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[1,5], Guangmei Tian[1,†], Zhangfeng Hu[1,3,4,†]

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

[4]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

[5]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

[#]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 displaying a high degree of conservation in basic morphology, eukaryotic cilia vary extensively in their sizes and functioning. Previously, we showed that an evolutionarily ancient transcription factor, X chromosome–associated protein 5 (Xap5) is required for motile ciliogenesis in unicellular and multicellular organisms. 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 factors networks determine ciliary diversity during evolution, and will have a profound impact on our understanding 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, 2, 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, 5, 6]. 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[7, 8, 9]. 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[10, 11, 12, 13], provides an important model for investigating this fundamental question, and its dysfunction can lead to a class of genetic disorders known as ciliopathies[14].

The assembly of the cilium, or ciliogenesis requires the precise transcriptional coordination of hundreds of genes[15, 16, 17]. In metazoans, the regulatory factor X (Rfx) and Forkhead box protein J1 (Foxj1) transcription factor families have been identified as key regulators of ciliogenesis[18, 19, 20, 21]. Rfx factors, such as DAF-19 in *C. elegans*, regulate the expression of genes essential for building the core ciliary structure[22], while Foxj1 directs the motile ciliogenic program[23, 24]. 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[11, 25].

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

69 fluid environments—a role conserved from the cilia of the unicellular alga *Chlamydomonas* to the
70 flagella of mammalian sperm[26, 27]. These findings cemented Xap5's role in an ancient program for
71 motility but, in doing so, created a profound biological enigma: why is this regulator, seemingly
72 dedicated to motility, highly expressed in somatic cells that exclusively form non-motile primary
73 cilia for sensory functions? Whether Xap5 plays a role in primary ciliogenesis in these cells, and
74 the mechanism by which it might be co-opted for this distinct function, remains to be determined[28].
75 In this study, we specifically address these questions by characterizing the function of Xap5 in
76 primary ciliogenesis using NIH/3T3 mouse fibroblasts, a standard model for primary cilium
77 formation upon serum starvation[29]. Here, we identify the Non-POU domain-containing octamer-
78 binding protein (Nono) as a critical interacting partner of Xap5. Our integrated analyses
79 demonstrate that Xap5 recruits its cofactor Nono to regulate primary cilia formation by modulating
80 a downstream transcriptional axis involving Sox5 and Sox9. While Sox9 has been previously
81 implicated in primary ciliogenesis[30, 31, 32], our findings for the first time identify Xap5 as a critical
82 upstream activator of these transcription factors in somatic cells. Together, these findings define
83 the complete Xap5-Nono-Sox5/9 axis for primary ciliogenesis. Moreover, they offer a powerful
84 illustration of evolutionary co-option in action, establishing a mechanistic framework for how
85 context-specific partnerships empower ancient transcription factors to drive cellular innovation.
86
87 **Results**
88 **Xap5 is essential for primary cilia formation**
89 To investigate the role of Xap5 in primary cilia formation, we first determined its subcellular
90 localization in NIH/3T3 cells. Immunofluorescence staining with a specific anti-Xap5 antibody
91 revealed that endogenous Xap5 is predominantly nuclear (Fig. 1a), consistent with our previous

92 work. We confirmed this finding via ectopic expression of a Flag-tagged Xap5 construct (Fig. 1b).
93 Consistent with its nuclear localization, bioinformatic analysis predicted a putative nuclear
94 localization signal (NLS) within the Xap5 protein (Supplementary Fig. 1a). To functionally
95 validate this sequence, we generated a site-directed mutant (Flag-Xap5-ΔNLS) (Supplementary
96 Fig. 1a). Crucially, this mutation abolished the nuclear import of Xap5, leading to its
97 mislocalization in the cytoplasm and confirming that this NLS is essential for its function (Fig.
98 1b).

99 To investigate the function of Xap5, we generated *Xap5* knockout (KO) NIH/3T3 cell lines
100 using the CRISPR/Cas9 system[33] (Supplementary Fig. 1b). The complete depletion of Xap5
101 protein product was confirmed by western blotting (Fig. 1c). Upon inducing ciliogenesis via serum
102 starvation (Fig. 1d), we first observed a striking defect in ciliary length (Fig. 1e). Although
103 comparable to wild-type (WT) cells before starvation, *Xap5* KO cells exhibited a significant
104 shortening of their primary cilia post-starvation (Fig. 1e, f). To determine if this length defect
105 stemmed from an overall failure in ciliogenesis, we next quantified the ciliation frequency.
106 Interestingly, both WT and *Xap5* KO cells responded robustly to the starvation stimulus, showing
107 a comparable increase in the percentage of ciliated cells over 48 hours (Fig. 1g). These results
108 collectively demonstrate that Xap5 is dispensable for the initiation of ciliogenesis but plays an
109 essential role in the subsequent elongation and/or maintenance of the ciliary axoneme.

110 **Xap5 forms a nuclear regulatory module with its cofactor Nono**

111 Having established the essential role of Xap5 in primary ciliogenesis, we next sought to uncover
112 its underlying regulatory mechanism by identifying its synergistic cofactors. An unbiased
113 immunoprecipitation-mass spectrometry (IP-MS) screen for Xap5 interactors identified several
114 nuclear proteins (Supplementary Fig. 2), among which we prioritized Non-POU domain-

115 containing octamer-binding protein (Nono), given its well-established, pleiotropic roles in
116 transcriptional regulation and RNA processing[34, 35, 36].

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

125 **The Xap5-Nono module is functionally required for ciliogenesis**

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

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

136 To determine if the Xap5-Nono module cooperates at the transcriptional level, we performed
137 RNA-sequencing (RNA-seq) on WT, *Xap5* KO, and *Nono* KO cells following 24 hours of serum

138 starvation, a time point where the phenotype was robustly established. This analysis revealed
139 widespread transcriptional changes; *Xap5* KO cells showed 1,116 upregulated and 1,071
140 downregulated genes, while *Nono* KO cells had 1,425 upregulated and 1,042 downregulated genes
141 relative to WT (fold change ≥ 1.5; Supplementary Fig. 4a, b; Supplementary Data 1, 2).

142 Given that loss of either Xap5 or Nono impairs ciliogenesis, we hypothesized that they function
143 together as transcriptional activators of a shared ciliogenic program. We therefore focused our
144 subsequent analysis on the genes commonly downregulated in both knockouts. Notably, Gene
145 Ontology (GO) analysis of these genes revealed a highly significant enrichment for ciliary
146 categories (Fig. 4a, b; Supplementary Fig. 4c, d). To investigate this further, we compared this
147 downregulated gene set against a comprehensive ciliary gene list (Supplementary Data 3), which
148 was assembled from the SYSCILIA gold standard (SCGSv2)[17] and manual literature curation. This
149 comparison confirmed a significant overlap between the two gene sets (Fig. 4c). We validated
150 several of these shared targets—including the transcription factors *Sox5* and *Sox9*, and the ciliary
151 component *Pkd2*—by qPCR (Fig. 4d, e).

152 Among this cohort of downregulated genes, two transcription factors, Sox5 and Sox9, stood out.
153 Sox9 is a known regulator of primary cilia in other contexts, and both factors are known to
154 cooperate in developmental programs. This made them compelling candidates for further
155 investigation as potential downstream effectors in our system. Consistent with this hypothesis,
156 western blot analysis confirmed that the protein levels of both Sox5 and Sox9 were markedly
157 reduced in *Xap5* and *Nono* KO cells (Fig. 4f). These data strongly indicate that Xap5 and Nono
158 co-regulate a transcriptional network that includes the downstream effectors Sox5 and Sox9,
159 thereby driving primary cilia formation.

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

161 To determine if Xap5 and Nono directly regulate their target genes, we mapped their genome-wide
162 chromatin occupancy using CUT&Tag[37] in NIH/3T3 cells (Supplementary Data 4, 5). This
163 analysis revealed a strong enrichment of both factors at promoter regions; approximately 30% of
164 Xap5 peaks and a striking 50% of Nono peaks were located at promoters (Supplementary Fig. 5a,
165 b). Consistent with this, genomic heatmaps demonstrated that both Xap5 and Nono binding signals
166 were highly enriched around the transcription start sites of their target genes (Fig. 5a, b;
167 Supplementary Fig. 5c, d).

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

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

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

181 Our finding that Xap5 and Nono directly co-regulate Sox5 and Sox9 prompted us to investigate
182 their respective roles in ciliogenesis. Given that the essential role of Sox9 in maintaining primary
183 cilia length is well-documented[30, 32], we focused our functional analysis on the uncharacterized

184 role of Sox5. To this end, we generated *Sox5* KO NIH/3T3 cells using the CRISPR/Cas9 system
185 (Fig. 6a; Supplementary Fig. 6).

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

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

195 To confirm that Sox5 functions transcriptionally downstream of the Xap5-Nono complex, we
196 performed RNA-seq on *Sox5* KO and WT cells after 24 hours of serum starvation. The analysis
197 identified 2,124 upregulated and 2,105 downregulated genes in *Sox5* KO cells (fold change ≥ 1.5;
198 Supplementary Fig. 7a; Supplementary Data 6). To assess the global impact on ciliary pathways,
199 we first performed Gene Set Enrichment Analysis (GSEA)[38] on all ranked differentially expressed
200 genes, which revealed a significant enrichment of the "Cilium organization" pathway (Fig. 7a).
201 Consistent with this, GO analysis of the strictly defined downregulated gene set (fold change ≥
202 1.5) also highlighted terms related to cilium assembly (Fig. 7b; Supplementary Fig. 7b).
203 Furthermore, direct comparison showed a significant overlap between the full set of genes
204 dysregulated by Sox5 loss and our curated ciliary gene list (Supplementary Fig. 7c), reinforcing
205 the strong link between Sox5 and ciliogenesis.

206 Critically, to place Sox5 within the Xap5-Nono pathway, we compared the downregulated gene

207 sets across all three KO lines. A Venn diagram revealed a significant and substantial overlap
208 among them (Fig. 7c). We confirmed this shared transcriptional dependency by qPCR, validating
209 that several key ciliary genes—including *Adcy3*, *Ccdc39*, and *Pkd2*—were indeed downregulated
210 across all three KO models (Fig. 7d). These results provide strong evidence that Sox5 acts as a key
211 transcriptional mediator downstream of Xap5 and Nono, executing a common program of ciliary
212 gene expression required for proper cilia elongation.
213

214 **Discussion**

215 The assembly of the cilium, a highly conserved organelle, requires a sophisticated and hierarchical
216 transcriptional program that is responsive to extracellular cues and precisely adapted to generate
217 enormous functional diversity[12, 18, 19, 20, 21, 25]. The transcriptional program's complexity is evident
218 in its role in regulating flagella in unicellular organisms like *Chlamydomonas reinhardtii* and
219 orchestrating the synchronous formation of hundreds of motile cilia in the mammalian airway[13, 39,
220 40, 41, 42]. In mammals, a core ciliogenic program is controlled by the Rfx family, which act as master
221 architects for the core ciliary structure[19, 43, 44, 45]. Built upon this foundation are multi-tiered
222 regulatory cascades, often initiated by morphogenetic signaling pathways, that deploy cell-type-
223 specific "selector cassettes"[13, 19]. For the highly specialized motile cilia, this cascade involves key
224 initiators like GMNC/MCIDAS (antagonized by GMNN)[46, 47, 48, 49], integrates context-specific
225 factors like NOTO[50, 51], and converges on downstream effectors TAp73/MYB[52, 53, 54], which in
226 turn activate the master regulator Foxj1 to execute the motile program[23, 24]. While this elegant
227 hierarchy explains the formation of multiciliated cells, how upstream signaling pathways initiate
228 the transcriptional programs for the vast diversity of primary cilia remains a key unanswered
229 question.

230 Central to this complexity is the ancient, evolutionarily conserved transcription factor Xap5.
231 Xap5 regulates flagellar assembly in *Chlamydomonas reinhardtii*, and is widely expressed in
232 mammalian tissues[26, 55, 56]. Our recent work revealed that during motile ciliogenesis in
233 spermatogenesis, Xap5 functions upstream of Rfx2 and Foxj1 to drive the sperm flagellar
234 program[27]. This finding solidified Xap5's role as a key regulator for motility but also sharpened
235 the central paradox addressed in this study: how does this same factor, which utilizes the
236 established motile cilia pathway in germ cells, function in somatic cells that build exclusively non-
237 motile primary cilia?

238 In this study, we resolve this fundamental paradox by uncovering a novel regulatory strategy
239 that enables Xap5's functional switch. Our findings demonstrate that in somatic cells, Xap5 is
240 repurposed by recruiting the nuclear protein Nono. The resulting Xap5-Nono complex forms a
241 regulatory module that regulates primary ciliogenesis by activating a downstream transcriptional
242 cascade involving Sox5 and Sox9. The discovery of this complete Xap5-Nono-Sox5/9 axis reveals
243 a parallel pathway for ciliogenesis and provides a powerful molecular blueprint for how a
244 conserved master regulator achieves functional diversity.

245 Our work identifies Sox5 and Sox9 as direct synergistic downstream effectors of the Xap5-
246 Nono module. The essential role of Sox9 in promoting primary cilia formation is well-established
247 in multiple developmental contexts[30, 31, 32]. Significantly, our finding that the Xap5-Nono complex
248 directly co-regulates Sox9 and Sox5 establishes Sox5 as a new, critical player in primary
249 ciliogenesis. This mechanism provides a compelling example of the "selector cassette" concept[19],
250 where cell-type-specific factors are deployed to modify a core developmental program, and we
251 propose that the Sox5/Sox9 partnership functions as such a cassette. This partnership is highly
252 plausible, given the known cooperation of Sox5 and Sox9 in other developmental programs (e.g.,

253 chondrogenesis[57]) and prior evidence linking Sox5 to ciliogenic genes in motile systems[58, 59]. Our
254 findings therefore position the Xap5-Nono module as an upstream integrator that co-opts this pair
255 of related Sox factors to regulate ciliogenesis.

256 Collectively, the discovery of this adaptable, Xap5-driven regulatory network has implications
257 for understanding tissue-specific manifestations of human ciliopathies. Given that Sox9
258 dysfunction is associated with cystic diseases characterized by ciliary defects, disruption of the
259 upstream Xap5-Nono module could represent a potential pathogenic mechanism. The
260 identification of upstream regulators for primary ciliogenesis is a rapidly advancing frontier, as
261 evidenced by the recent discovery of SP5/SP8 as key transcription factors in this process[60]. Unlike
262 the ancient Xap5 which is conserved from unicellular organisms, the SP family is metazoan-
263 specific[61], suggesting the evolution of multiple, distinct upstream regulatory layers for primary
264 ciliogenesis. Our identification of the distinct Xap5-Nono axis highlights the complexity of these
265 regulatory networks. Furthermore, our finding that systemic knockout of *Xap5* results in early
266 embryonic lethality, with lethality occurring prior to E7.5[27]—a phenotype potentially more severe
267 than that reported for SP5/SP8 knockouts—suggests a fundamental, and potentially earlier, role
268 for Xap5 in embryonic development. It will therefore be critical to identify the upstream signals
269 that govern the Xap5-Nono partnership and to determine whether Xap5 engages other partners to
270 drive its distinct functions in other tissues. Understanding this context-dependent interactome is
271 crucial for deciphering how ancient regulators like Xap5 are repurposed in different cellular
272 contexts, shedding light on their roles in development and disease.

273

274

275

276 **Methods**

277 **Antibodies**

278 The following primary antibodies were used in this study: from ABclonal Biotechnology, anti-
279 Flag (IF 1:1,000; AE005), anti-Gapdh (IB 1:10,000; AC002), anti-Nono (IB 1:1,000, IF 1:500;
280 A3800), and anti-Sox5 (IB 1:1,000; A23127); from Sigma-Aldrich, anti-Acetyl-α-Tubulin (IF
281 1:500; T7451); from Proteintech, anti-Arl13b (IF 1:500; 17711-1-AP); from HUABIO, anti-Sox9
282 (IB 1:1,000; ET1611-56); and from Prospertech, anti-Xap5 (IB 1:1,000, IF 1:500; HU-412003).

283 The following secondary antibodies were used: Alexa Fluor 488- and 594-conjugated goat anti-
284 mouse/rabbit IgG (IF 1:500; ThermoFisher Scientific); HRP-conjugated goat anti-mouse IgG (IB
285 1:2,000; Bio-Rad, #1706516) and HRP-conjugated goat anti-rabbit IgG (IB 1:2,000; Bio-Rad,
286 #1706515).

287 **Cell culture**

288 The NIH/3TT3 mouse embryonic fibroblast cell line (ATCC, CRL-1658) was cultured in
289 Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% calf serum (CS) and 1%
290 penicillin-streptomycin (all from Life Technologies). Cells were maintained in a humidified
291 incubator at 37°C and 5% CO_2.

292 **Plasmid construction, transfection and generation of knockout cell lines**

293 For protein expression, the coding sequences of mouse Xap5 and Nono were amplified from
294 NIH/3T3 cDNA and cloned into the pECMV-MCS-Flag and pCMV-C-GFP vectors, respectively,
295 to generate constructs encoding C-terminally tagged proteins. Site-directed mutagenesis to delete
296 the Xap5 nuclear localization signal (NLS) was performed using the Mut Express II Fast
297 Mutagenesis Kit V2 (Vazyme, C214). For CRISPR/Cas9-mediated gene knockout, the all-in-one
298 plasmid pU6-sgRosa26-1_CBh-Cas9-T2A-BFP (a gift from Ralf Kuehn; Addgene plasmid

299 #64216) was used as the backbone. Gene-specific sgRNA sequences were inserted by replacing
300 the original sgRosa26-1 guide RNA sequence via site-directed mutagenesis. The integrity of all
301 constructs was confirmed by Sanger sequencing. All primer and sgRNA sequences are listed in
302 Supplementary Tables 1 and 2, respectively.

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

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

314 **Immunofluorescence and microscopy**

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

319 Cells on coverslips were washed three times with PBS and fixed with 4% paraformaldehyde
320 (PFA) in PBS for 10 min at room temperature. Following fixation, residual PFA was quenched
321 with 50 mM NH_4Cl in PBS for 5 min. Cells were then permeabilized with 0.2% Triton X-100 for

322 15 min and blocked with 3% bovine serum albumin (BSA) in PBS for 1 h. Primary antibody
323 incubations were performed for 1 h at room temperature in a humidified chamber. After washing,
324 cells were incubated with Alexa Fluor-conjugated secondary antibodies for 1 h, followed by a 5
325 min counterstain with DAPI (Sigma-Aldrich). Finally, coverslips were washed and mounted onto
326 glass slides using Fluoromount-G mounting medium (SouthernBiotech). This protocol was used
327 for all cilia detection and Xap5 ectopic expression experiments. For the detection of ectopically
328 expressed Nono-GFP, cells were fixed and washed as above, then directly counterstained with
329 DAPI before mounting, omitting the permeabilization and antibody incubation steps.

330 General fluorescence for confirming ectopic protein expression was assessed using a Zeiss Axio
331 Vert. A1 inverted microscope. High-resolution imaging of primary cilia was performed on a Leica
332 SP8 confocal microscope using a 63× oil-immersion objective. The percentage of ciliated cells
333 and cilia length were quantified from confocal images using the Automated Cilia Detection and
334 Counter (ACDC) software (v0.93)[62]. The nuclear area was quantified from DAPI-stained images
335 using Fiji software.

336 **RNA isolation and RT-qPCR**

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

344 **Western blotting**

345 Cells were lysed on ice in RIPA buffer (Beyotime, P0013B) supplemented with 1× protease
346 inhibitor cocktail (Roche, 11873580001) and 1 mM PMSF (Sigma, 93482). The resulting lysates
347 were resolved by SDS-PAGE and transferred to polyvinylidene fluoride membranes (Millipore,
348 IPVH00010) using a wet transfer system (Tanon). Membranes were blocked with 5% non-fat milk
349 in TBST (20 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.1% Tween-20) for 1 h at room temperature,
350 followed by incubation with primary antibodies overnight at 4 °C. After washing, membranes were
351 incubated with HRP-conjugated secondary antibodies for 1 h at room temperature.
352 Chemiluminescence signals were developed using the ECL Super Kit (ABClonal, RM02867) and
353 visualized with a Tanon 5200 Chemiluminescence Imaging System.

354 **Co-immunoprecipitation (Co-IP) and mass spectrometry**

355 For Co-IP and IP-MS experiments, cultured cells were washed twice with cold 1× PBS and
356 harvested by scraping. Total proteins were extracted on ice using a Co-IP buffer containing 50 mM
357 HEPES (pH 7.5), 200 mM KCl, 1 mM EGTA, 1 mM $MgCl_2$, 10% (v/v) glycerol, 0.5 mM DTT,
358 0.3% (v/v) NP-40, and 1× protease inhibitor cocktail (Roche). For immunoprecipitation,
359 antibodies were first crosslinked to Protein A/G magnetic beads (Invitrogen, 80105G) using 5 mM
360 BS^3. The resulting antibody-conjugated beads were then incubated with the cell lysates for 2 h at
361 4°C with gentle rotation. After extensive washing, bound proteins were eluted with 0.1 M Glycine-
362 HCl (pH 2.5). For western blot analysis, the eluates and input samples were neutralized, mixed
363 with SDS loading buffer, boiled for 5 min, and subsequently analyzed by immunoblotting. For
364 protein identification, the eluates were sent for analysis by liquid chromatography with tandem
365 mass spectrometry (LC-MS/MS) at Novogene (Beijing, China).

366 **RNA-sequencing and analysis**

367 For transcriptomic analysis, cells were plated for 24 h and subsequently serum-starved for 24 h.

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

372 The raw sequencing reads were mapped to the Mus musculus reference genome
373 (GRCm39/mm39) using HISAT2[63] (v2.1.0). StringTie[64] (v1.3.3b) was subsequently used to
374 assemble the mapped reads. Gene expression levels were quantified as fragments per kilobase of
375 transcript per million fragments mapped (FPKM). Differential expression analysis between
376 experimental groups was performed using DESeq2[65] (v1.12.4). Genes with a false discovery rate
377 (FDR) < 0.01 and an absolute fold change ≥ 1.5 were considered significantly differentially
378 expressed. Gene Ontology (GO) enrichment analysis was performed using the DAVID database.
379 Gene Set Enrichment Analysis (GSEA) was conducted using the OECloud platform
380 (https://cloud.oebiotech.com).

381 **CUT&Tag-sequencing and analysis**

382 For CUT&Tag analysis, NIH/3T3 cells were plated for 24 h and subsequently serum-starved for
383 24 h. Following harvesting, 100,000 cells were collected for each sample. Library construction
384 was performed using the Hyperactive Universal CUT&Tag Assay Kit for Illumina Pro (Vazyme,
385 TD904) according to the manufacturer's protocol. The resulting libraries were sequenced on an
386 Illumina NovaSeq platform (150 bp paired-end reads) at Novogene Co., Ltd. (Beijing, China).

387 For bioinformatic processing, raw reads were assessed for quality with FastQC[66] (v0.11.9) and
388 trimmed with fastp[67] (v0.20.0; parameters: --length_required 15, --n_base_limit 6). Clean reads
389 were aligned to the Mus musculus reference genome (mm39) using BWA-MEM (v0.7.12;
390 parameters: -k 32, -T 30, -M). Peak calling was performed using MACS2[68] (v2.1.0) with a q-value

391 threshold of 0.05 and the following parameters: -f AUTO --call-summits --nomodel --shift -100 -
392 -extsize 200 --keep-dup all. For *de novo* motif discovery, peak summits were extended to 500 bp
393 and analyzed using the HOMER[69] (v4.9.1) findMotifsGenome.pl script. Peak annotation and
394 association with the nearest genes were performed using the ChIPseeker[70] R package. Genome
395 browser tracks were generated using the Integrative Genomics Viewer (IGV, v2.18.1)[71].

396 **CUT&Tag-qPCR**

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

403 **Statistical analysis**

404 Statistical analysis was performed using GraphPad Prism 6.01 or 8.0.2. All data were presented as
405 mean ± SEM. Values of $P < 0.05$ were considered statistically significant. Statistical significance
406 between two groups was calculated using an unpaired, parametric, two-sided student's t test.

407 **Data availability**

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

413

414 **References**

415 1. True JR, Carroll SB. Gene co-option in physiological and morphological evolution. *Annu*
416 *Rev Cell Dev Biol* **18**, 53-80 (2002).
417
418 2. Peter IS, Davidson EH. Evolution of gene regulatory networks controlling body plan
419 development. *Cell* **144**, 970-985 (2011).
420
421 3. Davidson EH, Erwin DH. Gene regulatory networks and the evolution of animal body plans.
422 *Science* **311**, 796-800 (2006).
423
424 4. Iwafuchi-Doi M, Zaret KS. Cell fate control by pioneer transcription factors. *Development*
425 **143**, 1833-1837 (2016).
426
427 5. Zaret KS. Pioneer Transcription Factors Initiating Gene Network Changes. *Annu Rev*
428 *Genet* **54**, 367-385 (2020).
429
430 6. Whyte WA*, et al.* Master transcription factors and mediator establish super-enhancers at
431 key cell identity genes. *Cell* **153**, 307-319 (2013).
432
433 7. Carroll SB. Evo-devo and an expanding evolutionary synthesis: a genetic theory of
434 morphological evolution. *Cell* **134**, 25-36 (2008).
435
436 8. Wagner GP, Lynch VJ. Evolutionary novelties. *Curr Biol* **20**, R48-52 (2010).
437
438 9. Arendt D*, et al.* The origin and evolution of cell types. *Nat Rev Genet* **17**, 744-757 (2016).
439
440 10. Carvalho-Santos Z, Azimzadeh J, Pereira-Leal JB, Bettencourt-Dias M. Evolution: Tracing
441 the origins of centrioles, cilia, and flagella. *J Cell Biol* **194**, 165-175 (2011).
442
443 11. Goetz SC, Anderson KV. The primary cilium: a signalling centre during vertebrate
444 development. *Nat Rev Genet* **11**, 331-344 (2010).
445
446 12. Anvarian Z, Mykytyn K, Mukhopadhyay S, Pedersen LB, Christensen ST. Cellular
447 signalling by primary cilia in development, organ function and disease. *Nat Rev Nephrol*
448 **15**, 199-219 (2019).
449
450 13. Spassky N, Meunier A. The development and functions of multiciliated epithelia. *Nat Rev*
451 *Mol Cell Biol* **18**, 423-436 (2017).
452
453 14. Reiter JF, Leroux MR. Genes and molecular pathways underpinning ciliopathies. *Nat Rev*
454 *Mol Cell Biol* **18**, 533-547 (2017).
455
456 15. van Dam TJP*, et al.* CiliaCarta: An integrated and validated compendium of ciliary genes.
457 *PLoS One* **14**, e0216705 (2019).
458
459 16. van Dam TJ, Wheway G, Slaats GG, Group SS, Huynen MA, Giles RH. The SYSCILIA
460 gold standard (SCGSv1) of known ciliary components and its applications within a
461 systems biology consortium. *Cilia* **2**, 7 (2013).
462
463 17. Vasquez SSV, van Dam J, Wheway G. 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 (2021).

18. Thomas J, Morle L, Soulavie F, Laurencon A, Sagnol S, Durand B. Transcriptional control of genes involved in ciliogenesis: a first step in making cilia. *Biol Cell* **102**, 499-513 (2010).

19. Choksi SP, Lauter G, Swoboda P, Roy S. Switching on cilia: transcriptional networks regulating ciliogenesis. *Development* **141**, 1427-1441 (2014).

20. Lewis M, Stracker TH. Transcriptional regulation of multiciliated cell differentiation. *Semin Cell Dev Biol* **110**, 51-60 (2021).

21. Collins C, Ventrella R, Mitchell BJ. Building a ciliated epithelium: Transcriptional regulation and radial intercalation of multiciliated cells. *Curr Top Dev Biol* **145**, 3-39 (2021).

22. Swoboda P, Adler HT, Thomas JH. The RFX-type transcription factor DAF-19 regulates sensory neuron cilium formation in C. elegans. *Mol Cell* **5**, 411-421 (2000).

23. Yu X, Ng CP, Habacher H, Roy S. Foxj1 transcription factors are master regulators of the motile ciliogenic program. *Nat Genet* **40**, 1445-1453 (2008).

24. Stubbs JL, Oishi I, Izpisua Belmonte JC, Kintner C. The forkhead protein Foxj1 specifies node-like cilia in Xenopus and zebrafish embryos. *Nat Genet* **40**, 1454-1460 (2008).

25. Nachury MV, Mick DU. Establishing and regulating the composition of cilia for signal transduction. *Nat Rev Mol Cell Biol* **20**, 389-405 (2019).

26. Li L*, et al.* 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 (2018).

27. Wang W*, et al.* Control of ciliary transcriptional programs during spermatogenesis by antagonistic transcription factors. *Elife* **13**, (2025).

28. Brown JM. Cilia and transcription: a mini review. *Front Cell Dev Biol* **13**, 1582796 (2025).

29. Gomez AE, Christman AK, Van De Weghe JC, Finn M, Doherty D. Systematic analysis of cilia characteristics and Hedgehog signaling in five immortal cell lines. *PLoS One* **17**, e0266433 (2022).

30. Xu WP*, et al.* Deletion of Sox9 in the liver leads to hepatic cystogenesis in mice by transcriptionally downregulating Sec63. *J Pathol* **254**, 57-69 (2021).

31. Lodh S, O'Hare EA, Zaghloul NA. Primary cilia in pancreatic development and disease. *Birth Defects Res C Embryo Today* **102**, 139-158 (2014).

32. Hrncir HR, Goodloe B, Bombin S, Hogan CB, Jadi O, Gracz AD. Sox9 inhibits Activin A to promote biliary maturation and branching morphogenesis. *Nat Commun* **16**, 1667 (2025).

33. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* **8**, 2281-2308 (2013).

34. Knott GJ, Bond CS, Fox AH. The DBHS proteins SFPQ, NONO and PSPC1: a multipurpose molecular scaffold. *Nucleic Acids Res* **44**, 3989-4004 (2016).

35. Amelio AL*, et al.* A coactivator trap identifies NONO (p54nrb) as a component of the cAMP-signaling pathway. *Proc Natl Acad Sci U S A* **104**, 20314-20319 (2007).

36. Fox AH, Nakagawa S, Hirose T, Bond CS. Paraspeckles: Where Long Noncoding RNA Meets Phase Separation. *Trends Biochem Sci* **43**, 124-135 (2018).

37. Kaya-Okur HS*, et al.* CUT&Tag for efficient epigenomic profiling of small samples and single cells. *Nat Commun* **10**, 1930 (2019).

38. Subramanian A*, et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545-15550 (2005).

39. Marshall WF. Chlamydomonas as a model system to study cilia and flagella using genetics, biochemistry, and microscopy. *Front Cell Dev Biol* **12**, 1412641 (2024).

40. Lyu Q, Li Q, Zhou J, Zhao H. Formation and function of multiciliated cells. *J Cell Biol* **223**, (2024).

41. Brooks ER, Wallingford JB. Multiciliated cells. *Curr Biol* **24**, R973-982 (2014).

42. Rayamajhi D, Roy S. Multiciliated cells: Development, functions and disease relevance. *Semin Cell Dev Biol* **175**, 103660 (2025).

43. Lemeille S*, et al.* Interplay of RFX transcription factors 1, 2 and 3 in motile ciliogenesis. *Nucleic Acids Res* **48**, 9019-9036 (2020).

44. Piasecki BP, Burghoorn J, Swoboda P. Regulatory Factor X (RFX)-mediated transcriptional rewiring of ciliary genes in animals. *Proc Natl Acad Sci U S A* **107**, 12969-12974 (2010).

45. Coyle MC*, et al.* An RFX transcription factor regulates ciliogenesis in the closest living relatives of animals. *Curr Biol* **33**, 3747-3758 e3749 (2023).

46. Ma L, Quigley I, Omran H, Kintner C. Multicilin drives centriole biogenesis via E2f proteins. *Genes Dev* **28**, 1461-1471 (2014).

47. Stubbs JL, Vladar EK, Axelrod JD, Kintner C. Multicilin promotes centriole assembly and ciliogenesis during multiciliate cell differentiation. *Nat Cell Biol* **14**, 140-147 (2012).

48. Arbi M*, et al.* GemC1 controls multiciliogenesis in the airway epithelium. *EMBO Rep* **17**, 400-413 (2016).

49. Terre B*, et al.* GEMC1 is a critical regulator of multiciliated cell differentiation. *EMBO J* **35**, 942-960 (2016).

50. Beckers A, Alten L, Viebahn C, Andre P, Gossler A. The mouse homeobox gene Noto regulates node morphogenesis, notochordal ciliogenesis, and left right patterning. *Proc*

Natl Acad Sci U S A **104**, 15765-15770 (2007).

51. Alten L*, et al.* Differential regulation of node formation, nodal ciliogenesis and cilia positioning by Noto and Foxj1. *Development* **139**, 1276-1284 (2012).

52. Nemajerova A*, et al.* TAp73 is a central transcriptional regulator of airway multiciliogenesis. *Genes Dev* **30**, 1300-1312 (2016).

53. Marshall CB*, et al.* p73 Is Required for Multiciliogenesis and Regulates the Foxj1-Associated Gene Network. *Cell Rep* **14**, 2289-2300 (2016).

54. Tan FE*, et al.* Myb promotes centriole amplification and later steps of the multiciliogenesis program. *Development* **140**, 4277-4286 (2013).

55. Sedlacek Z*, et al.* 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 (1999).

56. Martin-Tryon EL, Harmer SL. XAP5 CIRCADIAN TIMEKEEPER coordinates light signals for proper timing of photomorphogenesis and the circadian clock in Arabidopsis. *Plant Cell* **20**, 1244-1259 (2008).

57. Liu CF, Lefebvre V. The transcription factors SOX9 and SOX5/SOX6 cooperate genome-wide through super-enhancers to drive chondrogenesis. *Nucleic Acids Res* **43**, 8183-8203 (2015).

58. Kiselak EA*, et al.* 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 (2010).

59. Zhang L*, et al.* Transcriptional regulation of human sperm-associated antigen 16 gene by S-SOX5. *BMC Mol Biol* **18**, 2 (2017).

60. Liang Y*, et al.* Transcription factors SP5 and SP8 drive primary cilia formation in mammalian embryos. *Science* **389**, eadt5663 (2025).

61. Presnell JS, Schnitzler CE, Browne WE. KLF/SP Transcription Factor Family Evolution: Expansion, Diversification, and Innovation in Eukaryotes. *Genome Biol Evol* **7**, 2289-2309 (2015).

62. Lauring MC, Zhu T, Luo W, Wu W, Yu F, Toomre D. New software for automated cilia detection in cells (ACDC). *Cilia* **8**, 1 (2019).

63. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods* **12**, 357-360 (2015).

64. Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat Biotechnol* **33**, 290-295 (2015).

65. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for

RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).

66. Andrews S. FastQC: A Quality Control Tool for High Throughput Sequence Data. *http://wwwbioinformaticsbabrahamacuk/projects/fastqc/*, (2010).

67. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884-i890 (2018).

68. Zhang Y*, et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137 (2008).

69. Heinz S*, et al.* Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* **38**, 576-589 (2010).

70. Yu G, Wang LG, He QY. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* **31**, 2382-2383 (2015).

71. Robinson JT*, et al.* Integrative genomics viewer. *Nat Biotechnol* **29**, 24-26 (2011).

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).

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.

651 **Competing interests**

652 The authors declare no competing interests.

653

654 **Supplementary information**

655 **Reporting Summary**

656 **Supplementary Data 1**

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

658 **Supplementary Data 2**

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

660 **Supplementary Data 3**

661 Curated list of cilia-associated genes used in the study

662 **Supplementary Data 4**

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

664 **Supplementary Data 5**

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

666 **Supplementary Data 6**

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

668

669

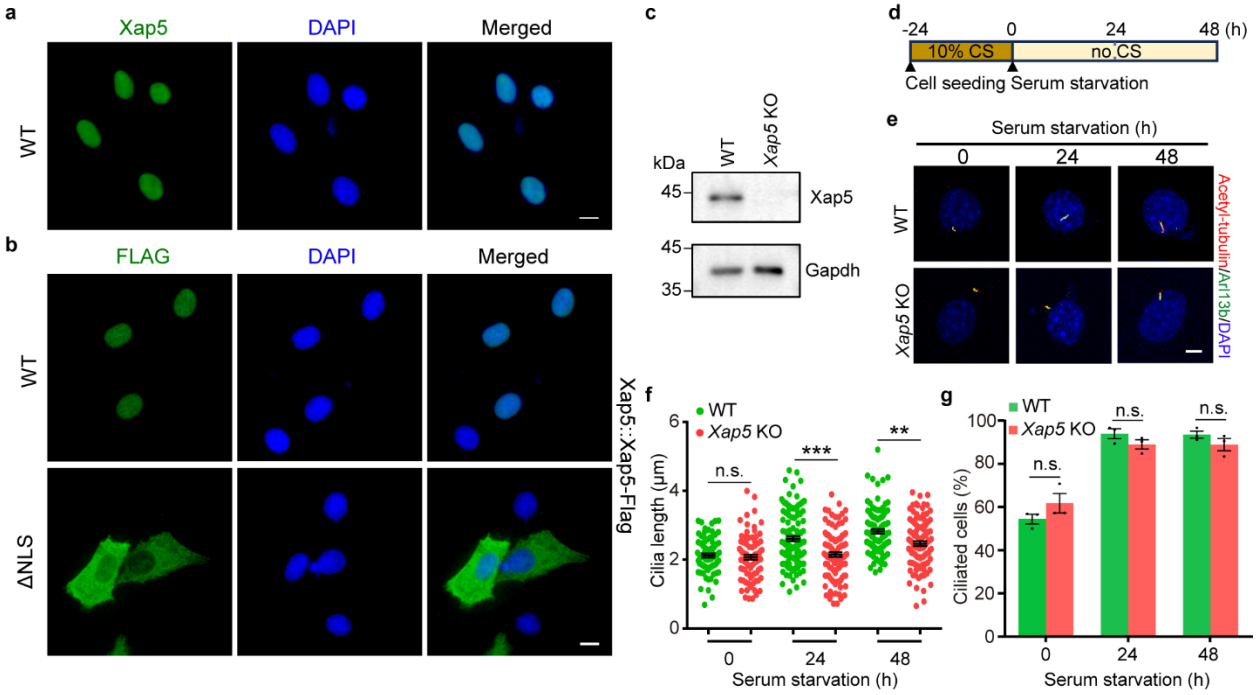
670

671

672

673

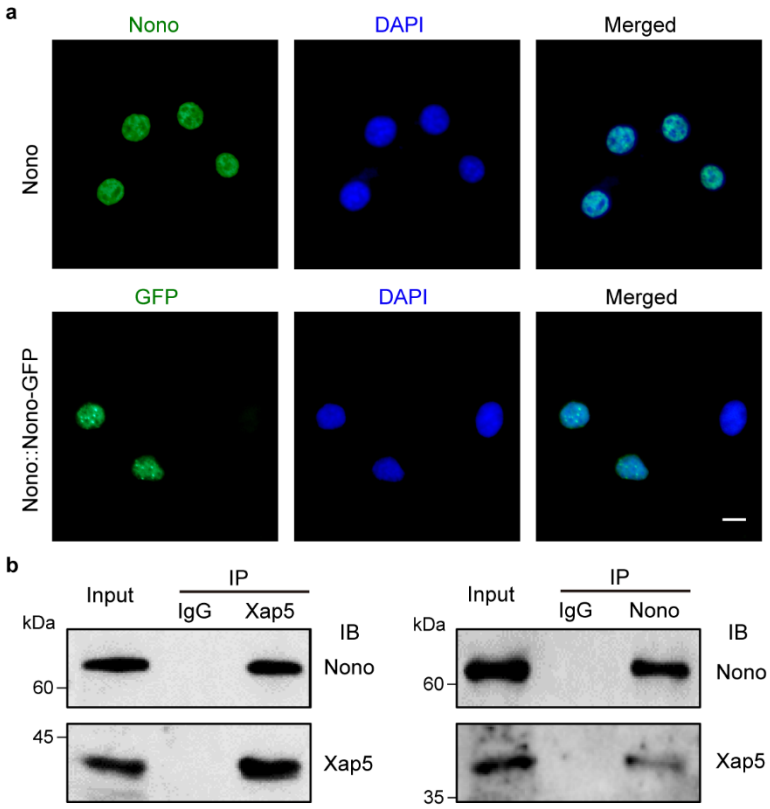
674 **Figures**



675

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

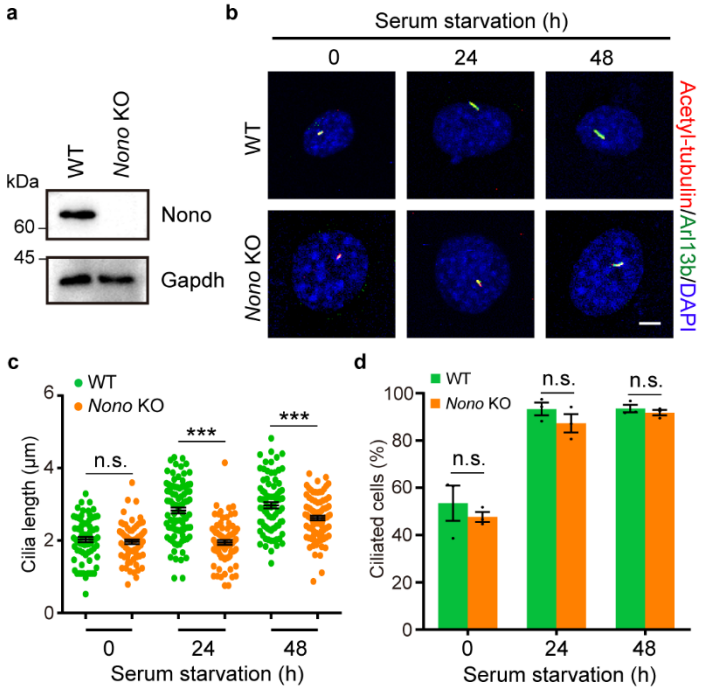
Data are presented as mean ± SEM (n = 3 independent experiments). All statistical significance between WT and KO at each time point was determined by multiple two-tailed unpaired t-tests. n.s., not significant; $^{**}P < 0.01$; $^{***}P < 0.001$. Source data are provided as a Source Data file.



703

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

715



716

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

729

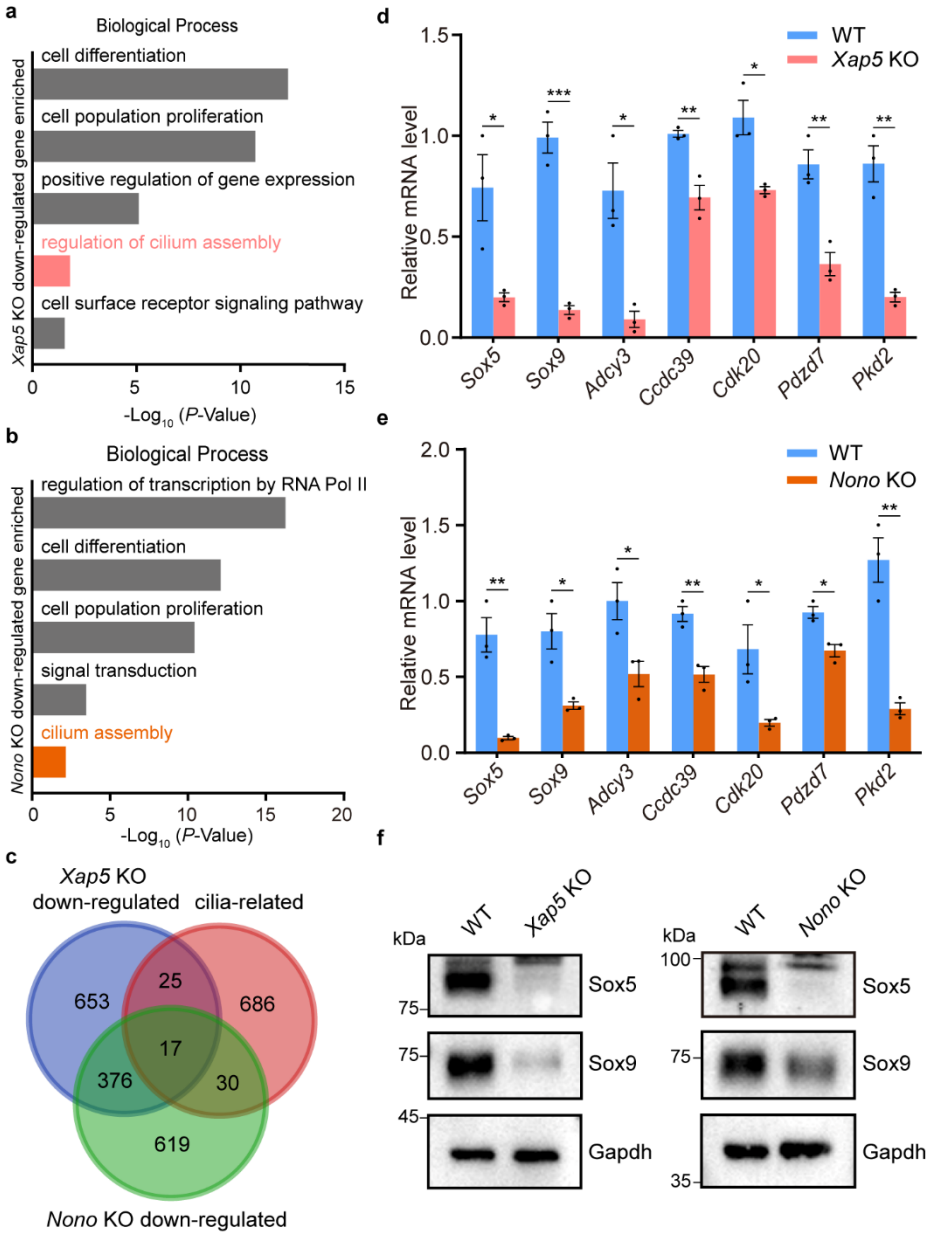


Fig. 4 Xap5 and Nono co-regulate a ciliogenic transcriptional network. All analyses were performed on NIH/3T3 cells following 24 h of serum starvation. **a, b** Representative enriched Gene Ontology (GO) terms (Biological Process) for genes downregulated in *Xap5* KO (**a**) and *Nono* KO (**b**) cells. **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 *Xap*5 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.

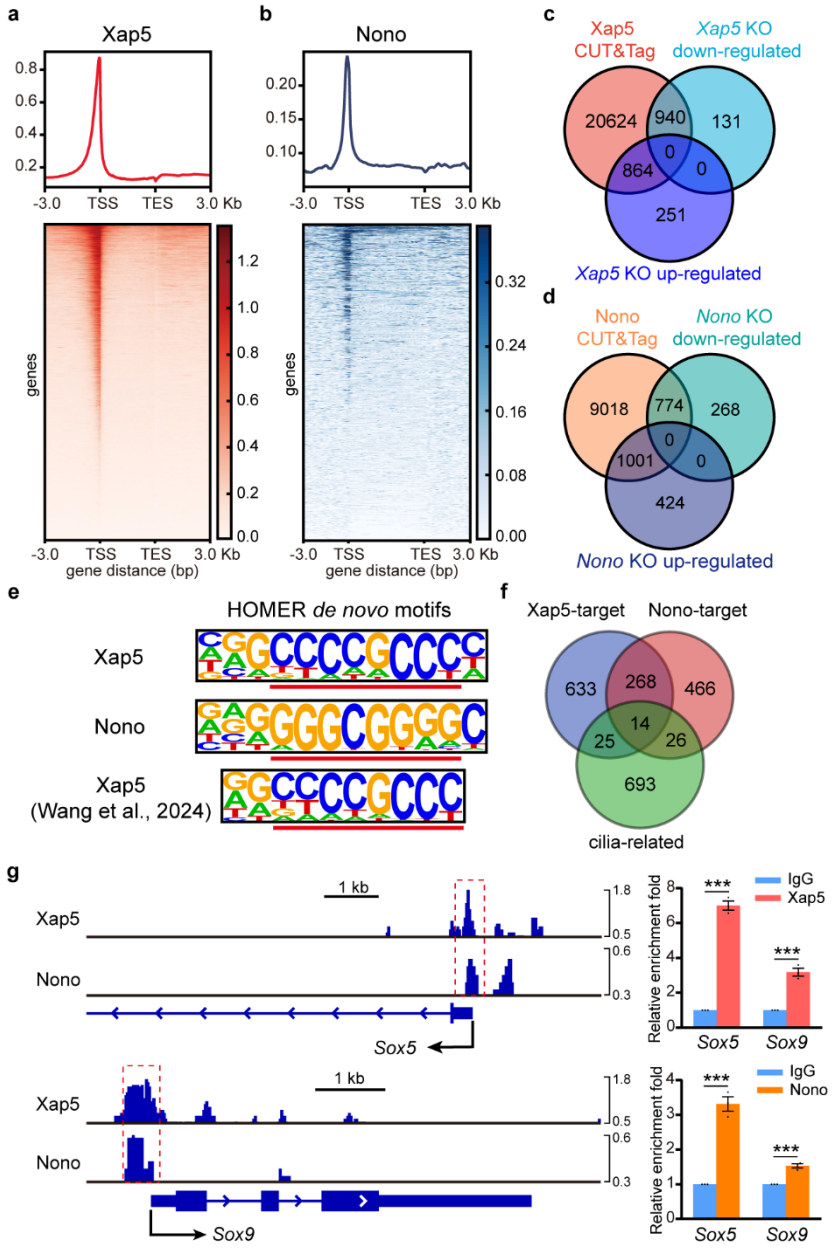


Fig. 5 Xap5 and Nono directly co-occupy the regulatory regions of ciliary genes. **a, b** Genome-wide binding profiles of Xap5 (**a**) and Nono (**b**) in NIH/3T3 cells determined by CUT&Tag. Top panels show the average enrichment profiles over gene bodies (from transcription start site, TSS, to transcription 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

767 knockout. **e** Comparison of *de novo* DNA-binding motifs derived from Xap5 and Nono CUT&Tag
768 peaks in this study, alongside the Xap5 motif from our previous work in testicular germ cells[27].
769 The high degree of similarity indicates a conserved DNA-binding preference across different
770 cellular contexts. **f** Venn diagram showing the overlap between direct Xap5 target genes (defined
771 as genes that are both bound by Xap5 and downregulated in *Xap5* KO), direct Nono target genes,
772 and the curated ciliary gene set (from Supplementary Data 3). **g** Direct co-occupancy of Xap5 and
773 Nono at the *Sox5* and *Sox9* gene loci. Left: Integrative Genomics Viewer (IGV) tracks showing
774 CUT&Tag signal enrichment. Right: CUT&Tag-qPCR validating the co-enrichment of Xap5 and
775 Nono at the *Sox5* and *Sox9* promoters. Red dashed box: Xap5 and Nono binding regions. Data are
776 presented as mean ± SEM from three independent experiments. *P* values were determined by two-
777 tailed unpaired t-test. ***$P < 0.001$. Source data are provided as a Source Data file.
778
779
780
781
782
783
784
785
786
787
788
789

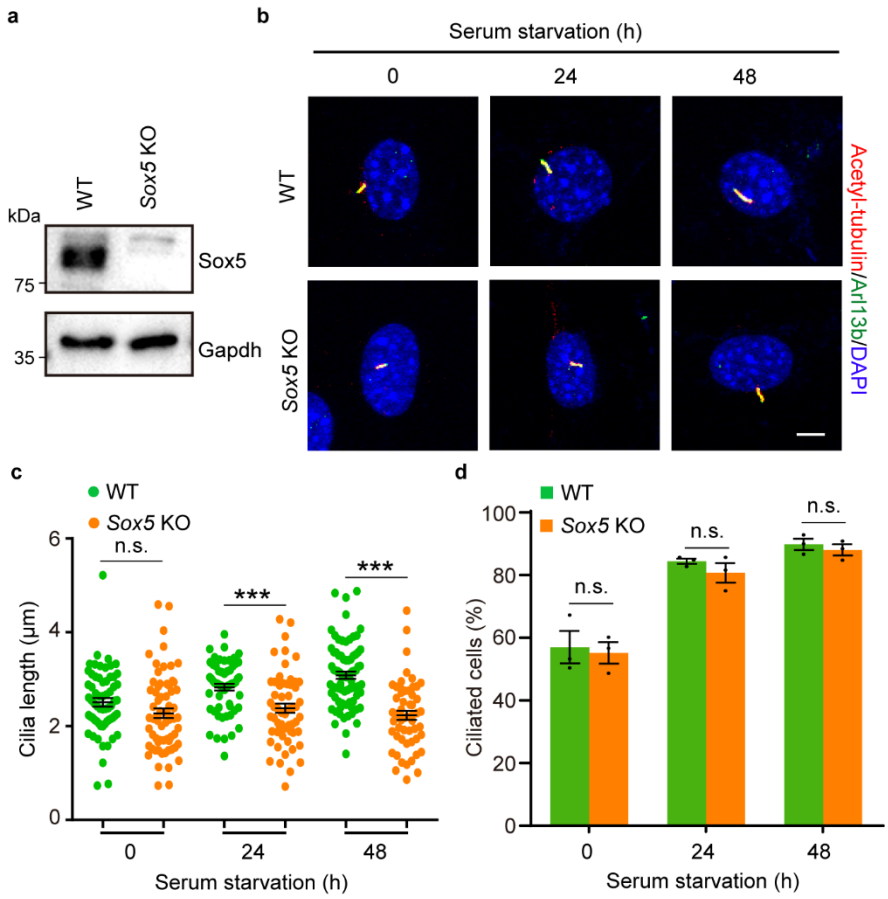


Fig. 6 Downstream effector Sox5 is essential for primary ciliogenesis. **a** Western blot analysis confirming the complete depletion of Sox5 protein in *Sox5* KO cell lysates compared to WT. The blot shown is representative of three independent biological replicates. Gapdh serves as a loading control. **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.

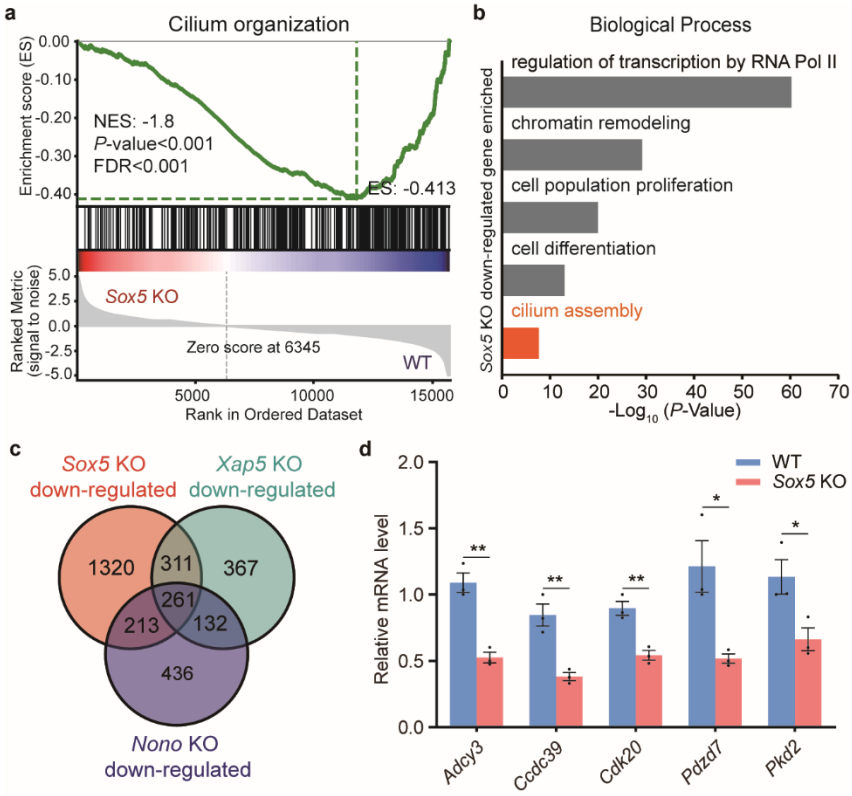
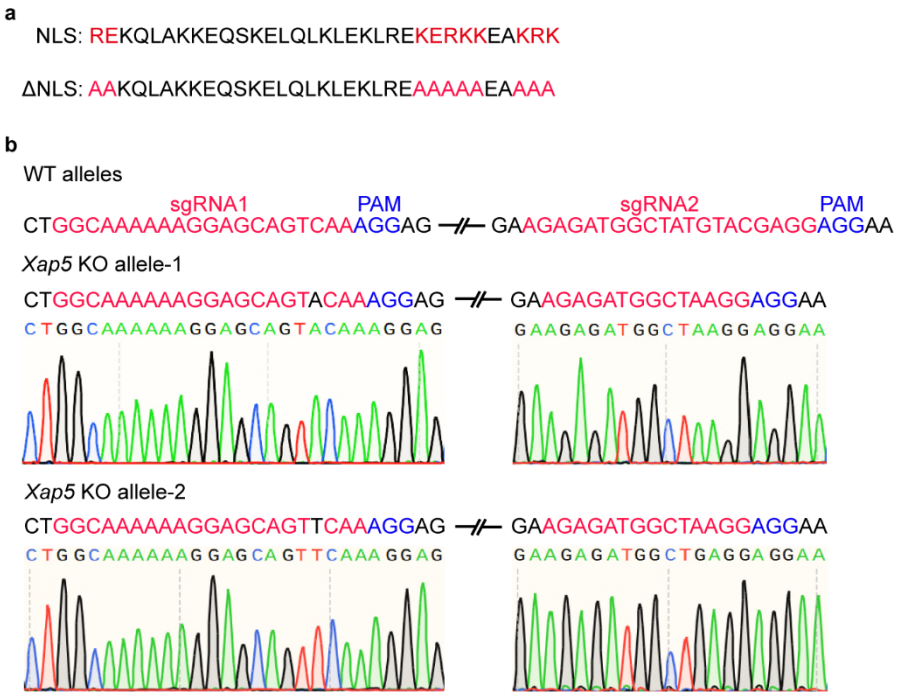


Fig. 7 Sox5 executes a ciliogenic transcriptional program downstream of the Xap5-Nono complex. **a** Gene Set Enrichment Analysis (GSEA) plot for the "cilium organization" pathway, performed on all ranked differentially expressed genes from the *Sox5* KO transcriptome. **b** Representative enriched GO 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.

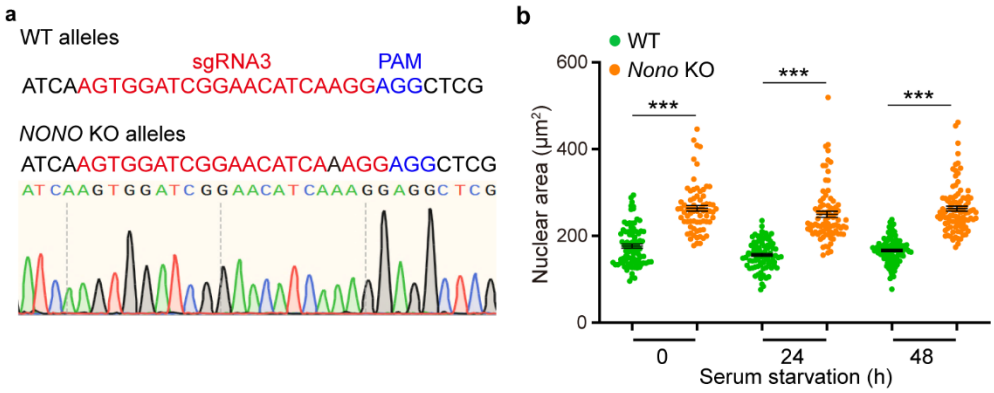
Supplementary Figures



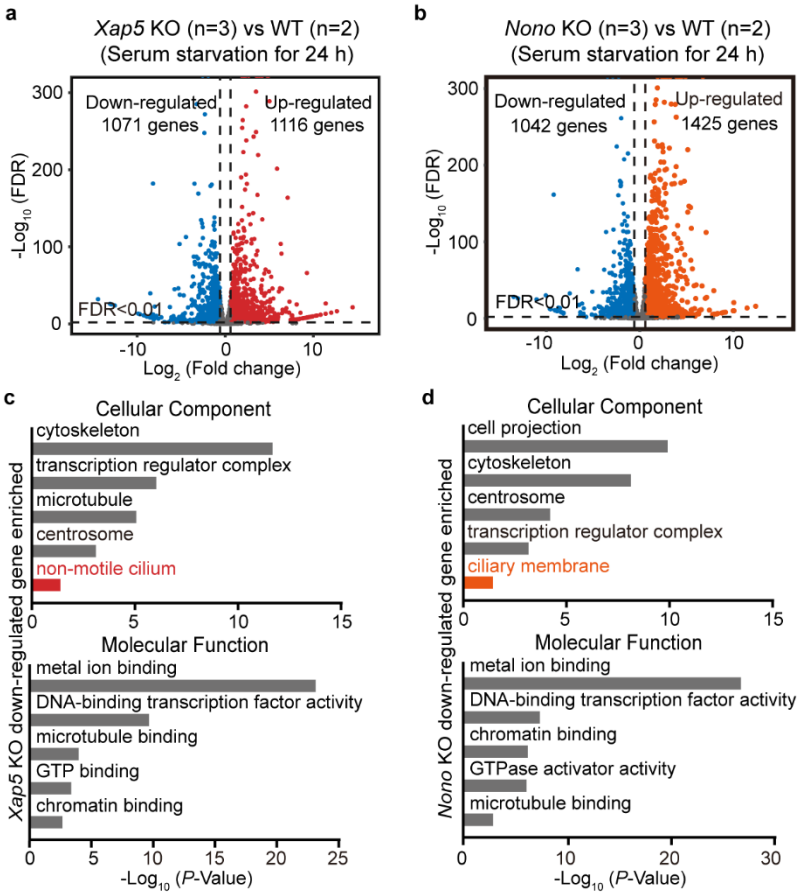
Supplementary Fig. 1 Design and validation of NLS-mutant and CRISPR/Cas9-mediated *Xap5* KO alleles. **a** Amino acid sequences of the wild-type nuclear localization signal (NLS) in Xap5 and the NLS-deleted mutant (ΔNLS) generated for this study. Key basic residues targeted for mutation are highlighted 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*

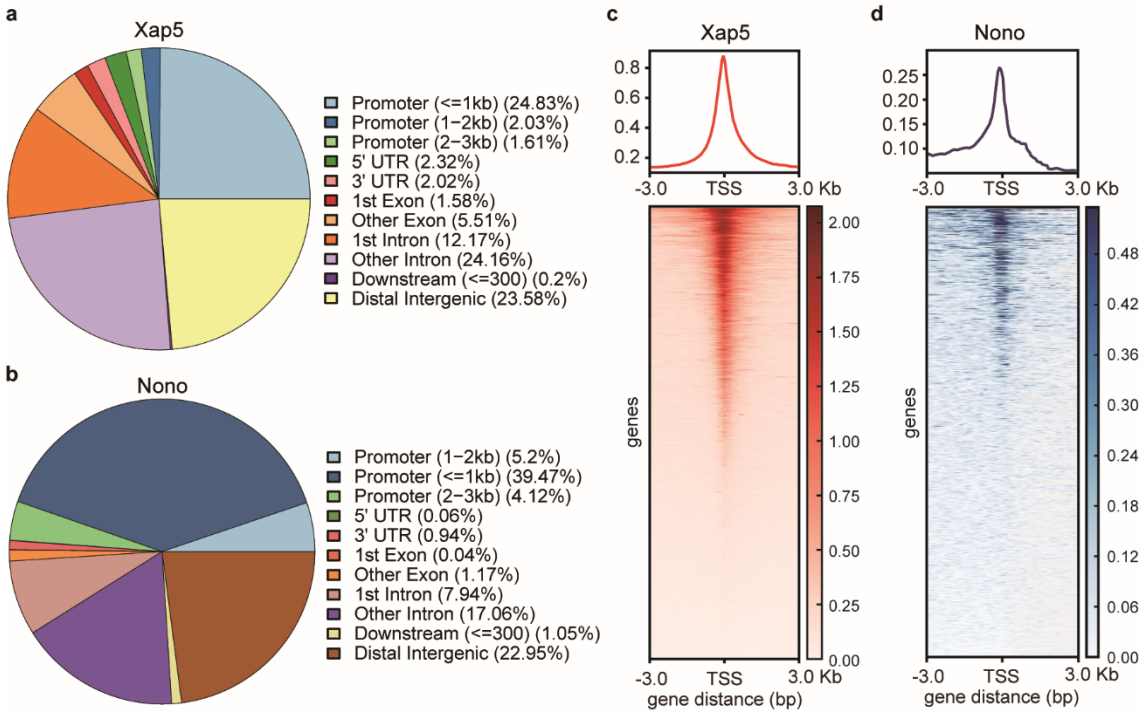
Supplementary Fig. 2 Identification of Xap5-interacting nuclear proteins by mass spectrometry. Table showing selected high-confidence nuclear proteins identified by immunoprecipitation (IP) of endogenous Xap5 from NIH/3T3 cell lysates followed by mass spectrometry (IP-MS). Notable hits include core histones (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.



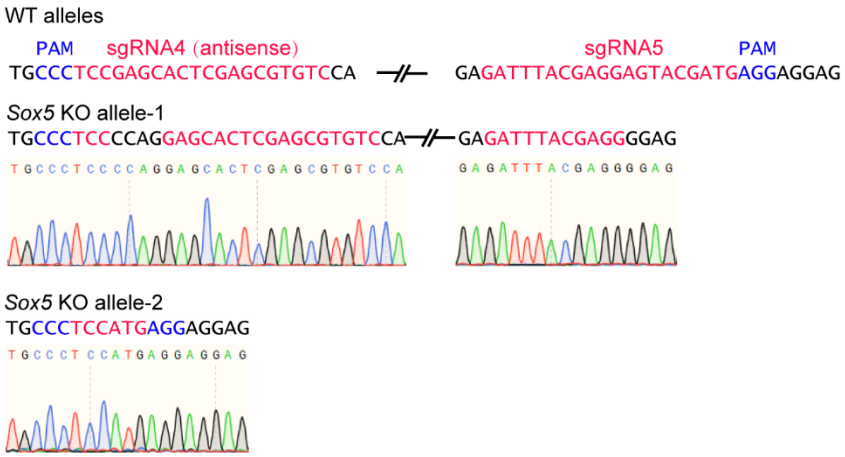
Supplementary Fig. 3 Validation of *Nono* KO and its impact on nuclear architecture. **a** Sanger sequencing chromatogram validating the CRISPR/Cas9-mediated KO of the *Nono* gene in a selected NIH/3T3 clone. The WT genomic sequence, sgRNA3 target site (red), and PAM sequence (blue) are indicated above. 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.



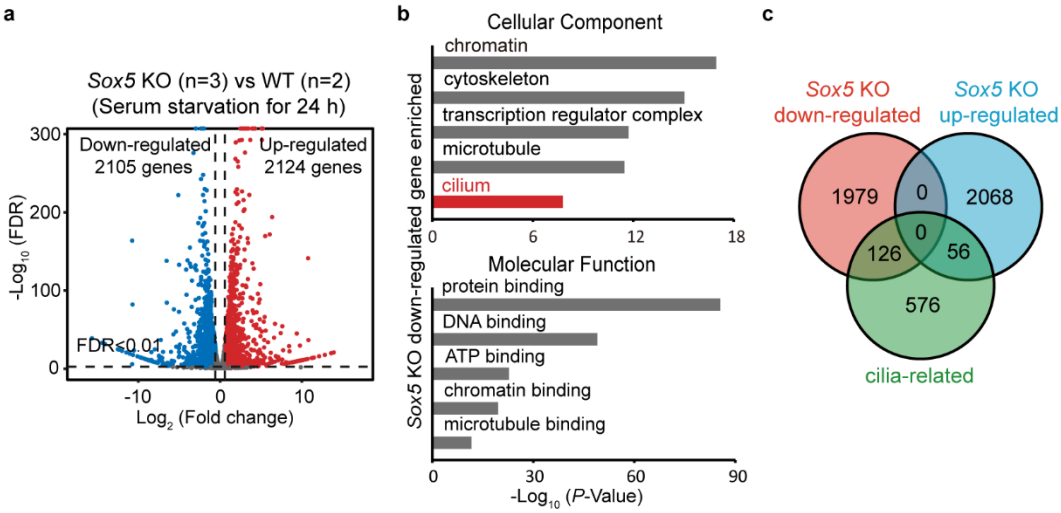
Supplementary Fig. 4 Transcriptomic profiling of *Xap5* and *Nono* KO cells. All analyses were performed on cells following 24 h of serum starvation. **a** Volcano plot of differentially expressed genes in *Xap5* KO (n=3 biological replicates) versus WT (n=2) cells. Using cutoffs of absolute fold change ≥ 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.

930 **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	

931

932

933

934

935

936

937

938

939

940

941

942

943

944

945

946

947 **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

948

949