase reporter transfection (pGL3.Prom.Tbx18BS2) [27] control. Four biological replicates were used for statistical analysis. A two-tailed Student’s t-test was carried between all normalized measurements between pairs of samples adjusting for equal or unequal variance according to an F-test. Significance was defined as P≤0.05 (*) significant, P≤0.01 (**) very significant or P≤0.001 (***) extremely significant.

RNA in situ hybridization
10-µm paraffin sections of the posterior trunk region of embryonic day (E) 12.5 wildtype NMRI mice were subjected to RNA hybridization with digoxigenin-labelled antisense riboprobes as previously described [38].

Proximity ligation assay (PLA)
PLA was performed with Duolink™ In Situ Red Starter Kit Mouse/Rabbit (DUO92101, Sigma Aldrich/Merck) as recently described [39]. Cells were seeded on collagen coated glass slides. For coating, collagen solution (C3867-1VL, Sigma Aldrich/Merck) was diluted to a concentration of 50 µg/ml in 0.1 N acetic acid, sterile filtered with 0.2 µm Whatman® Puradisc 30 syringe filters FP30/0.2 CA-S (Z612529, Sigma Aldrich/Merck) and applied to glass slides overnight. Plates were stored at room
temperature until further use. For endogenous ZMYM2 and TBX18, cells were permeabilized with 0.1% Triton™ X-100 for 15 minutes and washed in PBS for 3 times before application of the PLA kit without further fixation. Cells transfected with overexpression vectors were fixed in ice-cold 100% methanol for 5 minutes and 10% PFA for 20 minutes, and washed in PBS for 3 times. Epitopes were blocked with TNB blocking solution (#FP1012, Perkin-Elmer, Waltham, Massachusetts, USA) for 1 h. Primary antibodies were diluted in TNB blocking buffer and incubated for 2 h to overnight. The following steps were performed as specified in the manufacturer’s protocol.
Results
CRISPR/Cas9 technology confers a 3xFLAG tag to endogenous TBX18 in 293 and A549 cell lines

Identification of TBX18 interaction partners in embryonic contexts is hampered by the scarcity of the material, whereas overexpression approaches in vitro may be compromised by unphysiologically high protein levels forcing unnatural interactions. We reasoned that cell lines with endogenous TBX18 expression provide a useful alternative for identification of relevant binding partners of TBX18. We recently reported that 293 embryonic kidney cells show increased expression of TBX18 in a set of 6 human cell lines tested [27]. An expanded search identified elevated levels of TBX18 in A549 lung carcinomas and HPMEC pulmonary microcapillary endothelial cells as well (Supplementary Figure S1).

Since available antibodies proved unsuitable for precipitation of endogenous protein complexes of TBX18 in these cell lines in our hands, we decided to exploit the CRISPR/Cas9 technology to introduce a triple FLAG (3xFLAG) affinity tag fragment [40,41] into the TBX18 locus of both 293 and A549 cells [29,42]. Human TBX18 is localized on the long arm of chromosome 6 (6q14.3) and comprises 8 exons. To correctly integrate the 3xFLAG at the C-terminus of the encoded TBX18 protein, we designed guide RNA plasmids to direct the Cas9 nuclease to a protospacer adjacent motif (PAM) close to the TBX18 stop codon, and homology directed recombination (donor) plasmids containing a modified region of the 8th exon to erase the stop codon, followed by a 3xFLAG motif, a P2A cleavage sequence and a neomycin cassette for antibiotic resistance, flanked by 5´- and 3´-regions homologous to the TBX18 locus (Figure 1A, for details of these and all other constructs see Supplementary Table S1).

Cells transfected with these two constructs were grown under neomycin selection using G418, and individual surviving clones expanded. We validated the targeting event by RFLP analysis on genomic DNA extracted from the clones. For A549 cells, 4 out of 39 tested clones carried a heterozygous knock-in, 1 was homozygous; all 5 clones were devoid of additional integrations of the neo cassette (Supplementary Figure S2). For 293 cells, we identified 5 clones with correct heterozygous targeting of TBX18 and lack of additional off-target integrations of the selection cassette (n=45) (Supplementary Figure S3). We choose one 293 heterozygous clone (293-
P2C2), one heterozygous (A549-12C3) and one homozygous A549 clone (A549-12C4) for further characterization. Western blot analysis using an anti-FLAG antibody detected a band at approximately 90 kDa - which is slightly bigger than the calculated theoretical mass of 68.7 kDa - in the targeted but not in the parental cell lines (Figure 1B, Supplementary Figure S4). Immunofluorescent stainings revealed a speckle-like pattern of the FLAG tag in the nucleus of the TBX18.3xFLAG cell lines but not in the parental cell lines (Figure 1C) in agreement with a (predominant) nuclear localization of TBX18 in 293 and A549 cells.

**LC-MS identifies ZMYM2 as a novel TBX18 interacting protein in 293 and A549 cells**

To identify TBX18 interacting proteins in 293 and A549 cells, we used our newly generated TBX18.3xFLAG cell lines (heterozygous 293-P2C2, homozygous A549-12C4) for affinity chromatography of native complexes. After cell expansion and nuclear extraction, protein complexes harboring TBX18.3xFLAG were purified with an anti-FLAG batch affinity strategy [43,44]. The purified protein complexes were recovered by heat treatment and resolved by SDS-PAGE. The polyacrylamide gels of three independent experiments and controls for each cell line were sent to the Hannover Medical School Proteomics Facility for protein extraction and subsequent LC-MS analysis. A total of 115 human proteins were enriched in anti-FLAG immunoprecipitates from 293-P2C2 compared to parental 293 cells, from which 10 passed a stringent t-test statistical analysis of the intensity scores of the three samples of the 293-TBX18.3xFLAG cells compared to the three samples of the 293 parental cells. Four of the 10 candidates were Zn-finger proteins (PATZ2, ZNF143, ZMYM2, ZMYM3) (Figure 2, Supplementary Table S2). In anti-FLAG immunoprecipitates from A549-12C4 cells, 270 human proteins were more strongly recovered compared to the parental A549 cells. Here, 36 proteins were statistically significantly enriched in the three IP samples of the A549-TBX18.3xFLAG cells compared to A549 control cells (Figure 2, Supplementary Table S3). These included the transcription factor RFX1, nuclear corepressors like NCOR2 and EMSY, a cluster of proteins from the 60S ribosomal subunit (RPL4, RPL14, RPL15, RPL18) and the zinc finger proteins ZNF609 and ZMYM2. Importantly, TBX18 was the most enriched protein in immunoprecipitates from both cell lines confirming correct FLAG-tagging of the protein by our approach.
To recover cell context-independent binding partners of TBX18, we overlapped the lists of statistically significantly enriched candidates obtained from 293-P2C2 and A549-12C4 cell precipitates and identified besides TBX18 a single protein, Zinc finger, MYM-type 2 (ZMYM2), a component of the CoREST/BHC transcriptional corepressor complex [45-48]. Importantly, ZMYM2 was only found once as background in 136 datasets of FLAG precipitations in 293 cells as listed in the crapome database for common affinity purification contaminants [49] making it very unlikely that ZMYM2 is merely a contaminant in our IP experiment.

ZMYM2 interacts with endogenous TBX18 in the nucleus of 293 cells

To validate the finding of our proteomic approach, we independently assayed for endogenous interaction of ZMYM2 and TBX18 in 293 cells. Using a commercial antibody against ZMYM2 we detected a band running slightly higher than the expected size (154.9 kDa) in lysates of both 293 parental and 293-TBX18.3xFLAG knock-in cells (Figure 3A). In immunofluorescent analysis, anti-FLAG or anti-ZMYM2 antibodies detected predominantly nuclear antigens in 293-TBX18.3xFLAG cells and a minor cytoplasmic localization of the respective proteins (Figure 3B). To validate a possible direct interaction, we applied a proximity ligation assay. Using only a single antibody against TBX18.FLAG or ZMYM2 did not yield a staining. Combined application of both antibodies resulted in mainly nuclear PLA signals and little staining in the cytoplasm (Figure 3C). Thus, we identified ZMYM2 as a mainly nuclear protein that interacts with TBX18 in 293-TBX18.3xFLAG cells.

Overexpressed ZMYM2 binds to TBX18 in the nucleus of 293 cells

To further validate ZMYM2-TBX18 interaction, we transiently overexpressed ZMYM2.MYC-FLAG and TBX18.MYC proteins in 293 cells. After immunoprecipitation with an anti-FLAG antibody, we detected both ZMYM2.MYC-FLAG and TBX18.MYC protein on a western blot by an anti-MYC antibody indicating coprecipitation (Figure 4A). To further test whether overexpressed TBX18 productively interacts with ZMYM2 in 293 cells, we made use of a nuclear recruitment assay that we previously described [20]. It is based on the identification of a classical nuclear localization signal (NLS) at the N-terminus of the TBX18 protein. When this NLS is deleted, the resulting protein (TBX18ΔNLS) is excluded from the nucleus but can be shuttled back to this compartment by binding to a protein, which carries such a signal. Upon
co-expression of TBX18ΔNLS and ZMYM2.MYC-FLAG, we observed a statistically significant nuclear translocation of TBX18ΔNLS protein (Figure 4B,C, Supplementary Table S4A). Finally, we performed a PLA experiment with overexpressed proteins in 293 cells (Figure 4D). Since TBX18 forms homodimers [20], combined transfection of TBX18.MYC and TBX18.HA served as a positive control. In both, TBX18.MYC/TBX18.HA and TBX18.HA/ZMYM2.MYC-FLAG PLA, strong nuclear staining was detected, indicating close contact of overexpressed TBX18.HA and ZMYM2.MYC-FLAG proteins in the nucleus.

**ZMYM2 interacts with the T-box region of TBX18**

To delineate the region of the TBX18 protein, which interacts with ZMYM2, we performed pull-down assays using a previously described series of bacterially expressed fusion proteins of GST with the N- and C-terminal region and the T-box of TBX18 [20] and radioactively labeled ZMYM2 produced with an *in vitro* expression system. ZMYM2 interacted with the T-box and with the N-terminal plus T-box region of TBX18, indicating that the T-box of TBX18 serves as an interaction domain with ZMYM2 (Figure 5).

**ZMYM2 mutant proteins of CAKUT patients are partly compromised in binding to TBX18**

Recent work identified mutations in *ZMYM2* as a cause of congenital anomalies of the kidney and urinary tract (CAKUT) in human patients [37]. Since CAKUT have also been observed in patients with mutations in *TBX18* [9], ZMYM2 and TBX18 may functionally interact *in vivo*. To validate such a possibility, we analyzed whether Zmym2 is coexpressed with Tbx18 in the undifferentiated mesenchymal compartment of the embryonic murine ureter at a stage, when *TBX18* has a critical function in this tissue [10,11]. *In situ* hybridization on transverse sections of the proximal ureter region of E12.5 embryos indeed found coexpression of the two genes in the ureteric mesenchyme. Expression of Zmym2 was also found in the inner epithelial tissue in agreement with previous reports that Zmym2 has a widespread low expression in the developing urogenital system [37] and many developing and mature organs ([www.proteinatlas.com](http://www.proteinatlas.com)) [50].
We next asked whether mutations in ZMYM2 that have been characterized in CAKUT patients [37], compromise binding to TBX18. We therefore re-evaluated mutant ZMYM2 localization in 293 cells for the described truncated patients’ variants p.Gly257*, p.Gln398*, p.Cys536; Leufs*13, p.Arg540*, p.Tyr763Glnfs*6, p.(Lys812Aspfs*18) and p.Cys823* by immunofluorescence against the MYC-tag (depicted in Figure 6B) and applied a nuclear recruitment assay with TBX18 (Figure 6C). p.Tyr763* was the only truncated form that was predominantly expressed in the nucleus. Accordingly, TBX18ΔNLS.HA was used in the nuclear recruitment assay for this variant of ZMYM2. All other ZMYM2 variants, that were located almost completely in the cytosol, were tested with full length TBX18.HA protein for nuclear recruitment (Figure 6C-E). Immunofluorescent images (Figure 6C) and further statistical evaluation (Figure 6D,E, Supplementary Table S4B-D) demonstrated, that the short truncated versions of ZMYM2 including p.Gly257*, p.Gln398*, p.Cys536; Leufs*13 and p.Arg540* were recruited by TBX18 to the nucleus, while the longer variants p.Tyr763Glnfs*6, p.(Lys812Aspfs*18) and p.Cys823* were not. These results indicate that the binding domain of ZMYM2 to TBX18 is very likely located in the N-terminal region of the protein. The middle part of ZMYM2, which is still present in the longer ZMYM2 truncations, may either structurally mask the TBX18 binding site or antagonize binding.

**ZMYM2 acts as a corepressor for TBX18**

To analyze whether ZMYM2 modulates TBX18 transcriptional activity, we performed dual luciferase reporter assays in 293 cells. We used a recently described reporter plasmid in which a palindromic repeat of two TBEs is cloned in front of an SV40 minimal promoter and a firefly luciferase gene, together with expression constructs of TBX18 and candidate proteins in various combinations, and a Renilla luciferase control plasmid [27] (Figure 7, Supplementary Table S5). Under the chosen conditions, TBX18 reduced the reporter activity by approximately 40% in comparison to an empty expression vector. TLE3/GROUCHO further increased this repression as previously shown [20]. ZMYM2 dose-dependently increased TBX18 repression to a similar end-point as TLE3/GROUCHO. The corepressive activity of TLE and ZMYM2 on TBX18 was abolished when both proteins were coexpressed (Figure 7A). Of note, ZMYM2 had no effect on reporter expression in absence of TBX18 (Figure 7B).
ZMYM2 was shown to be a member of the CoREST corepressor complex, that harbors additional proteins such as KDM1A [45-48]. Since KDM1A was recovered in both of our proteomic screens in both 293 and A549 cells (Supplementary Table S2, S3), we tested it for its effect on TBX18 repressive activity. In fact, we found that KDM1A enhanced TBX18 repressive activity in a concentration-dependent manner. Notably, the addition of ZMYM2 to KDM1A relieved the corepression exerted by KDM1A similar to the effect of ZMYM2 on TLE3/GROUCHO (Figure 7C). Again, neither KDM1A alone nor the combination of KDM1A and ZMYM2 exerted transcriptional repressing activity on the luciferase reporter in absence of TBX18 protein (Figure 7D). We conclude that ZMYM2 and KDM1A, members of the CoREST corepressor complex, act as transcriptional corepressors of TBX18 in the chosen reporter system.
Discussion

Here, we used affinity purification and mass spectrometry to characterize the interactome of a tagged version of endogenous TBX18 in 293 and A549 cells. We found ZMYM2 as the only statistically enriched protein in anti-FLAG immunoprecipitates from both 293 and A549 cells with endogenously FLAG-tagged TBX18, and confirmed direct binding of ZMYM2 to the T-box of TBX18. Our work implicates the CoREST corepressor complex in TBX18 function, and suggests that TBX18 and ZMYM2 interact in the ureteric mesenchyme in vivo.

Immunoprecipitation of an endogenous FLAG tagged TBX18 protein identifies largely distinct sets of interacting proteins in A549 and 293 cells

Our affinity mediated immunoprecipitation approach against an engineered FLAG tag at the C-terminus of TBX18 protein identified both in 293 and A549 cells a large number of proteins coprecipitating with TBX18. These extensive lists (115 and 270 proteins, respectively) comprised proteins localized outside the nucleus (e.g. cytoskeletal proteins) indicating that the nuclear extracts used for the immunoprecipitations were not purified to homogeneity. Furthermore, we detected many proteins (e.g. the ribonucleoproteins RALY and HNRNPN) that have been frequently detected in similar anti-FLAG immunoprecipitation experiments as documented in the crapome database [49]. They may represent proteins that are unspecifically recognized by the anti-FLAG antibody, that adhere strongly to the binding matrix or that may preferentially denature to form adhesive aggregates. In any case, our findings enforce the notion that affinity purification against a single tag provides more false positive candidates compared to a tandem purification strategy available for dual tagged proteins.

Given these considerations, we decided to increase the stringency of our approach by selecting proteins that were statistically significantly enriched in the three individual immunoprecipitations from 293- and A549-TBX18.3xFLAG cells, respectively. By this we ended up with 10 candidate proteins for 293- and 36 proteins for A549-TBX18.3xFLAG cells, respectively, of which only two were in common: TBX18, as expected, and the transcriptional corepressor ZMYM2. One may argue that the poor overlap reflects differential gene expression in the two cell lines. In fact, A549 and 293 cells derive from two different organs and exhibit cell-type specific
signatures [50]. However, a comparison of the expression level of the genes encoding the 10 and 36 candidate interactors, respectively, in the two cell lines provides a complex picture (RNA-Seq data published in [50], Supplementary Table S6). Most of these genes exhibit similar expression levels in 293 and A549 cells despite the fact that the encoded proteins were enriched only in 293 or in A549 cells. Few candidates were enriched in cells despite low-level expression in these cells (e.g. TGM3 in 293 cells, RPL14 in A549 cells). Many candidates were enriched from the cells in which they show higher expression (e.g. NEFL in 293 cells, ITGB4 in A549 cells). This suggests that gene expression levels presents an important factor for identification of interactors but other parameters such as protein stability and modification, relative ratio and presence of auxiliary proteins may impinge on the possibility to productively interact with TBX18 in 293 and A549 cells.

We recently used overexpressed dual-tagged TBX18 as bait to identify by tandem affinity purification and subsequent LC-MS analysis TBX18 interacting proteins in 293 cells. We identified several transcriptional cofactors and a large number of homeobox transcription factors as preferred binding partners [27]. To our surprise, the interactome revealed for endogenous TBX18 in our current study is vastly different. Most notably, we found only few tissue-specific transcription factors, which all belonged to the Zn-finger family, and not to the homeobox protein family. It seems possible that in an overexpression approach, the endogenous binding partners of TBX18 are limited and that weak or even unnatural interactions are overrepresented or newly established. Such a behavior is eased by the fact that the T-box seems to act rather promiscuously by binding to many very different types of proteins or motifs. Alternatively, TBX18 expression in A549 and 293 cells may be extremely low to interact with very few other proteins. Irrespective of the precise mechanism, it shows that the expression level (e.g. low endogenous vs highly overexpressed) of a protein and the mode of purification (single tag vs dual tag) can dramatically affect the identification of relevant binding partners in cellular systems.

**ZMYM2 may present a biologically meaningful protein interaction partner of TBX18**

ZMYM2 was the protein with the highest significance score found in our proteomics analyses both from 293 and A549 immunoprecipitates. ZMYM2 interacted with endogenous and overexpressed TBX18 in the nucleus of 293 cells, and bound to the
T-box of TBX18 *in vitro*. Moreover, ZMYM2 enhanced the repressive activity of TBX18 in cellular transactivation assays but did not affect reporter activity on its own, providing a strong hint that this Zn-finger protein acts a corepressor for TBX18 in our experimental system.

Such a notion is supported by previous findings that ZMYM2 is a nuclear protein that binds to chromatin and recruits a repressive complex consisting of the histone demethylase KDM1A/LSD1, the cofactor CoREST (RCOR1 or RCOR2), and a histone deacetylase (HDAC1 or HDAC2) [48,51]. We recently showed that RCOR3, an occasional member of the CoREST complex [46] is able to interact with TBX18 in a nuclear recruitment assay, but this interaction was indirect and only altered transcription repression by TBX18 in a concentration-independent manner [27]. Moreover, our current proteomics analysis recovered with ZMYM3, ZMYM4 and KDM1A additional components of the CoREST/BHC transcriptional corepressor complex. KDM1A increased the TBX18 transcriptional repressive activity similar to ZMYM2. Hence, enrichment of 5 components of the CoREST complex in TBX18 immunoprecipitates (RCOR3, KDM1A, ZMYM2, ZMYM3, ZMYM4) and corepressive activity of KDM1A and ZMYM2 indicates the potential significance of this protein complex in mediating transcriptional repression by TBX18, at least in 293 and A549 cells.

Our transactivation assays showed that combinatorial expression of high doses of KDM1A and ZMYM2 attenuated the individual corepressive activity of the two proteins. At present, we have no explanation for this finding but we speculate that formation of transcriptionally unproductive complexes with unknown proteins occurs or that the stoichiometry between the components of the CoRest complex is disturbed.

Increasing concentrations of ZMYM2 similarly dampened TLE corepressive activity in our transactivation assay. Since both TLE and ZMYM2 bind to HDACs [21,48,51], competition for this class of proteins may underlie this effect. Alternatively, TLE and ZMYM2 may compete for binding to TBX18. Irrespective of the exact reason, this finding indicates that TLE and ZMYM2 corepressors act in a mutually exclusive i.e. cell-type specific manner. It is important to note that interaction of TBX18 with corepressive complexes has only been shown for cellular systems, such as 293 and A549 cells. In the future, TBX18 transcriptional complexes should be characterized for the various sites of *Tbx18* expression in the developing embryo to obtain a
conclusive idea on a possible heterogeneity of TBX18 associated chromatin complexes in vivo.

A recent genomic association study identified ZMYM2 as a gene underlying CAKUT in human patients [37]. In total, 14 different heterozygous loss-of-function mutations in ZMYM2 were identified in 15 unrelated families. The phenotypic changes in the urinary system of the patients included hydrourerter with and without hydronephrosis, ureteropelvic junction obstruction and renal agenesis. Mice with a heterozygous inactivation of Zmym2 showed a similar spectrum of CAKUT-like phenotypes including hydrourerter, duplex and cystic kidneys, and vesicoureteral reflux shortly at birth [37]. Importantly, these phenotypic changes overlap with those reported for mice and men with inactivating mutations in Tbx18/ TBX18 [9,10]. Moreover, our expression analysis showed that Tbx18 and Zmym2 are coexpressed in the ureteric mesenchyme of mouse embryos at E12.5, the stage at which Tbx18 is required for specification of this tissue. It is tempting to speculate that ZMYM2 and TBX18 physically interact in the ureteric mesenchyme and repress a set of genes possibly in combination or as part of the CoREST repressor complex to mediate early ureter development in mammals.

Importantly, in our nuclear recruitment assays short truncated versions of ZMYM2 were efficiently recruited into the nucleus by TBX18, indicating presence of an interaction domain in the N-terminal part of ZMYM2. However, missing the complete DNA binding domain, these truncated proteins probably remain non-functional in the nucleus. Unexpectedly, the middle part of ZMYM2, that largely consists of Zinc finger motifs and is responsible for DNA binding of ZMYM2, apparently hampers interaction with TBX18 in contrast to full length ZMYM2. This is similar to binding behaviour of truncated TBX18 versions to full length ZMYM2, where, in GST experiments, the T-box alone bound more efficiently to ZMYM2 than the N-terminal plus T-box truncation. Alterations in folding structure in these truncated proteins may mask the respective binding domains that become accessible with correct folding in the full-length proteins.

ZMYM2 patients also have extrarenal disease manifestations including skeletal abnormalities, hypoplastic hands, feet and nails and scoliosis but also cardiac septal defects [37]. Intriguingly, Tbx18-deficient mice display malformation of the vertebral column. Moreover, mice and men with homozygous loss of the closely related Tbx15/TBX15 gene exhibit numerous skeletal defects [52,53], while heterozygous
loss of another subfamily member, Tbx20/TBX20, results in a range of cardiac defects [54,55]. It is tempting to speculate that ZMYM2 interacts with TBX18 and other members of the protein family at numerous sites of developmental expression for transcriptional repression. Further genetic interaction and molecular studies in vivo may address this hypothesis in the future.
Data Availability
All supporting data are included within the main article and its Supplementary Files.

Competing Interests
The authors declare that there are no competing interests associated with the manuscript.

Funding
D.M.C. was funded by Health Research Board, Ireland (HPF-206-674), the International Pediatric Research Foundation Early Investigators’ Exchange Program and the Amgen® Irish Nephrology Society Specialist Registrar Bursary. She is now funded by the Eugen Drewlo Chair for Kidney Research and Innovation at the Schulich School of Medicine & Dentistry at Western University, London, Ontario, Canada. The work in the laboratory of A.K. was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) for the Cluster of Excellence REBIRTH (From Regenerative Biology to Reconstructive Therapy) at Hannover Medical School, and by the individual grant DFG KI728/7-2.

Author Contribution
Study concept and design, T.H.L., R.R.-R., M-J.K and A.K. Acquisitions of data, R.R.-R., M-J.K., T.H.L.. Drafting of the manuscript T.H.L., R.R.-R. and A.K. Providing essential reagents and information, D.M.C., S.S., F.H.. Critical revision of manuscript, all authors. Statistical analysis, T.H.L., R.R.-R., M-J.K.

Acknowledgements
We thank Christopher Binz and Imke Peters for technical assistance, Yang Shi (Harvard Medical School) for plasmids and Thomas von Hahn (Medizinische Hochschule Hannover) for cell lines. Proteomics data used in this publication were generated by the Research Core Unit Proteomics at Hannover Medical School.
References

1 Sebe-Pedros, A. and Ruiz-Trillo, I. (2017). Evolution and Classification of the T-Box Transcription Factor Family. Curr Top Dev Biol 122, 1-26.
2 Kispert, A. and Herrmann, B. G. (1993). The Brachyury gene encodes a novel DNA binding protein. EMBO J 12, 3211-3220.
3 Kispert, A., Koschorz, B. and Herrmann, B. G. (1995). The T protein encoded by Brachyury is a tissue-specific transcription factor. EMBO J 14, 4763-4772.
4 Papaioannou, V. E. (2014). The T-box gene family: emerging roles in development, stem cells and cancer. Development 141, 3819-3833.
5 Waldron, L., Steimle, J. D., Greco, T. M., Gomez, N. C., Dorr, K. M., Kweon, J., Temple, B., Yang, X. H., Wilczewski, C. M., Davis, I. J., Cristea, I. M., Moskowitz, I. P. and Conlon, F. L. (2016). The Cardiac TBX5 Interactome Reveals a Chromatin Remodeling Network Essential for Cardiac Septation. Dev Cell 36, 262-275.
6 Kaltenbrun, E., Greco, T. M., Slagle, C. E., Kennedy, L. M., Li, T., Cristea, I. M. and Conlon, F. L. (2013). A Gro/TLE-NuRD corepressor complex facilitates Tbx20-dependent transcriptional repression. J Proteome Res 12, 5395-5409.
7 Naiche, L. A., Harrelson, Z., Kelly, R. G. and Papaioannou, V. E. (2005). T-box genes in vertebrate development. Annu Rev Genet 39, 219-239.
8 Ghosh, T. K., Brook, J. D. and Wilsdon, A. (2017). T-Box Genes in Human Development and Disease. Curr Top Dev Biol 122, 383-415.
9 Vivante, A., Kleppa, M. J., Schulz, J., Kohl, S., Sharma, A., Chen, J., Shril, S., Hwang, D. Y., Weiss, A. C., Kaminski, M. M., Shukrun, R., Kemper, M. J., Lehnhardt, A., Beetz, R., Sanna-Cherchi, S., Verbitsky, M., Gharavi, A. G., Stuart, H. M., Feather, S. A., Goodship, J. A., Goodship, T. H., Woolf, A. S., Westra, S. J., Doody, D. P., Bauer, S. B., Lee, R. S., Adam, R. M., Lu, W., Reutter, H. M., Kehinde, E. O., Mancini, E. J., Lifton, R. P., Tasic, V., Lienkamp, S. S., Juppner, H., Kispert, A. and Hildebrandt, F. (2015). Mutations in TBX18 Cause Dominant Urinary Tract Malformations via Transcriptional Dysregulation of Ureter Development. Am J Hum Genet 97, 291-301.
10 Airik, R., Bussen, M., Singh, M. K., Petry, M. and Kispert, A. (2006). Tbx18 regulates the development of the ureteral mesenchyme. J Clin Invest 116, 663-674.
11 Bohnenpoll, T., Bettenhausen, E., Weiss, A. C., Foik, A. B., Trowe, M. O., Blank, P., Airik, R. and Kispert, A. (2013). Tbx18 expression demarcates multipotent precursor populations in the developing urogenital system but is exclusively required within the ureteric mesenchymal lineage to suppress a renal stromal fate. Dev Biol 380, 25-36.
12 Trowe, M. O., Maier, H., Schweizer, M. and Kispert, A. (2008). Deafness in mice lacking the T-box transcription factor Tbx18 in otic fibrocytes. Development 135, 1725-1734.
13 Bolt, C. C., Negi, S., Guimaraes-Camboa, N., Zhang, H., Troy, J. M., Lu, X., Kispert, A., Evans, S. M. and Stubbs, L. (2016). Tbx18 Regulates the Differentiation of Periductal Smooth Muscle Stroma and the Maintenance of Epithelial Integrity in the Prostate. PLoS One 11, e0154413.
14 Norden, J., Grieskamp, T., Christoffels, V. M., Moorman, A. F. and Kispert, A. (2012). Partial absence of pleuropericardial membranes in Tbx18- and Wt1-deficient mice. PLoS One 7, e45100.
15 Greulich, F., Farin, H. F., Schuster-Gossler, K. and Kispert, A. (2012). Tbx18 function in epicardial development. Cardiovasc Res 96, 476-483.
16 Wu, S. P., Dong, X. R., Regan, J. N., Su, C. and Majesky, M. W. (2013). Tbx18 regulates development of the epicardium and coronary vessels. Dev Biol 383, 307-320.
17 Christoffels, V. M., Mommersteeg, M. T., Trowe, M. O., Prall, O. W., de Gier-de Vries, C., Soufan, A. T., Bussen, M., Schuster-Gossler, K., Harvey, R. P., Moorman, A. F. and Kispert, A. (2006). Formation of the venous pole of the heart from an Nkx2-5-negative precursor population requires Tbx18. Circ Res 98, 1555-1563.
18 Wiese, C., Grieskamp, T., Airik, R., Mommersteeg, M. T., Gardiwal, A., de Gier-de Vries, C., Schuster-Gossler, K., Moorman, A. F., Kispert, A. and Christoffels, V. M. (2009). Formation of the sinus node head and differentiation of sinus node myocardium are independently regulated by Tbx18 and Tbx3. Circ Res 104, 388-397.
19 Bussen, M., Petry, M., Schuster-Gossler, K., Leitges, M., Gossler, A. and Kispert, A. (2004). The T-box transcription factor Tbx18 maintains the separation of anterior and posterior somite compartments. Genes Dev 18, 1209-1221.
20 Farin, H. F., Bussen, M., Schmidt, M. K., Singh, M. K., Schuster-Gossler, K. and Kispert, A. (2007). Transcriptional repression by the T-box proteins Tbx18 and Tbx15 depends on Groucho corepressors. J Biol Chem 282, 25748-25759.
21 Chen, G., Fernandez, J., Mische, S. and Courey, A. J. (1999). A functional interaction between the histone deacetylase Rpd3 and the corepressor groucho in Drosophila development. Genes Dev 13, 2218-2230.
22 Kaul, A., Schuster, E. and Jennings, B. H. (2014). The Groucho co-repressor is primarily recruited to local target sites in active chromatin to attenuate transcription. PLoS Genet 10, e1004595.
23 Kaul, A. K., Schuster, E. F. and Jennings, B. H. (2015). Recent insights into Groucho co-repressor recruitment and function. Transcription 6, 7-11.
24 Sekiya, T. and Zaret, K. S. (2007). Repression by Groucho/TLE/Grg proteins: genomic site recruitment generates compacted chromatin in vitro and impairs activator binding in vivo. Mol Cell 28, 291-303.
25 Farin, H. F., Mansouri, A., Petry, M. and Kispert, A. (2008). T-box protein Tbx18 interacts with the paired box protein Pax3 in the development of the paraxial mesoderm. J Biol Chem 283, 25372-25380.
26 Nie, X., Sun, J., Gordon, R. E., Cai, C. L. and Xu, P. X. (2010). SIX1 acts synergistically with TBX18 in mediating ureteral smooth muscle formation. Development 137, 755-765.
27 Rivera-Reyes, R., Kleppa, M. J. and Kispert, A. (2018). Proteomic analysis identifies transcriptional cofactors and homeobox transcription factors as TBX18 binding proteins. PLoS One 13, e0200964.
28 Pear, W. S., Nolan, G. P., Scott, M. L. and Baltimore, D. (1993). Production of high-titer helper-free retroviruses by transient transfection. Proc Natl Acad Sci U S A 90, 8392-8396.
29 Savic, D., Partridge, E. C., Newberry, K. M., Smith, S. B., Meadows, S. K., Roberts, B. S., Mackiewicz, M., Mendenhall, E. M. and Myers, R. M. (2015). CETCh-seq: CRISPR epitope tagging ChIP-seq of DNA-binding proteins. Genome Res 25, 1581-1589.
30 Chu, V. T., Graf, R., Wirtz, T., Weber, T., Favret, J., Li, X., Petsch, K., Tran, N. T., Sieweke, M. H., Berek, C., Kuhn, R. and Rajewsky, K. (2016). Efficient CRISPR-mediated mutagenesis in primary immune cells using CrispRGold and a C57BL/6 Cas9 transgenic mouse line. Proc Natl Acad Sci U S A 113, 12514-12519.
Bauer, D. E., Canver, M. C. and Orkin, S. H. (2015). Generation of genomic deletions in mammalian cell lines via CRISPR/Cas9. J Vis Exp, e52118.

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) Molecular cloning: a laboratory manual. Cold spring harbor laboratory press.

Wang, Z. (2009). Epitope tagging of endogenous proteins for genome-wide chromatin immunoprecipitation analysis. Methods Mol Biol 567, 87-98.

Jochim, N., Gerhard, R., Just, I. and Pich, A. (2011). Impact of clostridial glucosylating toxins on the proteome of colonic cells determined by isotope-coded protein labeling and LC-MALDI. Proteome Sci 9, 48.

Cox, J. and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol 26, 1367-1372.

Huang da, W., Sherman, B. T. and Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 4, 44-57.

Connaughton, D. M., Dai, R., Owen, D. J., Marquez, J., Mann, N., Graham-Paquin, A. L., Nakayama, M., Coyaund, E., Laurent, E. M. N., St-Germain, J. R., Blok, L. S., Vino, A., Klämbt, V., Deutsch, K., Wu, C. W. and Kolvenbach CM, K. F., Ottleswski I, Schneider R, Kitzler TM, Majmundar AJ, Buerguer F, Onuchic-Whitford AC, Youying M, Kolb A, Salmanullah D, Chen E, van der Ven AT, Rao J, Ityel H, Seltzsam S, Rieke JM, Chen J, Vivante A, Hwang DY, Kohl S, Dworschak GC, Hermle T, Alders M, Bartolomaeus T, Bauer SB, Baum MA, Brilstra EH, Chmallan TD, Zyskind J, Costin CE, Dipple KM, Duijkers FA, Ferguson M, Fitzpatrick DR, Fick R, Glass IA, Hulick PJ, Kline AD, Krey I, Kumar S, Lu W, Marco EJ, Wentzsen IM, Mefford HC, Platzer K, Povolotskaya IS, Savatt JM, Scherbakova NV, Senguttuvan P, Squire AE, Stein DR, Thiffault I, Voinova VY, Somers MJG, Ferguson MA, Traum AZ, Daouk GH, Daga A, Rodig NM, Terhal PA, van Binsbergen E, Eid LA, Tasic V, Rasouly HM, Lim TY, Ahram AG, Gharavi AG, Reuter HM, Rehm HL, MacArthur DG, Lek M, Laricchia KM, Lifton RP, Xu H, Mane SM, Sanna-Cherchi S, Sharrocks AD, Raught B, Fisher SE, Bouchard M, Khokha MK, Shril S, Hildebrandt F. (2020). Mutations of the Transcriptional Corepressor ZMYM2 Cause Syndromic Urinary Tract Malformations. Am J Hum Genet 107, 727-742.

Knappik, A. and Pluckthun, A. (1994). An improved affinity tag based on the FLAG peptide for the detection and purification of recombinant antibody fragments. Biotechniques 17, 754-761.

Tsai, A. and Carstens, R. P. (2006). An optimized protocol for protein purification in cultured mammalian cells using a tandem affinity purification approach. Nat Protoc 1, 2820-2827.

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A. and Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. Science 339, 819-823.
43 Wang, Y., Zhang, H., Chen, Y., Sun, Y., Yang, F., Yu, W., Liang, J., Sun, L., Yang, X., Shi, L., Li, R., Li, Y., Zhang, Y., Li, Q., Yi, X. and Shang, Y. (2009). LSD1 is a subunit of the NuRD complex and targets the metastasis programs in breast cancer. Cell 138, 660-672.

44 O'Brien, L. L., Guo, Q., Bahrami-Samani, E., Park, J. S., Hasso, S. M., Lee, Y. J., Fang, A., Kim, A. D., Guo, J., Hong, T. M., Peterson, K. A., Lozanoff, S., Raviram, R., Ren, B., Fogelgren, B., Smith, A. D., Valouev, A. and McMahon, A. P. (2018). Transcriptional regulatory control of mammalian nephron progenitors revealed by multi-factor cistromic analysis and genetic studies. PLoS Genet 14, e1007181.

45 Hakimi, M. A., Dong, Y., Lane, W. S., Speicher, D. W. and Shiekhattar, R. (2003). A candidate X-linked mental retardation gene is a component of a new family of histone deacetylase-containing complexes. J Biol Chem 278, 7234-7239.

46 Bantscheff, M., Hopf, C., Savitski, M. M., Dittmann, A., Grandi, P., Michon, A. M., Schlegl, J., Abraham, Y., Becher, I., Bergamin, G., Boesche, M., Delling, M., Dompelfeld, B., Eberhard, D., Huthmacher, C., Mathieson, T., Poeckel, D., Reader, V., Strunk, K., Sweetman, G., Kruse, U., Neubauer, G., Ramsden, N. G. and Drewes, G. (2011). Chemoproteomics profiling of HDAC inhibitors reveals selective targeting of HDAC complexes. Nat Biotechnol 29, 255-265.

47 Laurent, B., Ruitu, L., Murn, J., Hempel, K., Ferrao, R., Xiang, Y., Liu, S., Garcia, B. A., Wu, H., Wu, F., Steen, H. and Shi, Y. (2015). A specific LSD1/KDM1A isoform regulates neuronal differentiation through H3K9 demethylation. Mol Cell 57, 957-970.

48 Gocke, C. B. and Yu, H. (2008). ZNF198 stabilizes the LSD1-CoREST-HDAC1 complex on chromatin through its MYM-type zinc fingers. PLoS One 3, e3255.

49 Mellacheruvu, D., Wright, Z., Couzens, A. L., Lambert, J. P., St-Denis, N. A., Li, T., Miteva, Y. V., Hauri, S., Sardiou, M. E., Low, T. Y., Halim, V. A., Bagshaw, R. D., Hubner, N. C., Al-Hakim, A., Bouchard, A., Faubert, D., Fermin, D., Dunham, W. H., Goudreault, M., Lin, Z. Y., Badillo, B. G., Pawson, T., Durocher, D., Coulombe, B., Aebersold, R., Superti-Furga, G., Colinge, J., Heck, A. J., Choi, H., Gstaiger, M., Mohammed, S., Cristea, I. M., Bennett, K. L., Washburn, M. P., Raught, B., Ewing, R. M., Gingras, A. C. and Nesvizhskii, A. I. (2013). The CRAPome: a contaminant repository for affinity purification-mass spectrometry data. Nat Methods 10, 730-736.

50 Thul, P. J., Åkesson, L., Wiking, M., Mahdessian, D., Geladaki, A., Blal, H. A., Alm, T., Asplund, A., Björk, L., Breckels, L. M., Bäckström, A., Danielssson, F., Fagerberg, L., Fall, J., Gatto, L., Gnaann, C., Hober, S., Hjelmare, M., Johansson, F., Lee, S., Lindskog, C., Mulder, J., Mulvey, C. M., Nilsson, P., Oksvold, P., Rockberg, J., Schutten, R., Schenck, J. M., Silvertsson, Å., Sjöstedt, E., Skogs, M., Stadler, C., Sullivan, D. P., Tegel, H., Winsnes, C., Zhang, C., Zwahlen, M., Mardinoglu, A., Pontén, F., Feilizten, K. v., Lilley, K. S., Uhlén, M. and Lundberg, E. (2017). A subcellular map of the human proteome. Science 356, eaal3321.

51 Maiques-Diaz, A. and Somervaille, T. C. (2016). LSD1: biologic roles and therapeutic targeting. Epigenomics 8, 1103-1116.

52 Singh, M., Petry, M., Haenig, B., Lescher, B., Leitges, M. and Kispert, A. (2005). The T-box transcription factor Tbx15 is required for skeletal development. Mech Dev 122, 131-144.

53 Lausch, E., Hermanns, P., Farin, H. F., Al-Rashid, Y., Unger, S., Nikkel, S., Steinwender, C., Scherer, G., Spranger, J., Zabel, B., Kispert, A. and Superti-Furga, A. (2008). TBX15 mutations cause craniofacial dysmorphism, hypoplasia of scapula and pelvis, and short stature in Cousin syndrome. Am J Hum Genet 83, 649-655.
54 Stennard, F. A., Costa, M. W., Lai, D., Biben, C., Furtado, M. B., Solloway, M. J., McCulley, D. J., Leimena, C., Preis, J. I., Dunwoodie, S. L., Elliott, D. E., Prall, O. W., Black, B. L., Fatkin, D. and Harvey, R. P. (2005). Murine T-box transcription factor Tbx20 acts as a repressor during heart development, and is essential for adult heart integrity, function and adaptation. Development 132, 2451-2464.

55 Kirk, E. P., Sunde, M., Costa, M. W., Rankin, S. A., Wolstein, O., Castro, M. L., Butler, T. L., Hyun, C., Guo, G., Otway, R., Mackay, J. P., Waddell, L. B., Cole, A. D., Hayward, C., Keogh, A. P., M., Griffiths, L., Fatkin, D., Sholler, G. F., Zorn, A. M., Feneley, M. P., Winlaw, D. S. and Harvey, R. P. (2007). Mutations in cardiac T-box factor gene TBX20 are associated with diverse cardiac pathologies, including defects of septation and valvulogenesis and cardiomyopathy. Am J Hum Genet 81, 280-291.

56 Grossman, C. E., Qian, Y., Banki, K. and Perl, A. (2004). ZNF143 mediates basal and tissue-specific expression of human transaldolase. J Biol Chem 279, 12190-12205.

57 Thompson, J. J., Kaur, R., Sosa, C. P., Lee, J. H., Kashwagi, K., Zhou, D. and Robertson, K. D. (2018). ZBTB24 is a transcriptional regulator that coordinates with DNMT3B to control DNA methylation. Nucleic Acids Res 46, 10034-10051.

58 Peng, L., Chen, G., Zhu, Z., Shen, Z., Du, C., Zang, R., Su, Y., Xie, H., Li, H., Xu, X., Xia, Y. and Tang, W. (2017). Circular RNA ZNF609 functions as a competitive endogenous RNA to regulate AKT3 expression by sponging miR-150-5p in Hirschspring’s disease. Oncotarget 8, 808-818.

59 Armache, J. P., Jarasch, A., Anger, A. M., Villa, E., Becker, T., Bhushan, S., Jossinet, F., Habeck, M., Dindar, G., Franckenberg, S., Marquez, V., Mielke, T., Thomm, M., Berninghausen, O., Beatrix, B., Soding, J., Westhof, E., Wilson, D. N. and Beckmann, R. (2010). Localization of eukaryote-specific ribosomal proteins in a 5.5-A cryo-EM map of the 80S eukaryotic ribosome. Proc Natl Acad Sci U S A 107, 19754-19759.

60 Cheng, X. and Kao, H. Y. (2009). G protein pathway suppressor 2 (GPS2) is a transcriptional corepressor important for estrogen receptor alpha-mediated transcriptional regulation. J Biol Chem 284, 36395-36404.

61 Xu, Y., Sengupta, P. K., Seto, E. and Smith, B. D. (2006). Regulatory factor for X-box family proteins differentially interact with histone deacetylases to repress collagen alpha2(I) gene (COL1A2) expression. J Biol Chem 281, 9260-9270.
Figure legends

Figure 1. CRISPR/Cas9-mediated FLAG-tagging allows detection and localization of endogenous TBX18 protein in 293 and A549 cells. (A) Scheme of the CRISPR/Cas9 and homology mediated repair strategy to insert a triple FLAG peptide at the C-terminal end of TBX18. Red boxes indicate exon 8 of TBX18 on human chromosome 6 in which the stop codon is located. Thin dashed lines mark the homology region between the donor plasmid (white boxes) and the TBX18 locus before and after targeting to insert a cassette containing a triple FLAG tag, a P2A protein cleavage site, the neomycin resistance gene and the stop codon. (B) Western blot analysis of 293 parental and of 293-TBX18.3xFLAG heterozygous (293-P2C2) cells, and of A549 parental, A549-TBX18.3xFLAG heterozygous (A549-12C3) and homozygous (A549-12C4) cells with anti-FLAG antibodies detects a band slightly heavier (80-90kDa) than the calculated 68.7 kDa in the knock-in cells only (white arrowhead). ACTIN immunodetection serves as a loading control (black arrowhead). (C) Immunofluorescent staining of A549 and 293 parental and TBX18.3xFLAG knock-in cells with the anti-FLAG antibody detects a speckle-like pattern in the nucleus of TBX18.3xFLAG knock-in cells. Nuclei are counterstained with DAPI (blue).

Figure 2. LC-MS identifies novel protein binding partners of endogenous TBX18 protein in 293 and A549 cells. (A) Two lists showing 36 peptides statistically enriched in FLAG precipitates from A549-TBX18.3xFLAG cells (A549-12C4) and 10 proteins from 293-TBX18.3xFLAG cells (293-P2C2) compared to the unmodified parental lines. (B) Venn diagram depicting the overlap between the candidate protein lists for 293 (yellow) and A549 (blue) cells. (C) Manual literature mining identified proteins with transcriptional relevance; Zn finger proteins in red, ribosomal proteins in green, corepressors in blue. References: Gocke et al., 2008 [48]; Grossman et al., 2004 [56]; Thompson et al., 2018 [57]; Peng et al., 2017 [58]; Armache et al., 2010 [59]; Cheng and Kao, 2009 [60]; Xu et al., 2006 [61].

Figure 3. Endogenous TBX18 interacts with ZMYM2 in 293 cells. (A) Western blot analysis of endogenous ZMYM2 expression in extracts of HeLa, 293 and 293-TBX18.3xFLAG cells. The antibody against ZMYM2 detects a band of approximately 160 kDa size in 293-TBX18.3xFLAG and 293 parental control cells but not in HeLa
cells (arrow). **(B)** Immunofluorescent detection of endogenous TBX18.3xFLAG and ZMYM2 proteins with the indicated anti-FLAG and anti-ZMYM2 antibodies. Counterstaining of nuclei was performed with DAPI (blue). **(C)** Proximity ligation assay of endogenous FLAG-tagged TBX18 and ZMYM2 in 293-TBX18.3xFLAG knock-in cells using anti-FLAG and anti-ZMYM2 antibodies alone or in combination. Direct interaction is visualized by small red fluorescent dots. Nuclei are counterstained with DAPI (blue).

**Figure 4.** Overexpressed ZMYM2 directly binds to TBX18 in the nucleus of 293 cells. **(A)** Western blot of co-immunoprecipitation of TBX18.MYC and ZMYM2.MYC-FLAG via α-FLAG IP visualized by α-MYC. Input control shows presence of protein in cell lysates of 293 cells transfected with either TBX18 or ZMYM2 expression vectors. After co-incubation of both lysates, TBX18.MYC co-immunoprecipitates with ZMYM2.MYC-FLAG after α-FLAG IP, but not with negative IgG-control antibodies. **(B)** Nuclear recruitment assay in 293 cells. Cells were transfected with an expression construct for MYC-FLAG-tagged ZMYM2 or MYC-tagged TBX18 in the presence of HA-tagged TBX18 lacking the nuclear localization signal (TBX18ΔNLS.HA). Immunofluorescence analysis shows that NLS-deficient TBX18 protein (red) is efficiently shuttled from the cytoplasm to the nucleus by overexpressed full-length TBX18 or ZMYM2 (green). Nuclei are counterstained with DAPI (blue). **(C)** Statistical evaluation of the nuclear recruitment assay for 293 cells with overexpression of TBX18.MYC (first panel), ZMYM2.MYC-FLAG (second panel), TBX18.ΔNLS.HA alone or in combination with TBX18.MYC (positive control) and ZMYM2.MYC.FLAG (third panel). Shown is the percentage of cells which express the antigen (anti-MYC detection, first and second panel; anti-HA detection, third panel) in the nucleus (black bar) or in the cytosol (grey bar). Since the antigen can be found both in the nucleus and in the cytosol in some cells, the % do not add up to 100. Error bars show SD. *** indicates very high significance p<0.001 in Student’s t-test. **(D)** PLA in 293 cells transfected with over-expression vectors for TBX18.MYC, TBX18.HA and ZMYM2.MYC-FLAG or with combinations (TBX18.MYC with TBX18.HA, TBX18.HA with ZMYM2-MYC-FLAG) and probed with anti-HA and anti-MYC antibodies. Strong nuclear PLA signals were found in cells coexpressing TBX18.MYC and TBX18.HA (positive control) and in cells coexpressing TBX18.HA and ZMYM2.MYC-FLAG. Nuclei are counterstained with DAPI (blue).
Figure 5. ZMYM2 interacts with the T-box region of TBX18. (A) Scheme of TBX18 full-length and TBX18 truncated proteins used for the GST pull-down assay in (C). Indicated are the engrailed homology (eh1)-domain, the nuclear localization signal (NLS), and the T-box. (B) GST and fusion proteins of GST with N+T-, T-, N- and C-domains of TBX18 were purified from E. coli extracts and analyzed for integrity and quantity by Coomassie Brilliant Blue staining of SDS-polyacrylamide gels. Arrows indicate GST and GST-fusion proteins. (C) Autoradiographic analysis of pull-down assays performed with GST and fusion proteins of GST with N-terminal, N-terminal+T-box- (NT), T-box- (T), and C-terminal (C) regions of TBX18 obtained from E.coli extracts, and reticulocyte lysates programmed for in vitro translation of 35S-labelled tagged full-length ZMYM2. ZMYM2 preferentially interacts with the T-box region of TBX18. Arrows point to ZMYM2 at expected apparent size of approx. 154.9 kDa in 10% input control gel (C, left) and GST-pulldown (C, right).

Figure 6. Interaction of CAKUT mutant forms of ZMYM2 with TBX18 is partly compromised. (A) Comparative RNA in situ hybridization analysis on transverse sections of the proximal ureter of E12.5 wildtype mouse embryos reveals co-expression of Tbx18 and Zmym2 in the mesenchymal compartment. Note that expression of Tbx18 is confined to the ureteric mesenchyme (um) whereas Zmym2 expression is additionally found in the inner tissue compartment, the ureteric epithelium (ue). Size bar is 25 µm. (B) Domain structure of wildtype ZMYM2 and of truncated versions found in CAKUT patients. The localization of known protein domains of ZMYM2 is indicated, the name of each variant is given according to nomenclature recommended by the Human Genome Variation Society. The prefix p denotes protein reference followed by three letter code for the reference amino acid at the position of the mutation and a number for its position. Asterisk indicates nonsense substitution of a termination site (e.g. p.Arg540*). Addition of another three letter code indicates an amino acid exchange, fs codes for frameshift. A number following an asterisk indicates distance to new termination site from reference position in number of amino acids. Descriptions in brackets are predictions from DNA sequence data (e.g. p.(Lys812Aspfs*18)). The mutants represent a series of C-terminal deletions. (C) Immunofluorescent detection of TBX18 and ZMYM2 wildtype and mutant proteins after transfection of expression constructs into 293 cells.
ZMYM2 proteins were detected by their MYC-tag, TBX18 proteins by anti-HA immunofluorescence. Transfections and stainings are as indicated in the figure. ZMYM2 and truncated protein versions are shown in red immunofluorescence (anti-MYC), TBX18 full-length (TBX18FL) or ΔNLS (TBX18ΔNLS) proteins are shown in green (anti-HA). Upper rows show overlay pictures of DAPI nuclear stain (blue), ZMYM2 variant (red) and TBX18 (green), lower rows show single channel of recruited protein variant. Left images present single transfections to evaluate cellular localization of ZMYM2 variants, right images present co-transfections with TBX18 to examine nuclear recruitment of mutant ZMYM2 proteins. Wildtype ZMYM2 protein is localized mainly nuclear and recruits TBX18ΔNLS.HA to the nucleus. Mutant forms of ZMYM2 are almost completely cytosolic except p.Tyr763* that localized to the nucleus. TBX18FL protein was co-transfected to examine recruitment of cytosolic ZMYM2 variants into the nucleus. TBX18ΔNLS was used to test recruitment into the nucleus by the p.Tyr763* mutant ZMYM2. (D) Statistical analysis of ZMYM2 protein localization shown in C. Error bars show SD. (E) Statistical analysis of ZMYM2 mutant protein localization (upper row) and of TBX18ΔNLS.HA (lower row) in nuclear recruitment assays shown in C. Error bars show SD. Significance indicated as *, p<0.05 and ***, p<0.001 as determined in Student’s t-test. Note that protein expression can be found both in the nuclear and cytosolic compartment. Therefore, addition of nuclear and cytosolic localization can add to more than 100% in D and E. (F) Summary table of the TBX18/ZMYM2 nuclear recruitment assay.

**Figure 7.** ZMYM2 and KDM1 exert corepressor activity on TBX18 reporter constructs. Dual luciferase assays were performed with extracts of 293 cells cotransfected with constructs for the Tbx18 luciferase reporter and various combinations of expression constructs for full-length TBX18 with TLE3/GROUCHO (positive control), ZMYM2 and KDM1A as indicated. Numbers refer to ng of plasmid used for one transfection. Statistical analysis was performed on four biological replicates. Error bars show SD. Significance indicated as *, p<0.05 and ***, p<0.001 as determined in Student’s t-test. (A) TBX18 represses the luciferase reporter activity significantly by around 40%. TLE/GROUCHO and ZMYM2 enhance TBX18 repressive activity individually but not combinatorially. (B) ZMYM2 does not affect the activity of the luciferase reporter in absence of TBX18 protein. (C) KDM1A enhances the TBX18 repressive activity in a concentration dependent manner. Cotransfection
of a ZMYM2 expression construct relieves the co-repressive activity of KDM1A. (D) KDM1A does not affect reporter activity alone or in combination with ZMYM2 in absence of TBX18 protein.
Figure A: Schematic representation of the TBX18 locus with the Protospacer sequence (CCAAATGTCATTAACTTAAAGG) and the structure of the donor plasmid. The diagram illustrates the ORF, stop codon, 3′UTR, and exons (Exon 8) of the TBX18 gene.

Figure B: Western blot analysis showing the expression levels of TBX18 and ACTIN in different cell lines.

Figure C: Immunofluorescence images of different cell lines stained with anti-FLAG antibody and DAPI. The scale bar represents 25 μm.
### Table A

| Gene symbol | Protein name | Transcriptional relevance | Reference |
|-------------|--------------|---------------------------|-----------|
| TBX18       | T-box transcription factor TBX18 |                         |           |
| ZMYM2       | Zinc finger MYM-type protein 2 | Zn-finger DNA-binding TF | Gocke et al., 2008 |
| ZMYM3       | Zinc finger MYM-type protein 3 | Zn-finger DNA-binding TF | Gocke et al., 2008 |
| ZNF143      | Zinc finger protein 143 | Zn-finger DNA-binding TF | Grossman et al., 2004 |
| ZBTB24; PATZ2 | Zinc finger and BTB domain-containing protein 24 | Zn-finger DNA-binding TF | Thompson et al., 2018 |

### Table B

![Venn Diagram](https://via.placeholder.com/150)

### Diagram C

Gene symbol | Protein name | Reference |
------------|--------------|-----------|
TBX18       | T-box transcription factor TBX18 |           |
ZMYM2       | Zinc finger MYM-type protein 2 | Gocke et al., 2008 |
ZMYM3       | Zinc finger MYM-type protein 3 | Gocke et al., 2008 |
ZNF143      | Zinc finger protein 143 | Grossman et al., 2004 |
ZBTB24; PATZ2 | Zinc finger and BTB domain-containing protein 24 | Thompson et al., 2018 |
ZNF609      | Zinc finger protein 609 | Peng et al., 2017 |
RPL15       | 60S ribosomal protein L15 | Armache et al., 2010 |
NCOR2       | Nuclear receptor corepressor 2 | Cheng & Kao, 2009 |
RPL4        | 60S ribosomal protein L4 | Armache et al., 2010 |
RPL18       | 60S ribosomal protein L18 | Armache et al., 2010 |
RFX1        | MHC class II regulatory factor RFX1 | Xu et al., 2006 |
RPL14       | 60S ribosomal protein L14 | Armache et al., 2010 |
RRBP1       | Ribosome-binding protein 1 | Armache et al., 2010 |
