An ancient transcription factor functions as the master regulator of primary cilia formation

Weihua Wang[1,#,†], Xiqi Zhang[1,#], Yaxuan Qiu[2,#], Xiangrui Meng[2,#], Sitong Cheng[2], Yutong Chen[1], Siqi Liu[2], Wenhui Chen[1], Jiayan Yi[2], Xiwen You[2], Hongni Liu[1], Junqiao Xing[1], Cheng Xu[1], Haochen Jiang[1], Haibo Wang[1,3], Guangmei Tian[1,†], Zhangfeng Hu[1,4,5,6,†]

[1]Institute of Microalgae Synthetic Biology and Green Manufacturing, School of Life Sciences, Jianghan University, Wuhan, 430056, China

[2]School of Medicine, Jianghan University, Wuhan, 430056, China

[3]Cancer Institute (Key Laboratory of Cancer Prevention and Intervention, China National Ministry of Education) of the Second Affiliated Hospital and Institute of Translational Medicine, Zhejiang University School of Medicine, Hangzhou, 310029, China

[4]Hubei Key Laboratory of Environmental and Health Effects of Persistent Toxic Substances, Jianghan University, Wuhan, 430056, China

[5]Hubei Engineering Research Center for Protection and Utilization of Special Biological Resources in the Hanjiang River Basin, School of Life Sciences, Jianghan University, Wuhan, 430056, China

[6]China-Cuba Joint Natural Products Research Laboratory, Jianghan University, Wuhan, 430056, China

[#]These authors contributed equally: Weihua Wang, Xiqi Zhang, Yaxuan Qiu, Xiangrui Meng

[†]Correspondence: wangwh2019@163.com (W.W.); tiangm163@163.com (G.T.); huzhangfenghzf@163.com (Z.H.)

Abstract

How ancient transcription factors are repurposed during evolution to drive the functional diversification of conserved organelles remains a fundamental question in biology. Although highly conserved in basic morphology, eukaryotic cilia vary extensively in their sizes and functions. Previously, we showed that an evolutionarily ancient transcription factor, X chromosome–associated protein 5 (Xap5), controls motile ciliogenic transcriptional programs during mouse spermatogenesis (Wang et al., 2025). Here we show that Xap5 functions as the master regulator of primary ciliogenesis. Xap5 associates with the nuclear protein Non-POU domain–containing octamer-binding protein (Nono) to form a regulatory module that directly binds and activates the transcription factors Sox5 and Sox9. Genetic ablation of *Xap5* or *Nono* impairs primary ciliogenesis, and loss of Sox5 disrupts the downstream ciliogenic program, consistent with the established role of Sox9. Collectively, our results provide new insight into how complex ciliary transcription factor networks determine ciliary diversity during evolution, and suggest that defects in this regulatory axis may contribute to the etiology of human ciliopathies.

Introduction

A central challenge in evolutionary biology is to understand how a limited number of ancient conserved genes generate tremendous cellular diversity through functional redeployment[1-3]. This challenge is particularly evident in the roles of master transcription factors, which must be precisely regulated across different cellular contexts to control specific developmental programs[4-7]. Such functional plasticity is thought to arise through evolutionary co-option, the process by which existing regulators are recruited into new regulatory modules to drive lineage- or organelle-specific innovations[8-10]. Elucidating the molecular strategies underlying this repurposing is essential for understanding how complex biological systems evolve. The biogenesis of the cilium—an evolutionarily ancient organelle that plays key roles in signaling and motility[11-14], provides an important model for investigating this fundamental question, and its dysfunction can lead to a class of genetic disorders known as ciliopathies[15].

The assembly of the cilium, or ciliogenesis, requires the precise transcriptional coordination of hundreds of genes[16-18]. In metazoans, the regulatory factor X (Rfx) and Forkhead box protein J1 (Foxj1) transcription factor families have been identified as key regulators of ciliogenesis[19-22]. Rfx factors, such as DAF-19 in *C. elegans*, regulate the expression of genes essential for building the core ciliary structure[23], while Foxj1 directs the motile ciliogenic program[24,25]. However, this well-established regulatory network primarily applies to motile, multiciliated cells. How the upstream, cell-type-specific transcriptional programs are initiated, particularly for the non-motile primary cilium found on most somatic cells, remains poorly understood[12,26].

The conserved transcription factor Xap5 serves as a prime example of this challenge, presenting a compelling biological paradox. Our previous work showed that Xap5 is a critical

regulator for the assembly of motile appendages—dynamic structures that generate movement by beating in fluid environments—a role conserved from the cilia of the unicellular alga *Chlamydomonas* to the flagella of mammalian sperm[27,28]. These findings cemented Xap5's role in an ancient program for motility but, in doing so, created a profound biological enigma: why is this regulator, seemingly dedicated to motility, highly expressed in somatic cells that exclusively form non-motile primary cilia for sensory functions? Whether Xap5 plays a role in primary ciliogenesis in these cells, and the mechanism by which it might be co-opted for this distinct function, remains to be determined[29].

In this study, we specifically address these questions by characterizing the function of Xap5 in primary ciliogenesis using NIH/3T3 mouse fibroblasts, a standard model for primary cilium formation upon serum starvation[30]. Here, we identify the Non-POU domain-containing octamer-binding protein (Nono) as a critical interacting partner of Xap5. Our integrated analyses demonstrate that Xap5 recruits its cofactor Nono to regulate primary cilia formation by modulating a downstream transcriptional axis involving Sox5 and Sox9. While Sox9 has been previously implicated in primary ciliogenesis[31-33], our findings for the first time identify Xap5 as a critical upstream activator of these transcription factors in somatic cells. Together, these findings define the complete Xap5-Nono-Sox5/9 axis for primary ciliogenesis. Moreover, they offer a powerful illustration of evolutionary co-option in action, establishing a mechanistic framework for how ancient transcription factors are functionally repurposed to drive cellular innovation.

Results

Xap5 is essential for primary cilia formation

To investigate the role of Xap5 in primary cilia formation, we first determined its subcellular localization in NIH/3T3 cells. Immunofluorescence staining with a specific anti-Xap5 antibody revealed that endogenous Xap5 is predominantly nuclear (Fig. 1a), consistent with our previous work. We confirmed this finding via ectopic expression of a Flag-tagged Xap5 construct (Fig. 1b). Consistent with its nuclear localization, bioinformatic analysis predicted a putative nuclear localization signal (NLS) within the Xap5 protein (Supplementary Fig. 1a). To functionally validate this sequence, we generated a site-directed mutant (Flag-Xap5-ΔNLS) (Supplementary Fig. 1a). Crucially, this mutation abolished the nuclear import of Xap5, leading to its mislocalization in the cytoplasm and confirming that this NLS is essential for its function (Fig. 1b).

To investigate the function of Xap5, we generated *Xap5* knockout (KO) NIH/3T3 cell lines using the CRISPR/Cas9 system[34] (Supplementary Fig. 1b). The complete depletion of Xap5 protein product was confirmed by western blotting (Fig. 1c). Upon inducing ciliogenesis via serum starvation (Fig. 1d), we first observed a striking defect in ciliary length (Fig. 1e). Although comparable to wild-type (WT) cells before starvation, *Xap5* KO cells exhibited a significant shortening of their primary cilia post-starvation (Fig. 1e, f). To determine if this length defect stemmed from an overall failure in ciliogenesis, we next quantified the ciliation frequency. Interestingly, both WT and *Xap5* KO cells responded robustly to the starvation stimulus, showing a comparable increase in the percentage of ciliated cells over 48 h (Fig. 1g). These results collectively demonstrate that Xap5 is dispensable for the initiation of ciliogenesis but plays an essential role in the subsequent elongation and/or maintenance of the ciliary axoneme.

Xap5 forms a nuclear regulatory module with its cofactor Nono

Having established the essential role of Xap5 in primary ciliogenesis, we next sought to uncover its underlying regulatory mechanism by identifying its synergistic cofactors. An unbiased immunoprecipitation-mass spectrometry (IP-MS) screen for Xap5 interactors identified several nuclear proteins (Supplementary Fig. 2), among which we prioritized Non-POU domain-containing octamer-binding protein (Nono), given its well-established, pleiotropic roles in transcriptional regulation and RNA processing[35-37].

As a prerequisite for their interaction, we first confirmed that Nono is also a predominantly nuclear protein in NIH/3T3 cells, a localization we validated for both the endogenous protein via immunofluorescence and an ectopically expressed Nono-GFP fusion protein (Fig. 2a). To biochemically validate their physical association, we next performed endogenous co-immunoprecipitation (Co-IP) assays. These results confirmed a robust and specific interaction, as immunoprecipitation of Xap5 specifically co-precipitated Nono (Fig. 2b). Taken together, these localization and co-immunoprecipitation data provide strong evidence that Xap5 and Nono form a bona fide regulatory complex within the nucleus.

The Xap5-Nono module is functionally required for ciliogenesis

Given that Nono forms a complex with the essential ciliogenesis factor Xap5, we hypothesized that the entire module is functionally indispensable. To test this, we generated *Nono* KO NIH/3T3 cells using the CRISPR/Cas9 system (Supplementary Fig. 3a). Indeed, upon serum starvation, *Nono* KO cells phenocopied the *Xap5* KO. They exhibited significantly shorter primary cilia compared to wild-type (WT) controls, yet displayed a comparable increase in ciliation frequency (Fig. 3b-d). Notably, in addition to the ciliary defects, we observed a significant increase in nuclear size in *Nono* KO cells (Fig. 3b; Supplementary Fig. 3b). Together, these findings demonstrate that Nono, like its partner Xap5, is dispensable for ciliary initiation

but essential for ciliary elongation and maintenance, and may also have a broader role in maintaining nuclear architecture.

The Xap5-Nono module co-regulates a downstream ciliogenic transcriptional cascade

To determine if the Xap5-Nono module cooperates at the transcriptional level, we performed RNA-sequencing (RNA-seq) on WT, *Xap5* KO, and *Nono* KO cells following 24 h of serum starvation, a time point where the phenotype was robustly established. This analysis revealed widespread transcriptional changes; *Xap5* KO cells showed 1,116 upregulated and 1,071 downregulated genes, while *Nono* KO cells had 1,425 upregulated and 1,042 downregulated genes relative to WT (fold change ≥ 1.5; Supplementary Fig. 4a, b; Supplementary Data 1, 2).

Given that loss of either Xap5 or Nono impairs ciliogenesis, we hypothesized that they function together as transcriptional activators of a shared ciliogenic program. We therefore focused our subsequent analysis on the genes commonly downregulated in both knockouts. Notably, Gene Ontology (GO) analysis of these genes revealed a highly significant enrichment for ciliary categories (Fig. 4a, b; Supplementary Fig. 4c, d). To investigate this further, we compared this downregulated gene set against a comprehensive ciliary gene list (Supplementary Data 3), which was assembled from the SYSCILIA gold standard (SCGSv2)[18] and manual literature curation. This comparison confirmed a significant overlap between the two gene sets (Fig. 4c). We validated several of these shared targets—including the transcription factors *Sox5* and *Sox9*, and the ciliary component *Pkd2*—by qPCR (Fig. 4d, e).

Among this cohort of downregulated genes, two transcription factors, Sox5 and Sox9, stood out. Sox9 is a known regulator of primary cilia in other contexts, and both factors are known to cooperate in developmental programs. This made them compelling candidates for further investigation as potential downstream effectors in our system. Consistent with this hypothesis,

western blot analysis confirmed that the protein levels of both Sox5 and Sox9 were markedly reduced in *Xap5* and *Nono* KO cells (Fig. 4f). These data strongly indicate that Xap5 and Nono co-regulate a transcriptional network that includes the downstream effectors Sox5 and Sox9, thereby driving primary cilia formation.

The Xap5-Nono module directly co-occupies the regulatory regions of Sox5 and Sox9

To determine if Xap5 and Nono directly regulate their target genes, we mapped their genome-wide chromatin occupancy using CUT&Tag[38] in NIH/3T3 cells (Supplementary Data 4, 5). This analysis revealed a strong enrichment of both factors at promoter regions; approximately 30% of Xap5 peaks and a striking 50% of Nono peaks were located at promoters (Supplementary Fig. 5a, b). Consistent with this, genomic heatmaps demonstrated that both Xap5 and Nono binding signals were highly enriched around the transcription start sites of their target genes (Fig. 5a, b; Supplementary Fig. 5c, d).

Integrating these binding profiles with our RNA-seq data revealed a strong correlation between genes bound by Xap5 or Nono and those whose expression was altered upon KO (Fig. 5c, d), supporting a model of direct transcriptional regulation. De novo motif analysis revealed that Xap5 and Nono recognize a highly similar DNA-binding motif (Fig. 5e). Importantly, this Xap5 motif is virtually identical to that from our previous work in testicular germ cells (Fig. 5e), demonstrating that Xap5 maintains its DNA-binding specificity across distinct cellular contexts.

Crucially, we found a significant overlap between the Xap5/Nono co-occupied genes and our curated ciliary gene set (Fig. 5f). Among these, we observed direct co-occupancy of Xap5 and Nono at the promoter regions of the key downstream regulators Sox5 and Sox9, a finding we validated by CUT&Tag-qPCR (Fig. 5g). Collectively, these genomic data provide strong evidence that Xap5 and Nono form a complex that directly binds to the regulatory regions of a

common set of ciliary genes, including Sox5 and Sox9, to orchestrate primary cilia formation.

Sox5, a key downstream effector of the Xap5-Nono pathway, is essential for ciliogenesis

Our finding that Xap5 and Nono directly co-regulate Sox5 and Sox9 prompted us to investigate their respective roles in ciliogenesis. Given that the essential role of Sox9 in maintaining primary cilia length is well-documented[31,33], we focused our functional analysis on the uncharacterized role of Sox5. To this end, we generated *Sox5* KO NIH/3T3 cells using the CRISPR/Cas9 system (Fig. 6a; Supplementary Fig. 6).

Analysis after serum starvation revealed that *Sox5* KO cells precisely phenocopied the upstream knockouts. *Sox5* KO cells exhibited significantly shorter primary cilia than WT controls (Fig. 6b, c), yet this occurred despite a normal ciliation response, with both genotypes showing a robust and comparable increase in ciliation frequency (Fig. 6d). These results establish Sox5 as a critical effector responsible for ciliary elongation. Collectively, our data on Sox5, combined with the established literature on Sox9, strongly support a model where the Xap5-Nono complex acts through two essential downstream transcription factors, Sox5 and Sox9, to orchestrate the maintenance of primary cilia.

Sox5 executes a shared ciliogenic program downstream of the Xap5-Nono complex

To confirm that Sox5 functions transcriptionally downstream of the Xap5-Nono complex, we performed RNA-seq on *Sox5* KO and WT cells after 24 h of serum starvation. The analysis identified 2,124 upregulated and 2,105 downregulated genes in *Sox5* KO cells (fold change ≥ 1.5; Supplementary Fig. 7a; Supplementary Data 6). To assess the global impact on ciliary pathways, we first performed Gene Set Enrichment Analysis (GSEA)[39] on all ranked differentially expressed genes, which revealed a significant enrichment of the “Cilium organization” pathway (Fig. 7a). Consistent with this, GO analysis of the strictly defined

downregulated gene set (fold change ≥ 1.5) also highlighted terms related to cilium assembly (Fig. 7b; Supplementary Fig. 7b). Furthermore, direct comparison showed a significant overlap between the full set of genes dysregulated by Sox5 loss and our curated ciliary gene list (Supplementary Fig. 7c), reinforcing the strong link between Sox5 and ciliogenesis.

Critically, to place Sox5 within the Xap5-Nono pathway, we compared the downregulated gene sets across all three KO lines. A Venn diagram revealed a significant and substantial overlap among them (Fig. 7c). We confirmed this shared transcriptional dependency by qPCR, validating that several key ciliary genes—including *Adcy3*, *Ccdc39*, and *Pkd2*—were indeed downregulated across all three KO models (Fig. 7d). These results provide strong evidence that Sox5 acts as a key transcriptional mediator downstream of Xap5 and Nono, executing a common program of ciliary gene expression required for proper cilia elongation.

Discussion

The assembly of the cilium, a highly conserved organelle, requires a sophisticated and hierarchical transcriptional program that is responsive to extracellular cues and precisely adapted to generate enormous functional diversity[13,19-22,26]. The transcriptional program's complexity is evident in its role in regulating flagella in unicellular organisms like *Chlamydomonas reinhardtii* and orchestrating the synchronous formation of hundreds of motile cilia in the mammalian airway[14,40-43]. In mammals, a core ciliogenic program is controlled by the Rfx family, which act as master architects for the core ciliary structure[20,44-46]. Built upon this foundation are multi-tiered regulatory cascades, often initiated by morphogenetic signaling pathways, that deploy cell-type-specific "selector cassettes"[14,20]. For the highly specialized motile cilia, this cascade involves key initiators like GMNC/MCIDAS (antagonized by GMNN)[47-50], integrates context-

specific factors like NOTO[51,52], and converges on downstream effectors TAp73/MYB[53-55], which in turn activate the master regulator Foxj1 to execute the motile program[24,25]. While this elegant hierarchy explains the formation of multiciliated cells, how upstream signaling pathways initiate the transcriptional programs for the vast diversity of primary cilia remains a key unanswered question.

Central to this complexity is the ancient, evolutionarily conserved transcription factor Xap5. Xap5 regulates flagellar assembly in *Chlamydomonas reinhardtii*, and is widely expressed in mammalian tissues[27,56,57]. Our recent work revealed that during motile ciliogenesis in spermatogenesis, Xap5 functions upstream of Rfx2 and Foxj1 to drive the sperm flagellar program[28]. This finding solidified Xap5's role as a key regulator for motility but also sharpened the central paradox addressed in this study: how does this same factor, which utilizes the established motile cilia pathway in germ cells, function in somatic cells that build exclusively non-motile primary cilia?

In this study, we resolve this fundamental paradox by uncovering a regulatory strategy that enables Xap5's functional switch. In our somatic cell model, Xap5 is repurposed through recruitment of the nuclear protein Nono, forming an Xap5–Nono regulatory module that promotes primary ciliogenesis by directly engaging downstream Sox transcriptional effectors. This work provides a molecular blueprint for how a deeply conserved master regulator can achieve functional diversity through context-dependent cofactor assembly, while ultimately converging on core organelle biogenesis programs.

Specifically, we find that the Xap5–Nono module directly co-regulates Sox5 and Sox9, positioning them as downstream transcriptional effectors in primary ciliogenesis. The essential role of Sox9 in promoting primary cilia formation is well-established in multiple developmental

contexts[31-33]. Significantly, the finding that Xap5-Nono directly co-regulates Sox9 and Sox5 nominates Sox5 as a previously unrecognized, critical player in primary ciliogenesis. This organization exemplifies the "selector cassette" concept[20], in which context-specific deployment of transcription factors modifies the execution of a conserved program; we therefore propose that the context-dependent deployment of Sox5 and Sox9 downstream of Xap5–Nono functions as such a cassette. This model is supported by evidence that Sox5 can cooperate with Sox9 in other developmental programs (e.g., chondrogenesis[58]) and that Sox5 has been linked to ciliogenic gene regulation in motile contexts[59,60]. Together, these findings position Xap5–Nono as an upstream integrator that co-opts two related Sox factors to regulate primary ciliogenesis.

Our results suggest an evolutionary framework in which the DNA-recognition properties of Xap5 function as a conserved genomic anchor, while biological outputs diversify through context-dependent regulatory logic. Intriguingly, the GC-rich Xap5 binding motif identified here is nearly reverse-complementary to the CT-rich C-box core (5'-CCTCC-3') that functions as a DAF-19/RFX-dependent enhancer in *C. elegans* ciliome promoters[23,45,61]. Although direct binding of Xap5 family members to C-box elements remains to be demonstrated, this sequence-level complementarity, together with the deep conservation of the Xap5 family[27,56,57], supports the hypothesis that Xap5 could contribute to C-box–like regulatory architectures in distinct lineages. If so, it would be consistent with a broader model of evolutionary co-option[1-3,8-10], in which the core DNA-binding interface of Xap5 remains relatively constrained while regulatory outcomes may be reprogrammed via partner switching—for example, potential integration into RFX-driven X-box/C-box promoter modules in nematodes[23,45,61], versus association with Nono in mammalian somatic cells. This "rigid anchor, flexible cofactor" model provides a plausible route to expand regulatory complexity while preserving constraints imposed by DNA-binding

specificity.

Collectively, the discovery of this adaptable, Xap5-driven regulatory network offers new perspectives on human ciliopathies[15]. Given that Sox9 dysfunction is linked to cystic diseases with ciliary defects[31,32], disruption of the upstream Xap5–Nono module represents a plausible, yet unexplored, pathogenic mechanism. As upstream regulators of primary ciliogenesis continue to emerge—including the metazoan-specific SP5/SP8 factors[62,63]—our work highlights a distinct regulatory layer rooted in an evolutionarily ancient transcription factor[27,56,57]. Notably, systemic loss of Xap5 causes embryonic lethality prior to E7.5[28], indicating essential functions that likely extend beyond ciliary regulation alone. One possibility is that Xap5 establishes a permissive transcriptional state that enables (or constrains) the execution of ciliogenic gene programs, upon which later-evolved, developmental regulators such as SP5/SP8 impose tissue- and stage-specific logic[62]. Future work should define the upstream cues that govern Xap5–Nono assembly and test how alternative cofactors reprogram Xap5 outputs across tissues. Mapping the context-dependent partner repertoire of Xap5 will be key to explaining how an ancient regulator is repurposed to drive distinct transcriptional outputs across cell types, with direct implications for development and ciliopathy mechanisms.

Materials and Methods

Antibodies

The following primary antibodies were used in this study: from ABclonal Biotechnology, anti-Flag (IF 1:1,000; AE005), anti-Gapdh (IB 1:10,000; AC002), anti-Nono (IB 1:1,000, IF 1:500; A3800), and anti-Sox5 (IB 1:1,000; A23127); from Sigma-Aldrich, anti-Acetyl-α-Tubulin (IF 1:500; T7451); from Proteintech, anti-Arl13b (IF 1:500; 17711-1-AP); from HUABIO, anti-Sox9

(IB 1:1,000; ET1611-56); and from Prospertech, anti-Xap5 (IB 1:1,000, IF 1:500; HU-412003).

The following secondary antibodies were used: Alexa Fluor 488- and 594-conjugated goat anti-mouse/rabbit IgG (IF 1:500; ThermoFisher Scientific); HRP-conjugated goat anti-mouse IgG (IB 1:2,000; Bio-Rad, #1706516) and HRP-conjugated goat anti-rabbit IgG (IB 1:2,000; Bio-Rad, #1706515).

Cell culture

The NIH/3T3 mouse embryonic fibroblast cell line (ATCC, CRL-1658) was cultured in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% calf serum (CS) and 1% penicillin-streptomycin (all from Life Technologies). Cells were maintained in a humidified incubator at 37°C and 5% CO_2.

Plasmid construction, transfection and generation of knockout cell lines

For protein expression, the coding sequences of mouse Xap5 and Nono were amplified from NIH/3T3 cDNA and cloned into the pECMV-MCS-Flag and pCMV-C-GFP vectors, respectively, to generate constructs encoding C-terminally tagged proteins. Site-directed mutagenesis to delete the Xap5 nuclear localization signal (NLS) was performed using the Mut Express II Fast Mutagenesis Kit V2 (Vazyme, C214). For CRISPR/Cas9-mediated gene knockout, the all-in-one plasmid pU6-sgRosa26-1_CBh-Cas9-T2A-BFP (a gift from Ralf Kuehn; Addgene plasmid #64216) was used as the backbone. Gene-specific sgRNA sequences were inserted by replacing the original sgRosa26-1 guide RNA sequence via site-directed mutagenesis. The integrity of all constructs was confirmed by Sanger sequencing. All primer and sgRNA sequences are listed in Supplementary Tables 1 and 2, respectively.

For transient expression experiments, cells were transfected with the indicated DNA constructs using Lipofectamine 3000 (Life Technologies) according to the manufacturer's

protocol.

Stable knockout cell lines were generated by electroporation. Briefly, 1×10^6 NIH/3T3 cells were resuspended in 250 µL of Opti-MEM I Reduced Serum Medium (Gibco) containing 20 µg of the appropriate all-in-one CRISPR-Cas9 plasmids. The cell suspension was transferred to a 0.4 cm gap electroporation cuvette (Bio-Rad) and electroporated using a Gene Pulser Xcell Electroporation System (Bio-Rad) with a single exponential decay pulse (250 V, 950 µF, ∞ Ω resistance). Twenty-four hours post-electroporation, single cells exhibiting bright BFP fluorescence were identified by fluorescence microscopy and manually isolated. Individual cells were seeded into separate wells of a 96-well plate to grow into monoclonal colonies, which were subsequently expanded and validated for successful gene knockout.

Immunofluorescence and microscopy

To induce primary cilia formation, NIH/3T3 cells were seeded at low density (< 30% confluency) onto sterile, poly-D-lysine-coated (Gibco) glass coverslips (CITOTEST). Following 24 h of growth in standard medium, ciliogenesis was induced by switching to serum-free DMEM for an additional 24–48 h.

Cells on coverslips were washed three times with PBS and fixed with 4% paraformaldehyde (PFA) in PBS for 10 min at room temperature. Following fixation, residual PFA was quenched with 50 mM NH_4Cl in PBS for 5 min. Cells were then permeabilized with 0.2% Triton X-100 for 15 min and blocked with 3% bovine serum albumin (BSA) in PBS for 1 h. Primary antibody incubations were performed for 1 h at room temperature in a humidified chamber. After washing, cells were incubated with Alexa Fluor-conjugated secondary antibodies for 1 h, followed by a 5 min counterstain with DAPI (Sigma-Aldrich). Finally, coverslips were washed and mounted onto glass slides using Fluoromount-G mounting medium (SouthernBiotech). This protocol was used

for all cilia detection and Xap5 ectopic expression experiments. For the detection of ectopically expressed Nono-GFP, cells were fixed and washed as above, then directly counterstained with DAPI before mounting, omitting the permeabilization and antibody incubation steps.

General fluorescence for confirming ectopic protein expression was assessed using a Zeiss Axio Vert. A1 inverted microscope. High-resolution imaging of primary cilia was performed on a Leica SP8 confocal microscope using a 63× oil-immersion objective. The percentage of ciliated cells and cilia length were quantified from confocal images using the Automated Cilia Detection and Counter (ACDC) software (v0.93)[64]. The nuclear area was quantified from DAPI-stained images using Fiji software.

RNA isolation and RT-qPCR

Total RNA was isolated from cells using the FastPure Cell/Tissue Total RNA Isolation Kit (Vazyme, RC112) according to the manufacturer's protocol. For each sample, 500 ng of total RNA was reverse-transcribed into cDNA using the HiScript II 1st Strand cDNA Synthesis Kit (Vazyme, R212). RT-qPCR was then performed with ChamQ SYBR qPCR Master Mix (Vazyme, Q311) on an Applied Biosystems StepOnePlus Real-Time PCR Systems. Relative gene expression was calculated using the $2^{-\Delta\Delta Ct}$ method, with *Gapdh* serving as the internal normalization control. All primer sequences are listed in Supplementary Table 1.

Western blotting

Cells were lysed on ice in RIPA buffer (Beyotime, P0013B) supplemented with 1× protease inhibitor cocktail (Roche, 11873580001) and 1 mM PMSF (Sigma, 93482). The resulting lysates were resolved by SDS-PAGE and transferred to polyvinylidene fluoride membranes (Millipore, IPVH00010) using a wet transfer system (Tanon). Membranes were blocked with 5% non-fat milk in TBST (20 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.1% Tween-20) for 1 h at room

temperature, followed by incubation with primary antibodies overnight at 4 °C. After washing, membranes were incubated with HRP-conjugated secondary antibodies for 1 h at room temperature. Chemiluminescence signals were developed using the ECL Super Kit (ABClonal, RM02867) and visualized with a Tanon 5200 Chemiluminescence Imaging System.

Co-immunoprecipitation (Co-IP) and mass spectrometry

For Co-IP and IP-MS experiments, cultured cells were washed twice with cold 1× PBS and harvested by scraping. Total proteins were extracted on ice using a Co-IP buffer containing 50 mM HEPES (pH 7.5), 200 mM KCl, 1 mM EGTA, 1 mM $MgCl_2$, 10% (v/v) glycerol, 0.5 mM DTT, 0.3% (v/v) NP-40, and 1× protease inhibitor cocktail (Roche). For immunoprecipitation, antibodies were first crosslinked to Protein A/G magnetic beads (Invitrogen, 80105G) using 5 mM BS^3. The resulting antibody-conjugated beads were then incubated with the cell lysates for 2 h at 4°C with gentle rotation. After extensive washing, bound proteins were eluted with 0.1 M Glycine-HCl (pH 2.5). For western blot analysis, the eluates and input samples were neutralized, mixed with SDS loading buffer, boiled for 5 min, and subsequently analyzed by immunoblotting. For protein identification, the eluates were sent for analysis by liquid chromatography with tandem mass spectrometry (LC-MS/MS) at Novogene (Beijing, China).

RNA-sequencing and analysis

For transcriptomic analysis, cells were plated for 24 h and subsequently serum-starved for 24 h. At the end of the starvation period, cells were immediately washed three times with 1× PBS and then lysed directly in TRIzol Reagent (Invitrogen) for total RNA extraction. Library construction and paired-end (2 × 150 bp) sequencing on an Illumina NovaSeq 6000 platform were performed by BioMarker Technologies (Beijing, China).

The raw sequencing reads were mapped to the Mus musculus reference genome

(GRCm39/mm39) using HISAT2[65] (v2.1.0). StringTie[66] (v1.3.3b) was subsequently used to assemble the mapped reads. Gene expression levels were quantified as fragments per kilobase of transcript per million fragments mapped (FPKM). Differential expression analysis between experimental groups was performed using DESeq2[67] (v1.12.4). Genes with a false discovery rate (FDR) < 0.01 and an absolute fold change ≥ 1.5 were considered significantly differentially expressed. Gene Ontology (GO) enrichment analysis was performed using the DAVID database. Gene Set Enrichment Analysis (GSEA) was conducted using the OECloud platform (https://cloud.oebiotech.com).

CUT&Tag-sequencing and analysis

For CUT&Tag analysis, NIH/3T3 cells were plated for 24 h and subsequently serum-starved for 24 h. Following harvesting, 100,000 cells were collected for each sample. Library construction was performed using the Hyperactive Universal CUT&Tag Assay Kit for Illumina Pro (Vazyme, TD904) according to the manufacturer's protocol. The resulting libraries were sequenced on an Illumina NovaSeq platform (150 bp paired-end reads) at Novogene Co., Ltd. (Beijing, China).

For bioinformatic processing, raw reads were assessed for quality with FastQC[68] (v0.11.9) and trimmed with fastp[69] (v0.20.0; parameters: --length_required 15, --n_base_limit 6). Clean reads were aligned to the Mus musculus reference genome (mm39) using BWA-MEM (v0.7.12; parameters: -k 32, -T 30, -M). Peak calling was performed using MACS2[70] (v2.1.0) with a q-value threshold of 0.05 and the following parameters: -f AUTO --call-summits --nomodel --shift -100 --extsize 200 --keep-dup all. For *de novo* motif discovery, peak summits were extended to 500 bp and analyzed using the HOMER[71] (v4.9.1) findMotifsGenome.pl script. Peak annotation and association with the nearest genes were performed using the ChIPseeker[72] R package. Genome browser tracks were generated using the Integrative Genomics Viewer (IGV, v2.18.1)[73].

CUT&Tag-qPCR

For targeted validation, DNA from 100,000 cells was prepared in parallel using the same CUT&Tag kit according to the manufacturer's protocol. qPCR was performed using the resulting DNA. Immunoprecipitation with IgG sera served as a negative control, and an exogenous DNA spike-in was used as an internal normalization control. Enrichment at specific loci was calculated relative to the spike-in DNA control, and the $2^{-\Delta\Delta Ct}$ method was used to compare experimental samples. All primer sequences used for CUT&Tag-qPCR are listed in Supplementary Table 1.

Statistical analysis

Statistical analysis was performed using GraphPad Prism v6.01 or v8.0.2. All data were presented as mean ± SEM. Values of $P < 0.05$ were considered statistically significant. Statistical significance between two groups was calculated using an unpaired, parametric, two-sided Student's t-test.

Data availability

The RNA-sequencing data generated in this study have been deposited in the Gene Expression Omnibus (GEO) database under accession code GSE305355. The CUT&Tag-sequencing data reported in this paper are available in the GEO database under accession code GSE304980. All other data supporting the findings of this study are available within the article and its Supplementary Information files. Source data are provided with this paper.

References

1. True, J.R., and Carroll, S.B. (2002). Gene co-option in physiological and morphological evolution. Annu Rev Cell Dev Biol *18*, 53-80. 10.1146/annurev.cellbio.18.020402.140619.
2. Peter, I.S., and Davidson, E.H. (2011). Evolution of gene regulatory networks controlling body plan development. Cell *144*, 970-985. 10.1016/j.cell.2011.02.017.
3. Davidson, E.H., and Erwin, D.H. (2006). Gene regulatory networks and the evolution of animal body plans. Science *311*, 796-800. 10.1126/science.1113832.
4. Iwafuchi-Doi, M., and Zaret, K.S. (2016). Cell fate control by pioneer transcription factors.

Development *143*, 1833-1837. 10.1242/dev.133900.

5. Zaret, K.S. (2020). Pioneer Transcription Factors Initiating Gene Network Changes. Annu Rev Genet *54*, 367-385. 10.1146/annurev-genet-030220-015007.
6. Whyte, W.A., Orlando, D.A., Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H., Rahl, P.B., Lee, T.I., and Young, R.A. (2013). Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell *153*, 307-319. 10.1016/j.cell.2013.03.035.
7. Levine, M., and Davidson, E.H. (2005). Gene regulatory networks for development. Proc Natl Acad Sci U S A *102*, 4936-4942. 10.1073/pnas.0408031102.
8. Carroll, S.B. (2008). Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. Cell *134*, 25-36. 10.1016/j.cell.2008.06.030.
9. Wagner, G.P., and Lynch, V.J. (2010). Evolutionary novelties. Curr Biol *20*, R48-52. 10.1016/j.cub.2009.11.010.
10. Arendt, D., Musser, J.M., Baker, C.V.H., Bergman, A., Cepko, C., Erwin, D.H., Pavlicev, M., Schlosser, G., Widder, S., Laubichler, M.D., and Wagner, G.P. (2016). The origin and evolution of cell types. Nat Rev Genet *17*, 744-757. 10.1038/nrg.2016.127.
11. Carvalho-Santos, Z., Azimzadeh, J., Pereira-Leal, J.B., and Bettencourt-Dias, M. (2011). Evolution: Tracing the origins of centrioles, cilia, and flagella. J Cell Biol *194*, 165-175. 10.1083/jcb.201011152.
12. Goetz, S.C., and Anderson, K.V. (2010). The primary cilium: a signalling centre during vertebrate development. Nat Rev Genet *11*, 331-344. 10.1038/nrg2774.
13. Anvarian, Z., Mykytyn, K., Mukhopadhyay, S., Pedersen, L.B., and Christensen, S.T. (2019). Cellular signalling by primary cilia in development, organ function and disease. Nat Rev Nephrol *15*, 199-219. 10.1038/s41581-019-0116-9.
14. Spassky, N., and Meunier, A. (2017). The development and functions of multiciliated epithelia. Nat Rev Mol Cell Biol *18*, 423-436. 10.1038/nrm.2017.21.
15. Reiter, J.F., and Leroux, M.R. (2017). Genes and molecular pathways underpinning ciliopathies. Nat Rev Mol Cell Biol *18*, 533-547. 10.1038/nrm.2017.60.
16. van Dam, T.J.P., Kennedy, J., van der Lee, R., de Vrieze, E., Wunderlich, K.A., Rix, S., Dougherty, G.W., Lambacher, N.J., Li, C., Jensen, V.L., et al. (2019). CiliaCarta: An integrated and validated compendium of ciliary genes. PLoS One *14*, e0216705. 10.1371/journal.pone.0216705.
17. van Dam, T.J., Wheway, G., Slaats, G.G., Group, S.S., Huynen, M.A., and Giles, R.H. (2013). The SYSCILIA gold standard (SCGSv1) of known ciliary components and its applications within a systems biology consortium. Cilia *2*, 7. 10.1186/2046-2530-2-7.
18. Vasquez, S.S.V., van Dam, J., and Wheway, G. (2021). An updated SYSCILIA gold standard (SCGSv2) of known ciliary genes, revealing the vast progress that has been made in the cilia research field. Mol Biol Cell *32*, br13. 10.1091/mbc.E21-05-0226.
19. Thomas, J., Morle, L., Soulavie, F., Laurencon, A., Sagnol, S., and Durand, B. (2010). Transcriptional control of genes involved in ciliogenesis: a first step in making cilia. Biol Cell *102*, 499-513. 10.1042/BC20100035.
20. Choksi, S.P., Lauter, G., Swoboda, P., and Roy, S. (2014). Switching on cilia: transcriptional networks regulating ciliogenesis. Development *141*, 1427-1441. 10.1242/dev.074666.
21. Lewis, M., and Stracker, T.H. (2021). Transcriptional regulation of multiciliated cell differentiation. Semin Cell Dev Biol *110*, 51-60. 10.1016/j.semcdb.2020.04.007.
22. Collins, C., Ventrella, R., and Mitchell, B.J. (2021). Building a ciliated epithelium: Transcriptional regulation and radial intercalation of multiciliated cells. Curr Top Dev Biol *145*, 3-39. 10.1016/bs.ctdb.2020.08.001.
23. Swoboda, P., Adler, H.T., and Thomas, J.H. (2000). The RFX-type transcription factor DAF-19 regulates sensory neuron cilium formation in C. elegans. Mol Cell *5*, 411-421. 10.1016/s1097-2765(00)80436-0.
24. Yu, X., Ng, C.P., Habacher, H., and Roy, S. (2008). Foxj1 transcription factors are master regulators of the motile ciliogenic program. Nat Genet *40*, 1445-1453. 10.1038/ng.263.

25. Stubbs, J.L., Oishi, I., Izpisua Belmonte, J.C., and Kintner, C. (2008). The forkhead protein Foxj1 specifies node-like cilia in Xenopus and zebrafish embryos. Nat Genet *40*, 1454-1460. 10.1038/ng.267.
26. Nachury, M.V., and Mick, D.U. (2019). Establishing and regulating the composition of cilia for signal transduction. Nat Rev Mol Cell Biol *20*, 389-405. 10.1038/s41580-019-0116-4.
27. Li, L., Tian, G., Peng, H., Meng, D., Wang, L., Hu, X., Tian, C., He, M., Zhou, J., Chen, L., et al. (2018). New class of transcription factors controls flagellar assembly by recruiting RNA polymerase II in Chlamydomonas. Proc Natl Acad Sci U S A *115*, 4435-4440. 10.1073/pnas.1719206115.
28. Wang, W., Xing, J., Zhang, X., Liu, H., Liu, X., Jiang, H., Xu, C., Zhao, X., and Hu, Z. (2025). Control of ciliary transcriptional programs during spermatogenesis by antagonistic transcription factors. Elife *13*. 10.7554/eLife.94754.
29. Brown, J.M. (2025). Cilia and transcription: a mini review. Front Cell Dev Biol *13*, 1582796. 10.3389/fcell.2025.1582796.
30. Gomez, A.E., Christman, A.K., Van De Weghe, J.C., Finn, M., and Doherty, D. (2022). Systematic analysis of cilia characteristics and Hedgehog signaling in five immortal cell lines. PLoS One *17*, e0266433. 10.1371/journal.pone.0266433.
31. Xu, W.P., Cui, Y.L., Chen, L.L., Ding, K., Ding, C.H., Chen, F., Zhang, X., and Xie, W.F. (2021). Deletion of Sox9 in the liver leads to hepatic cystogenesis in mice by transcriptionally downregulating Sec63. J Pathol *254*, 57-69. 10.1002/path.5636.
32. Lodh, S., O'Hare, E.A., and Zaghloul, N.A. (2014). Primary cilia in pancreatic development and disease. Birth Defects Res C Embryo Today *102*, 139-158. 10.1002/bdrc.21063.
33. Hrncir, H.R., Goodloe, B., Bombin, S., Hogan, C.B., Jadi, O., and Gracz, A.D. (2025). Sox9 inhibits Activin A to promote biliary maturation and branching morphogenesis. Nat Commun *16*, 1667. 10.1038/s41467-025-56813-x.
34. Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. Nat Protoc *8*, 2281-2308. 10.1038/nprot.2013.143.
35. Knott, G.J., Bond, C.S., and Fox, A.H. (2016). The DBHS proteins SFPQ, NONO and PSPC1: a multipurpose molecular scaffold. Nucleic Acids Res *44*, 3989-4004. 10.1093/nar/gkw271.
36. Amelio, A.L., Miraglia, L.J., Conkright, J.J., Mercer, B.A., Batalov, S., Cavett, V., Orth, A.P., Busby, J., Hogenesch, J.B., and Conkright, M.D. (2007). A coactivator trap identifies NONO (p54nrb) as a component of the cAMP-signaling pathway. Proc Natl Acad Sci U S A *104*, 20314-20319. 10.1073/pnas.0707999105.
37. Fox, A.H., Nakagawa, S., Hirose, T., and Bond, C.S. (2018). Paraspeckles: Where Long Noncoding RNA Meets Phase Separation. Trends Biochem Sci *43*, 124-135. 10.1016/j.tibs.2017.12.001.
38. Kaya-Okur, H.S., Wu, S.J., Codomo, C.A., Pledger, E.S., Bryson, T.D., Henikoff, J.G., Ahmad, K., and Henikoff, S. (2019). CUT&Tag for efficient epigenomic profiling of small samples and single cells. Nat Commun *10*, 1930. 10.1038/s41467-019-09982-5.
39. Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A *102*, 15545-15550. 10.1073/pnas.0506580102.
40. Marshall, W.F. (2024). Chlamydomonas as a model system to study cilia and flagella using genetics, biochemistry, and microscopy. Front Cell Dev Biol *12*, 1412641. 10.3389/fcell.2024.1412641.
41. Lyu, Q., Li, Q., Zhou, J., and Zhao, H. (2024). Formation and function of multiciliated cells. J Cell Biol *223*. 10.1083/jcb.202307150.
42. Brooks, E.R., and Wallingford, J.B. (2014). Multiciliated cells. Curr Biol *24*, R973-982. 10.1016/j.cub.2014.08.047.
43. Rayamajhi, D., and Roy, S. (2025). Multiciliated cells: Development, functions and disease

relevance. Semin Cell Dev Biol *175*, 103660. 10.1016/j.semcdb.2025.103660.
44. Lemeille, S., Paschaki, M., Baas, D., Morle, L., Duteyrat, J.L., Ait-Lounis, A., Barras, E., Soulavie, F., Jerber, J., Thomas, J., et al. (2020). Interplay of RFX transcription factors 1, 2 and 3 in motile ciliogenesis. Nucleic Acids Res *48*, 9019-9036. 10.1093/nar/gkaa625.
45. Piasecki, B.P., Burghoorn, J., and Swoboda, P. (2010). Regulatory Factor X (RFX)-mediated transcriptional rewiring of ciliary genes in animals. Proc Natl Acad Sci U S A *107*, 12969-12974. 10.1073/pnas.0914241107.
46. Coyle, M.C., Tajima, A.M., Leon, F., Choksi, S.P., Yang, A., Espinoza, S., Hughes, T.R., Reiter, J.F., Booth, D.S., and King, N. (2023). An RFX transcription factor regulates ciliogenesis in the closest living relatives of animals. Curr Biol *33*, 3747-3758 e3749. 10.1016/j.cub.2023.07.022.
47. Ma, L., Quigley, I., Omran, H., and Kintner, C. (2014). Multicilin drives centriole biogenesis via E2f proteins. Genes Dev *28*, 1461-1471. 10.1101/gad.243832.114.
48. Stubbs, J.L., Vladar, E.K., Axelrod, J.D., and Kintner, C. (2012). Multicilin promotes centriole assembly and ciliogenesis during multiciliate cell differentiation. Nat Cell Biol *14*, 140-147. 10.1038/ncb2406.
49. Arbi, M., Pefani, D.E., Kyrousi, C., Lalioti, M.E., Kalogeropoulou, A., Papanastasiou, A.D., Taraviras, S., and Lygerou, Z. (2016). GemC1 controls multiciliogenesis in the airway epithelium. EMBO Rep *17*, 400-413. 10.15252/embr.201540882.
50. Terre, B., Piergiovanni, G., Segura-Bayona, S., Gil-Gomez, G., Youssef, S.A., Attolini, C.S., Wilsch-Brauninger, M., Jung, C., Rojas, A.M., Marjanovic, M., et al. (2016). GEMC1 is a critical regulator of multiciliated cell differentiation. EMBO J *35*, 942-960. 10.15252/embj.201592821.
51. Beckers, A., Alten, L., Viebahn, C., Andre, P., and Gossler, A. (2007). The mouse homeobox gene Noto regulates node morphogenesis, notochordal ciliogenesis, and left right patterning. Proc Natl Acad Sci U S A *104*, 15765-15770. 10.1073/pnas.0704344104.
52. Alten, L., Schuster-Gossler, K., Beckers, A., Groos, S., Ulmer, B., Hegermann, J., Ochs, M., and Gossler, A. (2012). Differential regulation of node formation, nodal ciliogenesis and cilia positioning by Noto and Foxj1. Development *139*, 1276-1284. 10.1242/dev.072728.
53. Nemajerova, A., Kramer, D., Siller, S.S., Herr, C., Shomroni, O., Pena, T., Gallinas Suazo, C., Glaser, K., Wildung, M., Steffen, H., et al. (2016). TAp73 is a central transcriptional regulator of airway multiciliogenesis. Genes Dev *30*, 1300-1312. 10.1101/gad.279836.116.
54. Marshall, C.B., Mays, D.J., Beeler, J.S., Rosenbluth, J.M., Boyd, K.L., Santos Guasch, G.L., Shaver, T.M., Tang, L.J., Liu, Q., Shyr, Y., et al. (2016). p73 Is Required for Multiciliogenesis and Regulates the Foxj1-Associated Gene Network. Cell Rep *14*, 2289-2300. 10.1016/j.celrep.2016.02.035.
55. Tan, F.E., Vladar, E.K., Ma, L., Fuentealba, L.C., Hoh, R., Espinoza, F.H., Axelrod, J.D., Alvarez-Buylla, A., Stearns, T., Kintner, C., and Krasnow, M.A. (2013). Myb promotes centriole amplification and later steps of the multiciliogenesis program. Development *140*, 4277-4286. 10.1242/dev.094102.
56. Sedlacek, Z., Munstermann, E., Dhorne-Pollet, S., Otto, C., Bock, D., Schutz, G., and Poustka, A. (1999). Human and mouse XAP-5 and XAP-5-like (X5L) genes: identification of an ancient functional retroposon differentially expressed in testis. Genomics *61*, 125-132. 10.1006/geno.1999.5931.
57. Martin-Tryon, E.L., and Harmer, S.L. (2008). XAP5 CIRCADIAN TIMEKEEPER coordinates light signals for proper timing of photomorphogenesis and the circadian clock in Arabidopsis. Plant Cell *20*, 1244-1259. 10.1105/tpc.107.056655.
58. Liu, C.F., and Lefebvre, V. (2015). The transcription factors SOX9 and SOX5/SOX6 cooperate genome-wide through super-enhancers to drive chondrogenesis. Nucleic Acids Res *43*, 8183-8203. 10.1093/nar/gkv688.
59. Kiselak, E.A., Shen, X., Song, J., Gude, D.R., Wang, J., Brody, S.L., Strauss, J.F., 3rd, and Zhang, Z. (2010). Transcriptional regulation of an axonemal central apparatus gene, sperm-associated antigen 6, by a SRY-related high mobility group transcription factor, S-SOX5. J Biol Chem *285*,

30496-30505. 10.1074/jbc.M110.121590.
60. Zhang, L., Liu, Y., Li, W., Zhang, Q., Li, Y., Liu, J., Min, J., Shuang, C., Song, S., and Zhang, Z. (2017). Transcriptional regulation of human sperm-associated antigen 16 gene by S-SOX5. BMC Mol Biol *18*, 2. 10.1186/s12867-017-0082-3.
61. Burghoorn, J., Piasecki, B.P., Crona, F., Phirke, P., Jeppsson, K.E., and Swoboda, P. (2012). The in vivo dissection of direct RFX-target gene promoters in C. elegans reveals a novel cis-regulatory element, the C-box. Dev Biol *368*, 415-426. 10.1016/j.ydbio.2012.05.033.
62. Liang, Y., Koche, R., Chalamalasetty, R.B., Stephen, D.N., Kennedy, M.W., Lao, Z., Pang, Y., Kuo, Y.Y., Lee, M., Lobo, F.P., et al. (2025). Transcription factors SP5 and SP8 drive primary cilia formation in mammalian embryos. Science *389*, eadt5663. 10.1126/science.adt5663.
63. Presnell, J.S., Schnitzler, C.E., and Browne, W.E. (2015). KLF/SP Transcription Factor Family Evolution: Expansion, Diversification, and Innovation in Eukaryotes. Genome Biol Evol *7*, 2289-2309. 10.1093/gbe/evv141.
64. Lauring, M.C., Zhu, T., Luo, W., Wu, W., Yu, F., and Toomre, D. (2019). New software for automated cilia detection in cells (ACDC). Cilia *8*, 1. 10.1186/s13630-019-0061-z.
65. Kim, D., Langmead, B., and Salzberg, S.L. (2015). HISAT: a fast spliced aligner with low memory requirements. Nat Methods *12*, 357-360. 10.1038/nmeth.3317.
66. Pertea, M., Pertea, G.M., Antonescu, C.M., Chang, T.C., Mendell, J.T., and Salzberg, S.L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol *33*, 290-295. 10.1038/nbt.3122.
67. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol *15*, 550. 10.1186/s13059-014-0550-8.
68. Andrews, S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.
69. Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics *34*, i884-i890. 10.1093/bioinformatics/bty560.
70. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol *9*, R137. 10.1186/gb-2008-9-9-r137.
71. Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell *38*, 576-589. 10.1016/j.molcel.2010.05.004.
72. Yu, G., Wang, L.G., and He, Q.Y. (2015). ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. Bioinformatics *31*, 2382-2383. 10.1093/bioinformatics/btv145.
73. Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nat Biotechnol *29*, 24-26. 10.1038/nbt.1754.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (grant number: 32170702), the Natural Science Foundation for Distinguished Young Scholars of Hubei Province (grant number: 2025AFA060) and the Wuhan Municipal Education Bureau's Program for the Integration of Research and Education (grant number: 2025KCJ03). We thank Dr. Peter

Swoboda (Karolinska Institute) for his constructive comments on the preprint version of this manuscript, specifically for pointing out the sequence similarity between the Xap5 motif and the *C. elegans* C-box.

Author Contributions

W.W., X.Z., Y.Q., X.M., G.T. and Z.H. conceived and designed the research. W.W., X.Z., Y.Q., X.M., S.C., Y.C., S.L., W.C., J.Y., X.Y., H.L., J.X., C.X. and H.J. performed the research. W.W., G.T. and Z.H. contributed new reagents/analytic tools. W.W., H.W., G.T. and Z.H. analyzed the data. W.W. and Z.H. wrote the manuscript. All authors reviewed the manuscript and approved the final version.

Competing interests

The authors declare no competing interests.

Supplementary information

Supplementary Data 1

Differentially expressed genes between WT and *Xap5* KO identified by RNA-seq

Supplementary Data 2

Differentially expressed genes between WT and *Nono* KO identified by RNA-seq

Supplementary Data 3

Curated list of cilia-associated genes used in the study

Supplementary Data 4

List of Xap5-bound genes identified by CUT&Tag analysis

Supplementary Data 5

List of Nono-bound genes identified by CUT&Tag analysis

Supplementary Data 6

Differentially expressed genes between WT and *Sox5* KO identified by RNA-seq

Figures

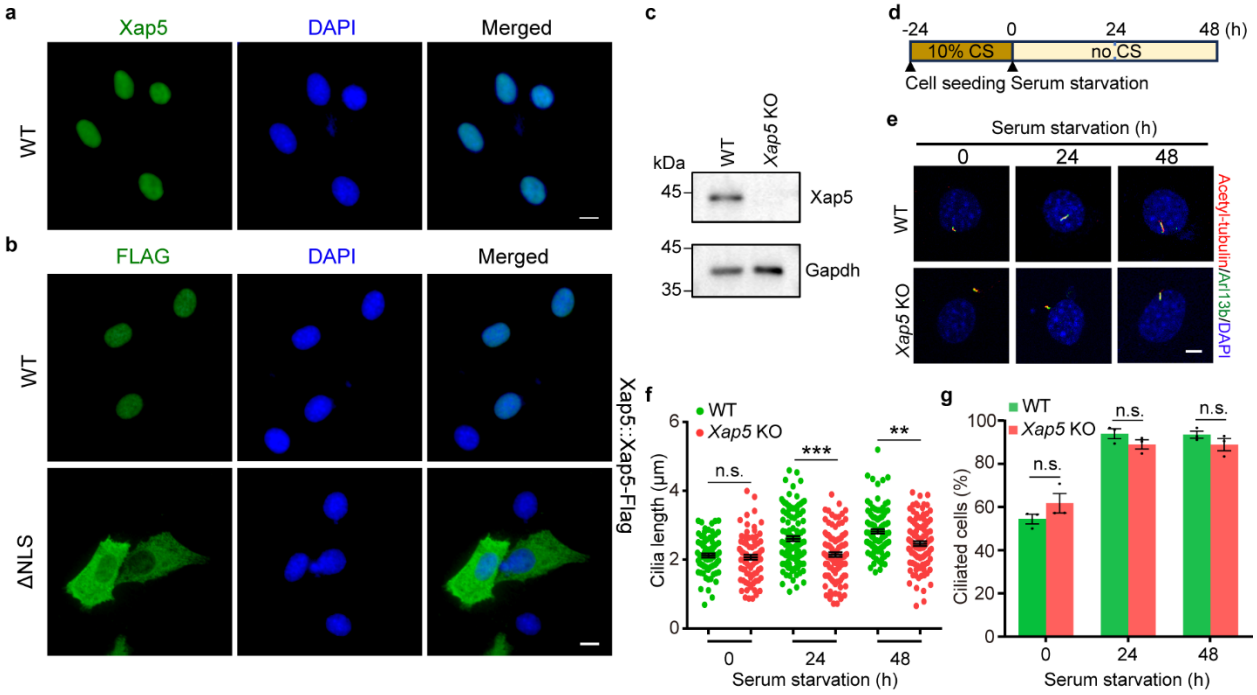


Fig. 1 Xap5 is a nuclear protein essential for primary ciliogenesis. **a** Immunofluorescence analysis of endogenous Xap5 (green) in NIH/3T3 cells. Nuclei were counterstained with DAPI (blue). Scale bar, 10 μm. **b** Validation of the Xap5 nuclear localization signal (NLS). Representative images of NIH/3T3 cells expressing either Flag-tagged wild-type Xap5 (WT) or an NLS-deleted mutant (ΔNLS) (green). Scale bar, 10 μm. **c** Western blot analysis confirming the complete depletion of Xap5 protein in the *Xap5* knockout (KO) clonal cell line. Gapdh serves as a loading control. The blot shown is representative of three independent biological replicates. **d** Schematic illustrating the time-course assay for primary cilia assembly induced by serum starvation. CS, calf serum. **e** Representative immunofluorescence images of cilia from WT and *Xap5* KO cells at the indicated time points. Cilia were co-stained for acetylated tubulin (red) and Arl13b (green). Nuclei were counterstained with DAPI (blue). Scale bar, 5 μm. **f** Quantification of primary cilia length in WT and *Xap5* KO cells over a 48-h time course of serum starvation. Data are presented as means ± SEM (n = 3 independent experiments). **g** Quantification of the

percentage of ciliated cells as in **f**. Data are presented as mean ± SEM (n = 3 independent experiments). All statistical significance between WT and KO at each time point was determined by multiple two-tailed unpaired t-tests. n.s., not significant; $^{**}P < 0.01$; $^{***}P < 0.001$. Source data are provided as a Source Data file.

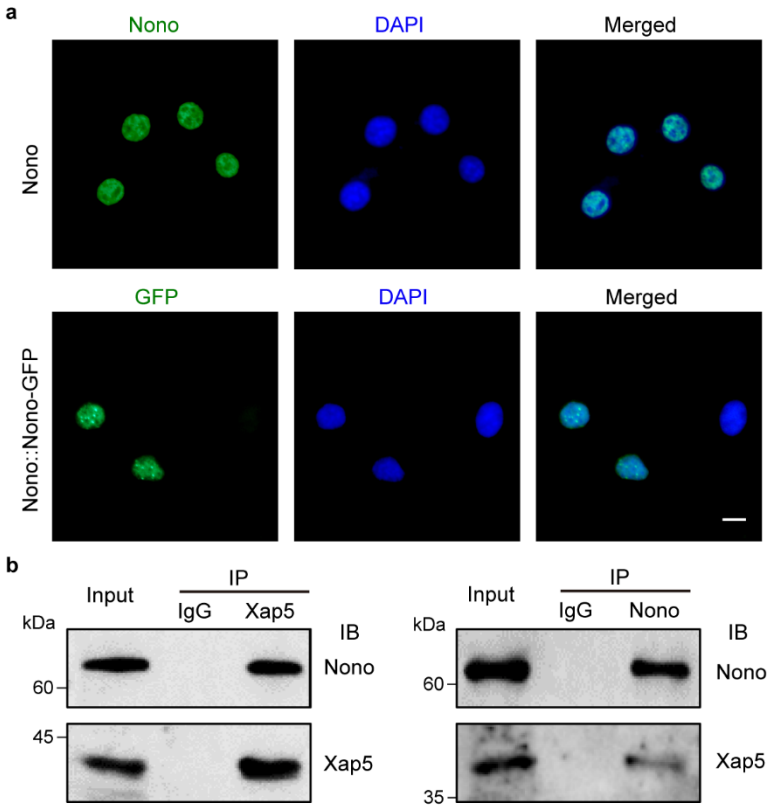


Fig. 2 Xap5 forms a nuclear complex with its cofactor Nono. **a** Representative immunofluorescence images showing the predominantly nuclear localization of Nono in NIH/3T3 cells. Top: Endogenous Nono detected with a specific anti-Nono antibody (green). Bottom: Ectopically expressed Nono-GFP fusion protein (green). Nuclei were counterstained with DAPI (blue). Scale bars, 10 μm. **b** Endogenous co-immunoprecipitation (Co-IP) of Xap5 and Nono from NIH/3T3 cell lysates. Left: Immunoprecipitation (IP) with an anti-Xap5 antibody specifically co-precipitated endogenous Nono. Right: Conversely, IP with an anti-Nono antibody co-precipitated endogenous Xap5. Control IgG was used as a negative control for the IP. The resulting precipitates (IP) and input lysates were analyzed by western blotting (IB) with the indicated antibodies. Data are representative of three independent experiments. Source data are provided as a Source Data file.

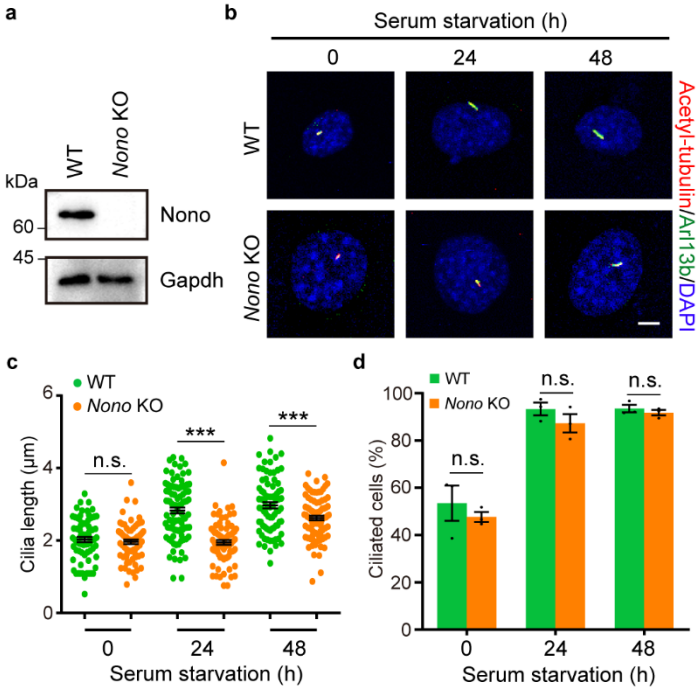


Fig. 3 Loss of Nono phenocopies Xap5 depletion, impairing primary cilia elongation. **a** Western blot analysis confirming the complete depletion of Nono protein in *Nono* KO cell lysates compared to WT. The blot shown is representative of three independent biological replicates. Gapdh serves as a loading control. **b** Representative immunofluorescence images of primary cilia in WT and *Nono* KO cells after 48 h of serum starvation. Cilia were co-stained for acetyl-tubulin (red) and Arl13b (green). Nuclei were counterstained with DAPI (blue). Scale bar, 5 μm. **c** Quantification of primary cilia length in WT and *Nono* KO cells over a 48-h time course of serum starvation. Data are presented as mean ± SEM (n = 3 independent experiments). **d** Quantification of the percentage of ciliated cells as in **c**. Data are presented as mean ± SEM (n = 3 independent experiments). All statistical significance between WT and KO at each time point was determined by multiple two-tailed unpaired t-tests. n.s., not significant; $^{**}P < 0.01$; $^{***}P < 0.001$. Source data are provided as a Source Data file.

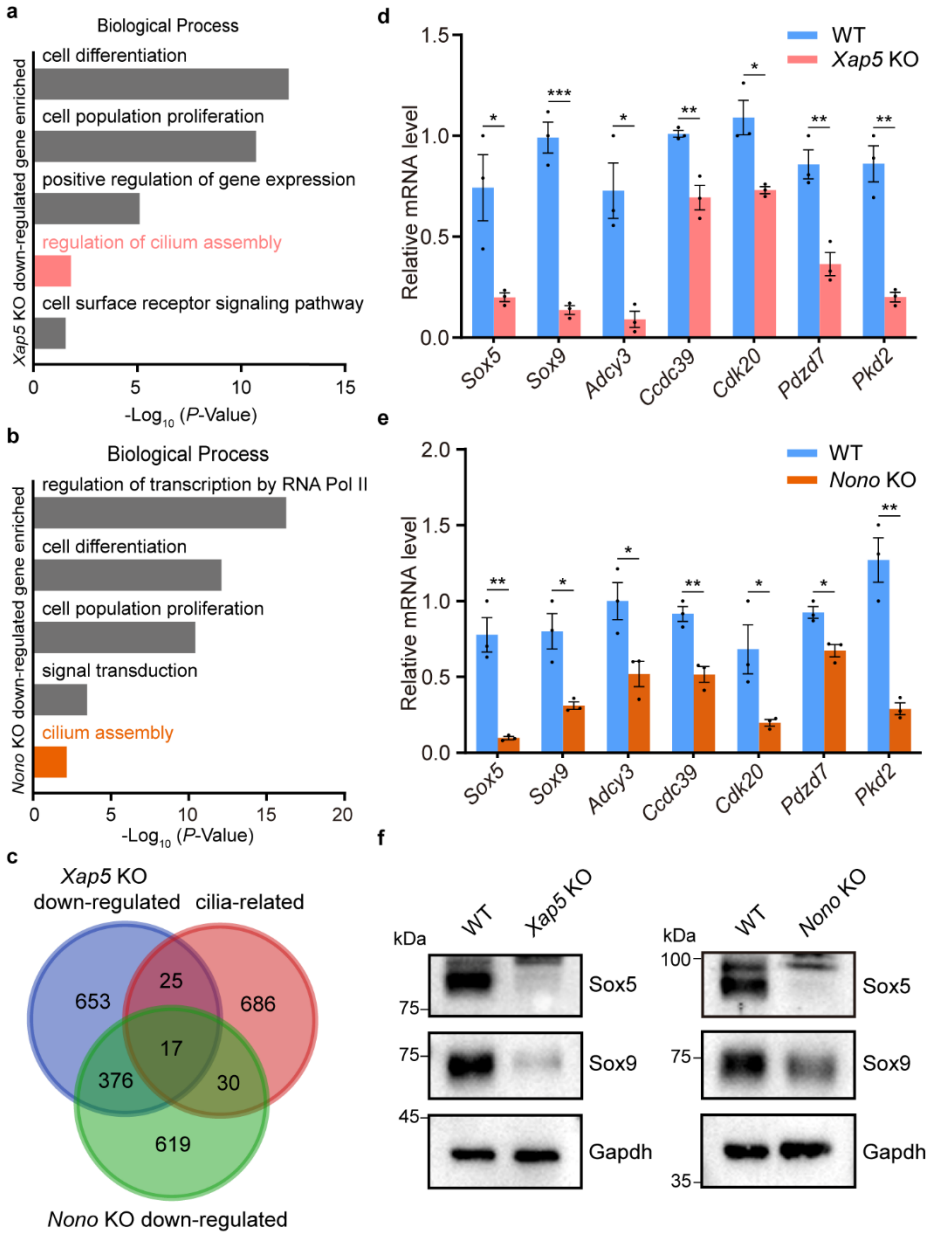


Fig. 4 Xap5 and Nono co-regulate a ciliogenic transcriptional network. All analyses were performed on NIH/3T3 cells following 24 h of serum starvation. **a, b** Representative enriched Gene Ontology (GO) terms (Biological Process) for genes downregulated in *Xap5* KO (**a**) and *Nono* KO (**b**) cells. **c** Venn diagram illustrating the significant overlap between genes downregulated in *Xap5* KO cells, *Nono* KO cells, and a curated list of cilia-associated genes (see also Supplementary Data 3). **d, e** qPCR validation of shared, downregulated ciliary target genes

in *Xap5* KO (**d**) and *Nono* KO (**e**) cells. Data are presented as mean ± SEM from three independent experiments. *P* values were determined by two-tailed unpaired t-test. **P* < 0.05, ***P* < 0.01, ****P* < 0.001. **f** Representative western blots showing reduced protein levels of Sox5 and Sox9 in *Xap5* KO (left) and *Nono* KO (right) cells compared to their respective WT controls. Gapdh serves as a loading control. Blots are representative of three independent experiments. Source data are provided as a Source Data file.

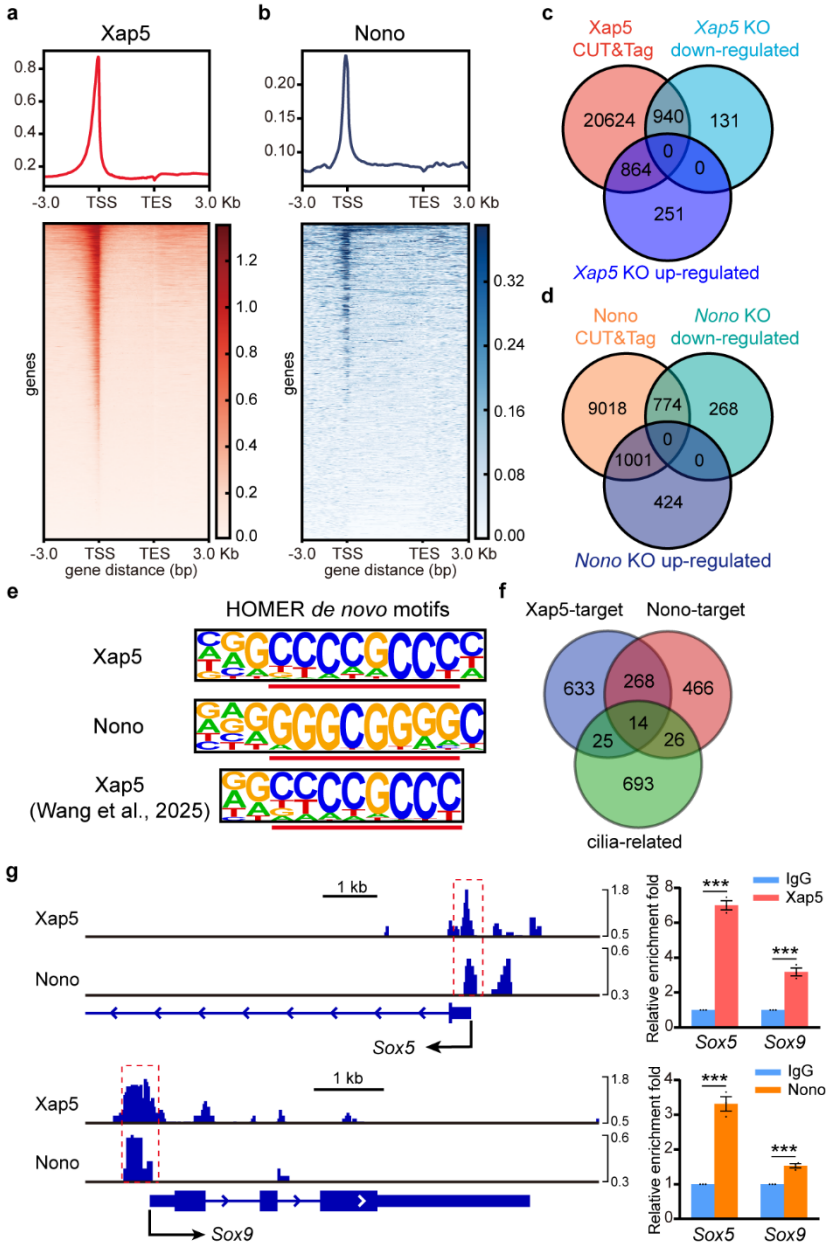


Fig. 5 Xap5 and Nono directly co-occupy the regulatory regions of ciliary genes. **a, b** Genome-wide binding profiles of Xap5 (**a**) and Nono (**b**) in NIH/3T3 cells determined by CUT&Tag. Top panels show the average enrichment profiles over gene bodies (from transcription start site, TSS, to transcription end site, TES). Bottom panels show heatmaps of binding signals centered on the TSS (±3 kb). **c, d** Venn diagrams illustrating the significant overlap between genes bound by Xap5 (**c**) or Nono (**d**) and the respective sets of differentially expressed genes identified upon

their knockout. **e** Comparison of *de novo* DNA-binding motifs derived from Xap5 and Nono CUT&Tag peaks in this study, alongside the Xap5 motif from our previous work in testicular germ cells[28]. The high degree of similarity indicates a conserved DNA-binding preference across different cellular contexts. **f** Venn diagram showing the overlap between direct Xap5 target genes (defined as genes that are both bound by Xap5 and downregulated in *Xap5* KO), direct Nono target genes, and the curated ciliary gene set (from Supplementary Data 3). **g** Direct co-occupancy of Xap5 and Nono at the *Sox5* and *Sox9* gene loci. Left: Integrative Genomics Viewer (IGV) tracks showing CUT&Tag signal enrichment. Right: CUT&Tag-qPCR validating the co-enrichment of Xap5 and Nono at the *Sox5* and *Sox9* promoters. Red dashed box: Xap5 and Nono binding regions. Data are presented as mean ± SEM from three independent experiments. *P* values were determined by two-tailed unpaired t-test. ***$P < 0.001$. Source data are provided as a Source Data file.

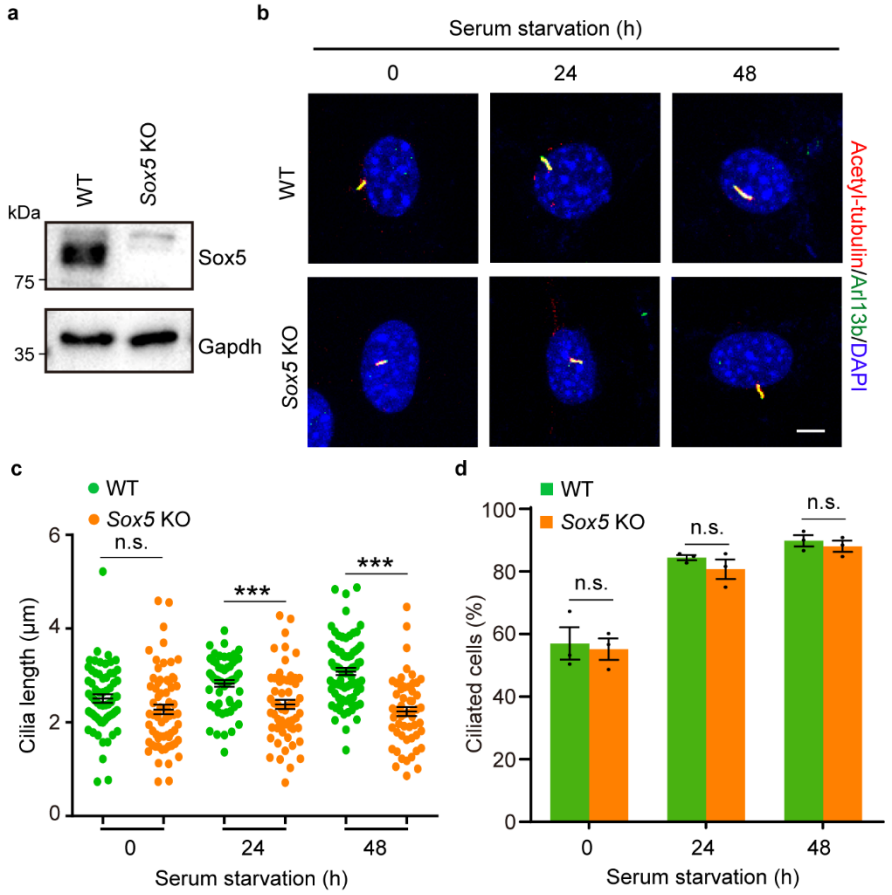


Fig. 6 Downstream effector Sox5 is essential for primary ciliogenesis. **a** Western blot analysis confirming the complete depletion of Sox5 protein in *Sox5* KO cell lysates compared to WT. The blot shown is representative of three independent biological replicates. Gapdh serves as a loading control. **b** Representative immunofluorescence images of primary cilia in WT and *Sox5* KO cells after serum starvation. Cilia were co-stained for acetylated tubulin (red) and Arl13b (green). Nuclei were counterstained with DAPI (blue). Scale bar, 5 μm. **c** Quantification of primary cilia length in WT and *Sox5* KO cells over a 48-h time course of serum starvation. Data are presented as mean ± SEM (n = 3 independent experiments). **d** Quantification of the percentage of ciliated cells as in **c**. Data are presented as mean ± SEM from three independent experiments. *P* values were determined by two-tailed unpaired t-test. n.s., not significant; ***$P < 0.001$. Source data are

provided as a Source Data file.

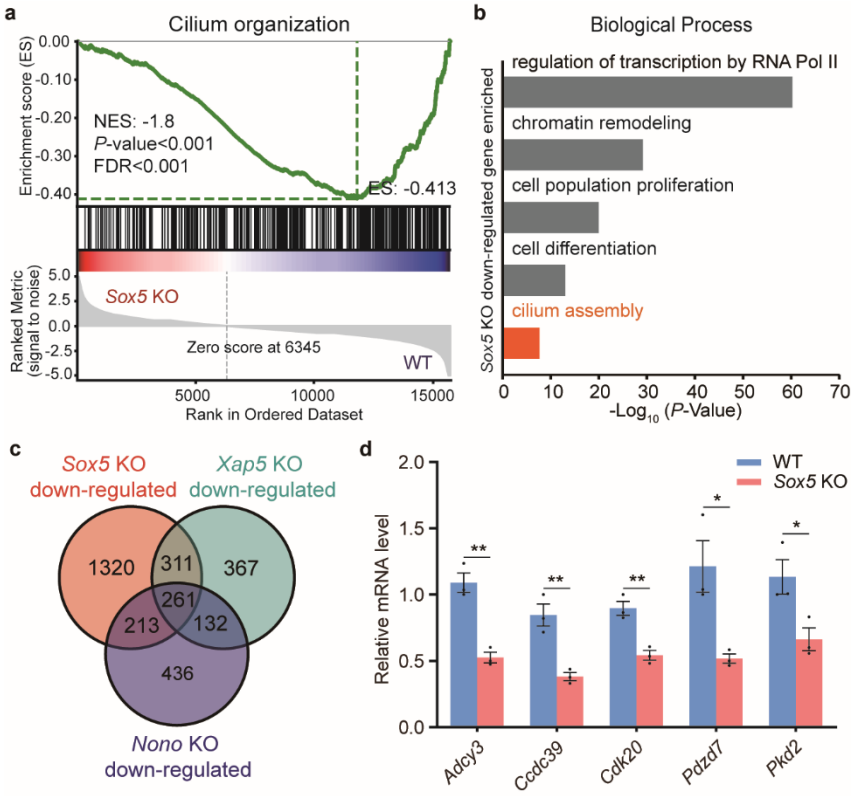
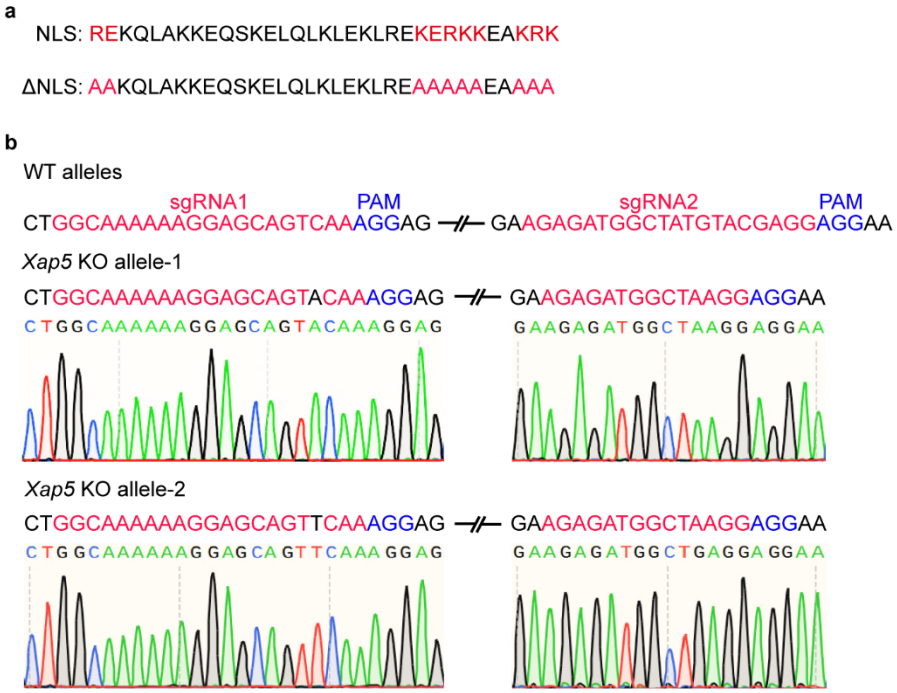


Fig. 7 Sox5 executes a ciliogenic transcriptional program downstream of the Xap5-Nono complex. **a** Gene Set Enrichment Analysis (GSEA) plot for the "cilium organization" pathway, performed on all ranked differentially expressed genes from the *Sox5* KO transcriptome. **b** Representative enriched GO terms (Biological Process) for genes significantly downregulated (fold change ≥ 1.5) in *Sox5* KO cells. **c** Venn diagram illustrating the significant overlap among genes downregulated in *Xap5* KO, *Nono* KO, and *Sox5* KO cells, identifying a common set of co-regulated genes. **d** qPCR validation confirming the reduced expression of shared downstream target genes in *Sox5* KO cells compared to WT. Data are presented as mean ± SEM from three independent experiments. *P* values were determined by two-tailed unpaired t-test. **P* < 0.05, ***P* < 0.01. Source data are provided as a Source Data file.

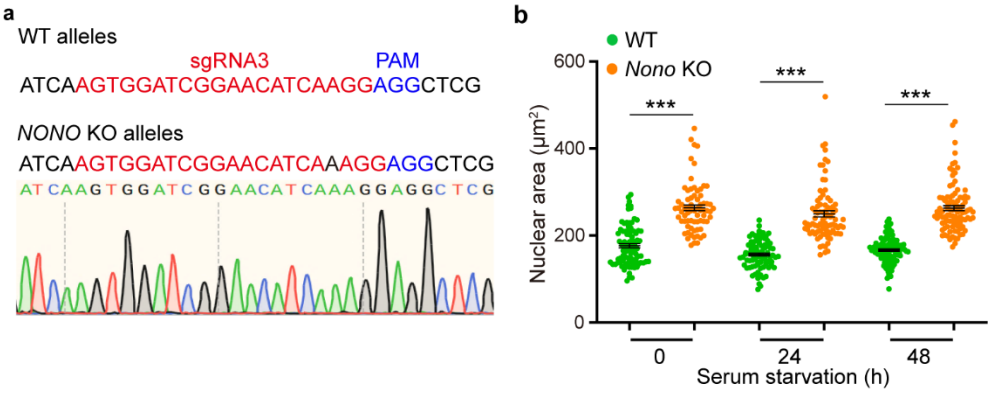
Supplementary Figures



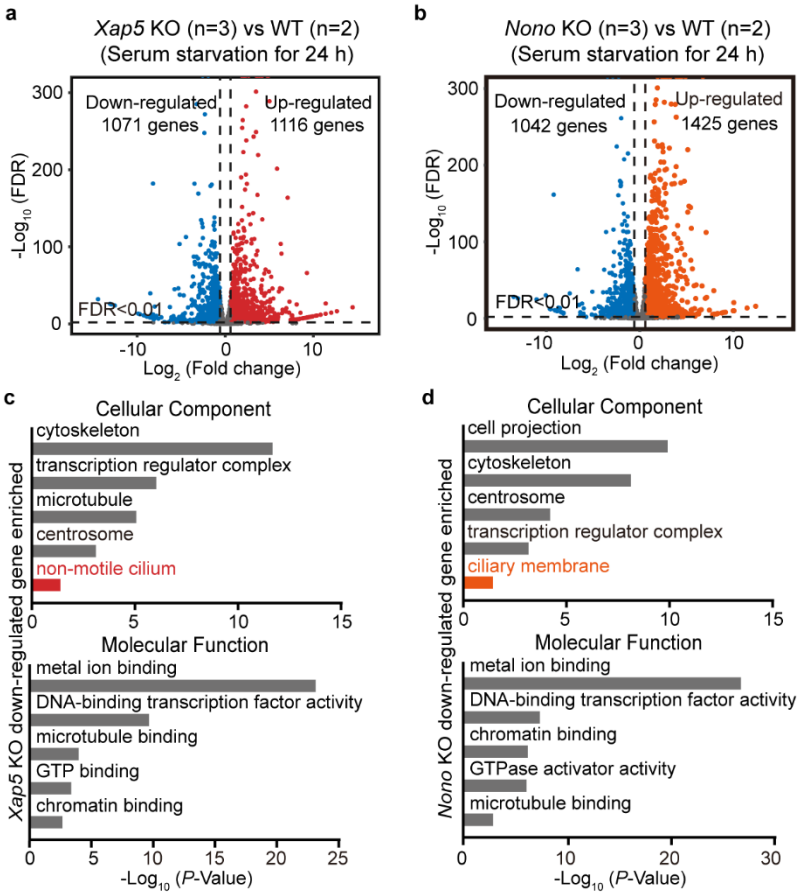
Supplementary Fig. 1 Design and validation of NLS-mutant and CRISPR/Cas9-mediated *Xap5* KO alleles. **a** Amino acid sequences of the wild-type nuclear localization signal (NLS) in Xap5 and the NLS-deleted mutant (ΔNLS) generated for this study. Key basic residues targeted for mutation are highlighted in red. **b** Validation of CRISPR/Cas9-mediated knockout of *Xap5*. Top: Schematic of the wild-type genomic locus showing the target sites for the two guide RNAs (sgRNA1 and sgRNA2, red) and their adjacent PAM sequences (blue). Bottom: Sanger sequencing chromatograms from a selected knockout clone revealing two distinct mutant alleles (allele-1 and allele-2). Both alleles share an identical mutation at the sgRNA2 target site but harbor different frameshift-inducing mutations at the sgRNA1 site, consistent with independent non-homologous end joining (NHEJ) repair events at this locus.

Accession	Description	Gene names
P62806	Histone H4	*Hist1h4a*
Q3TF40	Uncharacterized protein (Fragment)	*Nono*
Q9WV03	Protein Xap5	*Xap5*
P27661	Histone H2ax	*H2afx*
Q921L4	Histone H2b	*LOC665622*
P97350	Plakophilin-1	*Pkp1*
Q8VIJ6	Splicing factor, proline-and glutamine-rich	*Sfpq*
G3UXW9	COP9 signalosome complex subunit 1	*Gps1*

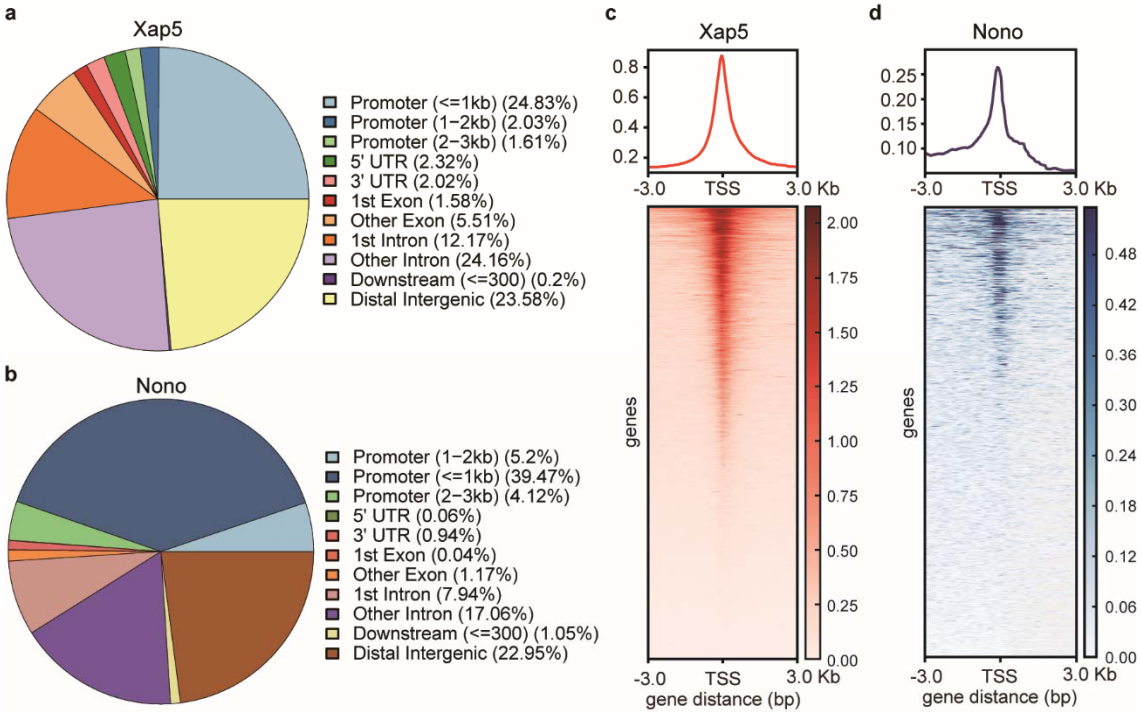
Supplementary Fig. 2 Identification of Xap5-interacting nuclear proteins by mass spectrometry. Table showing selected high-confidence nuclear proteins identified by immunoprecipitation (IP) of endogenous Xap5 from NIH/3T3 cell lysates followed by mass spectrometry (IP-MS). Notable hits include core histones (e.g., Hist1h4a, H2afx) and key RNA-binding proteins such as Nono and Sfpq, suggesting Xap5 may function in complexes involved in chromatin regulation and RNA processing.



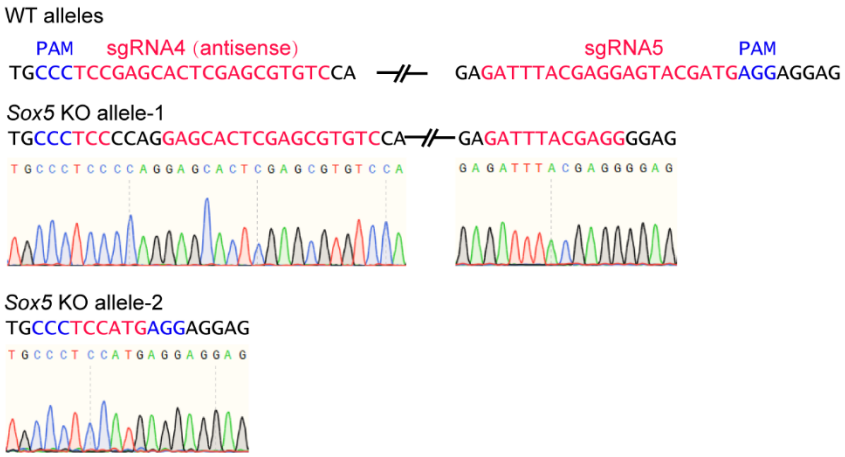
Supplementary Fig. 3 Validation of *Nono* KO and its impact on nuclear architecture. **a** Sanger sequencing chromatogram validating the CRISPR/Cas9-mediated KO of the *Nono* gene in a selected NIH/3T3 clone. The WT genomic sequence, sgRNA3 target site (red), and PAM sequence (blue) are indicated above. The resulting KO allele harbors a frameshift-inducing indel at the target site, consistent with successful gene disruption. **b** Quantification of nuclear area (μm²) in WT and *Nono* KO cells over a 48-h time course of serum starvation. Data are presented as the mean ± SEM (n = 3, independent biological replicates). Statistical significance between WT and KO cells at each time point was determined by multiple two-tailed unpaired t-tests. ***$P < 0.001$. Source data are provided as a Source Data file.



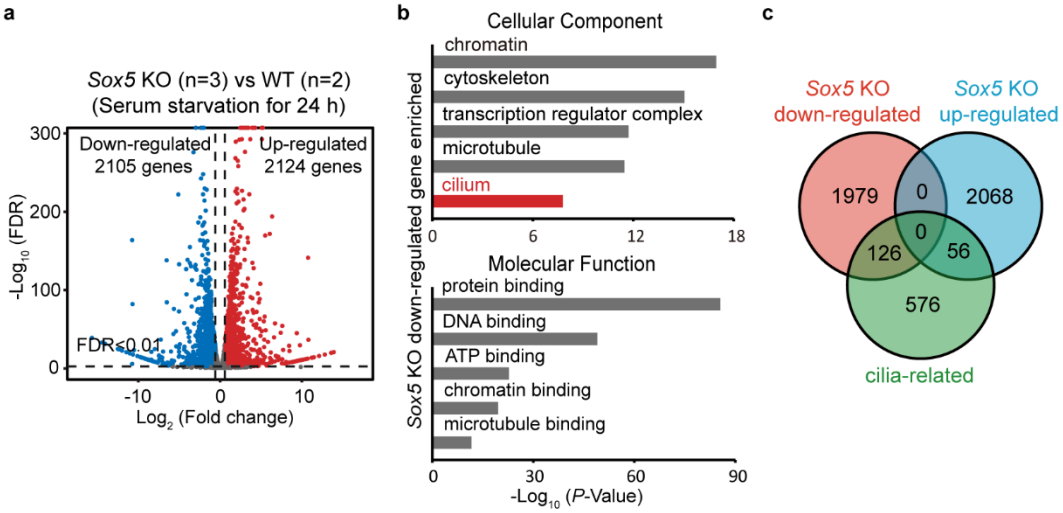
Supplementary Fig. 4 Transcriptomic profiling of *Xap5* and *Nono* KO cells. All analyses were performed on cells following 24 h of serum starvation. **a** Volcano plot of differentially expressed genes in *Xap5* KO (n=3 biological replicates) versus WT (n=2) cells. Using cutoffs of absolute fold change ≥ 1.5 and a false discovery rate (FDR) < 0.01 (highlighted dots), we identified 1,116 upregulated and 1,071 downregulated genes. **b** Volcano plot for *Nono* KO (n=3) versus WT (n=2) cells, analyzed using the same conditions and cutoffs as in **a**. This analysis identified 1,425 upregulated and 1,042 downregulated genes. **c**, **d** Representative enriched Gene Ontology (GO) terms for Cellular Component and Molecular Function categories among genes downregulated in *Xap5* KO (**c**) and *Nono* KO (**d**) cells. Cilia- and transcription-related terms are prominent in both datasets, reinforcing their dual roles in ciliogenesis and nuclear function.



Supplementary Fig. 5 Genomic distribution and metagene profiles of Xap5 and Nono CUT&Tag peaks. **a**, **b** Pie charts showing the genomic distribution of CUT&Tag peaks for Xap5 (**a**) and Nono (**b**). A substantial fraction of peaks for both Xap5 and Nono are located in promoter regions (defined as ≤3 kb upstream of a TSS). **c**, **d** Metagene plots (top) and heatmaps (bottom) showing the enrichment of Xap5 (**c**) and Nono (**d**) CUT&Tag signals centered on transcription start sites (TSS ±3 kb). The strong signal enrichment at the TSS is consistent with their roles as direct transcriptional regulators.



Supplementary Fig. 6 Validation of *Sox5* KO alleles by Sanger sequencing. Sanger sequencing chromatograms confirming CRISPR/Cas9-mediated deletion of *Sox5*. The wild-type genomic sequence, positions of the sgRNA4 (antisense strand) and sgRNA5 target sites (red), and the PAM sequences (blue) are indicated. Two independent *Sox5* KO alleles (KO allele-1 and KO allele-2) from the same clone display distinct mutations between the sgRNA cut sites. KO allele-1 harbors distinct frameshift-inducing mutations at each of the two sgRNA target sites, while KO allele-2 contains a large deletion spanning the entire region between the sgRNA sites. These results confirm successful biallelic disruption of the *Sox5* locus.



Supplementary Fig. 7 Transcriptomic analysis of *Sox5* KO cells. All analyses were performed on cells following 24 h of serum starvation. **a** Volcano plot of differentially expressed genes in *Sox5* KO (n=3 biological replicates) versus WT (n=2) cells. Using cutoffs of absolute fold change ≥ 1.5 and FDR < 0.01 (highlighted dots), we identified 2,124 upregulated and 2,105 downregulated genes. **b** Representative enriched GO terms for Cellular Component and Molecular Function categories among genes downregulated in *Sox5* KO cells. The significant enrichment of cilia-related terms reinforces the link between Sox5 and ciliogenesis. **c** Venn diagram illustrating the overlap between the curated ciliary gene set (from Supplementary Data 3) and genes either upregulated or downregulated in *Sox5* KO cells. The analysis reveals that a total of 182 ciliary genes were dysregulated, consisting of 126 downregulated and 56 upregulated genes, indicating a broad impact of Sox5 on the ciliary transcriptome.

Supplementary Table 1. Primers used in the study

Name	Sequences (5' to 3')	Usage
Xap5-Flag-F	GGGAATTCGTCGACTGGATCCATGGCCCAGTACAAGGGCG	Cloning primer
Xap5-Flag-R	TGAGATGAGTTTCTGCTCGAGCCGGATCGTGTACTTGTCCCA	
Nono-GFP-F	TCCAAGCTTCTGCAGGAATTCATGCAGAGCAATAAAGCCTTTAACT	
Nono-GFP-R	GCCCTTGCTCACCATTCTAGAATATCGGCGGCGTTTATTTG	
Sox5-RT-R	CCAGCCTTGTAGCTGAAACCA	qPCR primer
Sox9-RT-F	GAGCCGGATCTGAAGAGGGA	
Sox9-RT-R	GCTTGACGTGTGGCTTGTTC	
Adcy3-RT-F	AATGGCCTCAACGGCTCGGC	
Adcy3-RT-R	ATGAGGGCAGGGGAGCTGGG	
Ccdc39-RT-F	GAGCGAAGAACATTTTAAGGCCA	
Ccdc39-RT-R	TTCAAGTCATCCAGCTTTTGAGT	
Cdk20-RT-F	CAAAGCCAAGCACGTAGAGAC	
Cdk20-RT-R	TGCACCACATATTGACTGTCC	
Pdzd7-RT-F	TCTCTGGGGGCATCGAGTC	
Pdzd7-RT-R	GCTCAAAGCCAGCCTGTAGAT	
Pkd2-RT-F	GGGGAACAAGACTCATGGAAG	
Pkd2-RT-R	GCCGTAGGTCAAGATGCACAA	
Gapdh-RT-F	ATGGTGAAGGTCGGTGTGAAC	

Gapdh-RT-R	GCCTTGACTGTGCCGTTGAAT	
Sox5-CUT&Tag-F	TAGTACACTTCCAGCAGCGGT	
Sox5-CUT&Tag-R	GCCTGGAGGTTGTCTAATGCC	
Sox9-CUT&Tag-F	AAATCCGGTCCAATCAGCGA	
Sox9-CUT&Tag-R	ACGTTAGATACCTCGGGCCA	
DNA Spike in-CUT&Tag-F	GCCTTCTTCCCATTTCTGATCC	
DNA Spike in-CUT&Tag-R	CACGAATCAGCGGTAAAGGT	

Supplementary Table 2. sgRNAs used in the study

Name	Sequences (5' to 3')
Xap5-sgRNA1	GGCAAAAAAGGAGCAGTCAA
Xap5-sgRNA2	AGAGATGGCTATGTACGAGG
Nono-sgRNA3	AGTGGATCGGAACATCAAGG
Sox5-sgRNA4	GACACGCTCGAGTGCTCGGA
Sox5-sgRNA5	GATTTACGAGGAGTACGATG