Cohesin composition and dosage independently affect early development in zebrafish

Authors

Anastasia A. Labudina,[1] Michael Meier,[1] Gregory Gimenez,[1] David Tatarakis,[2,3] Sarada Ketharnathan,[1,4] Bridget Mackie,[1] Thomas F. Schilling,[2] Jisha Antony,[1] Julia A. Horsfield[1]*

Affiliations

1. Department of Pathology, Dunedin School of Medicine, University of Otago, P.O. Box 913, Dunedin 9016, New Zealand.
2. Department of Developmental and Cell Biology, University of California, Irvine, Irvine, CA 92697-2300, USA.
3. Present Address: Denali Therapeutics, 161 Oyster Point Blvd. South San Francisco, CA 94080, USA.
4. Present address: Children's Hospital of Eastern Ontario Research Institute, 401 Smyth Rd, Ottawa, Ottawa, ON K1H 5B2, Canada.

* To whom correspondence should be addressed: julia.horsfield@otago.ac.nz

Key words: cohesin, Rad21, Stag2, cell fate, zebrafish

22

23 **Summary Statement**

24 Viable zebrafish mutants show that cohesin complex quantity versus composition lead to
25 different transcriptional and developmental outcomes in the early embryo.

26

27 **Abstract**

28 Cohesin, a chromatin-associated protein complex with four core subunits (Smc1a, Smc3,
29 Rad21 and either Stag1 or 2), has a central role in cell proliferation and gene expression in
30 metazoans. Human developmental disorders termed "cohesinopathies" are characterised
31 by germline mutations in cohesin or its regulators that do not entirely eliminate cohesin
32 function. However, it is not clear if mutations in individual cohesin subunits have
33 independent developmental consequences. Here we show that zebrafish *rad21* or *stag2b*
34 mutants independently influence embryonic tailbud development. Both mutants have
35 altered mesoderm induction, but only homozygous or heterozygous *rad21* mutation affects
36 cell cycle gene expression. *stag2b* mutants have narrower notochords and reduced Wnt
37 signaling in neuromesodermal progenitors as revealed by single cell RNA-sequencing.
38 Stimulation of Wnt signaling rescues transcription and morphology in *stag2b*, but not
39 *rad21* mutants. Our results suggest that mutations altering the quantity versus composition
40 of cohesin have independent developmental consequences, with implications for the
41 understanding and management of cohesinopathies.

42

43

44

45 **MAIN TEXT**

46

47 **Introduction**

48 Cohesin is a multiprotein ring-shaped complex that is highly conserved from yeast to
49 humans. The vertebrate mitotic cohesin ring consists of two structural maintenance of
50 chromosomes subunits, Smc1a, Smc3 and an α-kleisin subunit Rad21 (Huis et al., 2014; Nasmyth
51 and Haering, 2009). In vertebrates Rad21 interacts with either one of two Stromalin subunits,
52 Stag1 or Stag2, and collectively these subunits are necessary for cohesin's association with DNA
53 (Dorsett and Strom, 2012; Gruber et al., 2003; Horsfield et al., 2012). Additionally, Nipbl and
54 Wapl modulate cohesin's residency on chromatin: Nipbl facilitates loading of cohesin onto DNA
55 (Ciosk et al., 2000), while Wapl facilitates its release (Kueng et al., 2006).

56 Cohesin is best known for its role in physically linking replicated sister chromatids to
57 ensure the accurate transmission of genetic material to daughter cells during cell division
58 (Nasmyth and Haering, 2009). In addition to mediating sister chromatin cohesion, the cohesin
59 complex also functions to repair DNA double strand breaks (Hou et al., 2022; Sjögren and
60 Nasmyth, 2001; Watrin and Peters, 2006). Loss of functional cohesin results in mitotic arrest and
61 cell death (Cukrov et al., 2018; Horsfield et al., 2007; Percival et al., 2015). Only a small fraction
62 of cohesin is necessary for sister chromatid cohesion (Gerlich et al., 2006) suggesting that the
63 observed high levels of cohesin in certain non-dividing cell types has important non-cell cycle
64 functions.

65 Cohesin also functions in three-dimensional (3D) genome organization and the regulation
66 of gene expression (Ball et al., 2014; Cuadrado and Losada, 2020; Dorsett and Strom, 2012;
67 Horsfield, 2023; Nishiyama, 2019; Zhu and Wang, 2019). Loop extrusion activity by cohesin
68 organizes DNA into topologically associated domains (TADs) that constrain the regulation of
69 gene expression (Davidson et al., 2019; Fudenberg et al., 2016; Hnisz et al., 2016; Krijger and de
70 Laat, 2016; Sanborn et al., 2015). The CCCTC-binding factor, CTCF, acts as a barrier to cohesin
71 and limits loop extrusion between convergent CTCF sites (Mayerova et al., 2020; Rao et al.,
72 2014), leading to the overlap of cohesin and CTCF at TAD boundaries (Dixon et al., 2012;
73 Parelho et al., 2008; Rao et al., 2014; Wendt et al., 2008). In addition, cohesin has gene regulatory
74 functions that are independent of CTCF (Meier et al., 2018; Schmidt et al., 2010). Sites bound by
75 cohesin but not CTCF are frequent at tissue-specific enhancers and promoters (Kagey et al.,
76 2010). Intra-TAD loops formed by cohesin can regulate transcription by mediating enhancer-

77 promoter contacts (Marsman and Horsfield, 2012; Ochi et al., 2020). However, only a subset of
78 enhancer-promoter contacts and DNA looping events appear to depend on cohesin (Friman et al.,
79 2023; Goel et al., 2023; Kane et al., 2022).

80 Germline cohesin insufficiency gives rise to a spectrum of multifactorial developmental
81 disorders collectively known as 'cohesinopathies' (Ball et al., 2014; Horsfield et al., 2012).
82 Typically, cohesinopathies result from heterozygous mutations in cohesin subunits or their
83 regulators (Horsfield et al., 2012). Cohesinopathies are associated with developmental delay, a
84 diverse range of developmental anomalies, and intellectual disability (Piché et al., 2019). The best
85 known cohesinopathy is Cornelia de Lange syndrome (CdLS, MIM #122470), a multisystem
86 disorder encompassing delayed growth, neurological and intellectual dysfunction, limb
87 abnormalities and gastrointestinal defects (Ireland et al., 1993; Jackson et al., 1993; Liu et al.,
88 2009; Opitz and Reynolds, 1985). Well over half of CdLS cases are caused by mutations in
89 *NIPBL* (Krantz et al., 2004; Tonkin et al., 2004), with mutations in other cohesin-associated
90 proteins accounting for a smaller subset of individuals with overlapping phenotypes. The specific
91 presentation of CdLS varies according to the cohesin-associated protein affected by genetic
92 changes (Cheung and Upton, 2015; Deardorff et al., 2020).

93 *RAD21* (MIM #606462) is among the five extensively studied genes associated with CdLS
94 (Deardorff et al., 2012a; Deardorff et al., 2007; Krantz et al., 2004). Individuals with *RAD21*
95 mutations display growth retardation, minor skeletal anomalies, and facial features that overlap
96 with CdLS, but lack severe intellectual disabilities (Deardorff et al., 2012b). Mutations in *RAD21*
97 are also linked with Mungan Syndrome (MIM #611376) (Mungan et al., 2003), sclerocornea
98 (Zhang et al., 2019) and holoprosencephaly (Kruszka et al., 2019). Most *RAD21* mutations
99 associated with cohesinopathy are truncations, missense mutations or in-frame deletions that are
100 predicted to interrupt the interaction between RAD21 and SMC1A, SMC3, or STAG1/2 (Cheng
101 et al., 2020; Krab et al., 2020). RAD21 physically bridges the SMC1A/SMC3 heads and
102 facilitates the cohesin loading process, likely by controlling the amount that complexes with DNA
103 (Sun et al., 2023). Therefore RAD21 abundance has potential to directly modulate the quantity of
104 cohesin complexes on DNA and its mutation or deficiency would result in reduction in cohesin
105 dose. Interestingly, the RAD21 protein must be intact for stable cohesin binding and looping at
106 CTCF-CTCF sites, and must be present but not necessarily intact for looped contacts inside of
107 CTCF domains (Liu and Dekker, 2022). Further supporting evidence suggests that the cohesion
108 and loop extrusion activities of cohesin can be separated experimentally and that cohesin uses
109 distinct mechanisms to perform these two functions (Nagasaka et al., 2023).

110 Individuals with STAG2 deficiency also display features of cohesinopathies (Cratsenberg
111 et al., 2021; Kruszka et al., 2019; Mullegama et al., 2017; Mullegama et al., 2019; Soardi et al.,
112 2017). Loss-of-function mutations in *STAG2* on the X chromosome are associated with
113 Mullegama-Klein-Martinez syndrome (MKMS, MIM #301022) in females but only missense
114 mutations are tolerated in males (Freyberger et al., 2021). Exome sequencing further established
115 *STAG1* and *STAG2* variants in patients with cohesinopathy phenotypes as loss-of-function (Yuan
116 et al., 2019) and recently, loss-of-function variants of *STAG2* have been categorized as X-linked
117 cohesinopathies with features of CdLS (Mullegama et al., 2017; Soardi et al., 2017). For example
118 an individual with a mosaic *STAG2* variant was described to have developmental delay,
119 microcephaly, and hemihypotrophy of the right side (Schmidt et al., 2022). A distinctive
120 cohesinopathy involving Xq25 microduplication that exclusively affects *STAG2* gives rise to
121 moderate intellectual disability, speech delay and facial dysmorphism (Gokce-Samar et al., 2022).
122 Additionally, some cases exhibit structural brain malformations consistent with
123 holoprosencephaly (Cratsenberg et al., 2021; Kruszka et al., 2019; Mullegama et al., 2017; Soardi
124 et al., 2017). Several molecular studies show that STAG1 and STAG2 paralogues have distinct
125 roles in 3D genome organization, but overlapping roles in the cell cycle (Cheng et al., 2022;
126 Cuadrado and Losada, 2020; Kojic et al., 2018; Porter et al., 2023; Remeseiro et al., 2012; Viny et
127 al., 2019). Moreover, STAG subunits can be detected at specific locations on DNA independently
128 of the rest of the cohesin complex (Pherson et al., 2019; Porter et al., 2023). Deficiency in STAG2
129 leads to the upregulation of STAG1 and the substitution of STAG1 for STAG2 in the cohesin
130 complex such that *STAG2* mutation leads to altered cohesin composition (Adane et al., 2021;
131 Bailey et al., 2021).

132 Dysregulated expression of multiple genes downstream of cohesin deficiency is thought to
133 be the predominant cause of cohesinopathies (Dorsett and Krantz, 2009; Horsfield et al., 2007;
134 Liu and Krantz, 2008; McNairn and Gerton, 2008; Muto et al., 2011). Because human
135 cohesinopathies with different genetic causes present with diverse phenotypes, it is possible that
136 cohesin subunits independently modulate the transcription function of cohesin during
137 development. This idea has not yet been tested in the early embryo when the developmental
138 changes in cohesinopathies are determined. In this study we compare the transcriptional and
139 developmental consequences of depleting Rad21 with depletion of Stag2. Rad21 controls cohesin
140 quantity on DNA (Sun et al., 2023) while Stag2 is thought to bind DNA independently and locate
141 cohesin to enhancers (Kojic et al., 2018; Pherson et al., 2019; Porter et al., 2023). Therefore, we
142 expect *stag2* mutants to interfere with cohesin's gene expression functions without interfering

143 with the cell cycle. Because *stag1b* and *stag2b* mutants are viable (Ketharnathan et al., 2020) and
144 the effects of *rad21* deficiency are dose-dependent (Schuster et al., 2015), zebrafish offer a unique
145 opportunity to investigate how cohesin complex quantity, versus cohesin complex composition,
146 affects cell fate decisions in the early embryo (Muto and Schilling, 2017). To explore this
147 question, we focused on the tailbud as a stem cell model.

148 The tailbud, located at the posterior end of the developing embryo, contains two
149 populations of bipotent stem cells known as neuromesodermal progenitors (NMPs) and midline
150 progenitor cells (MPCs) (Row et al., 2016; Steventon and Martin, 2022). These cells continuously
151 divide and differentiate into neuroectoderm and mesoderm by activating cell type-specific
152 transcription. By analysing *rad21* heterozygous and homozygous mutants (reflecting cohesin
153 dose) and *stag2b* mutants (reflecting cohesin type), we compare how the amount and composition
154 of the cohesin complex affect transcription in tailbud cells. We find that although *rad21*
155 heterozygous mutants are viable and fertile, they exhibit altered expression of thousands of genes
156 in the tailbud including cell cycle regulators, demonstrating that decreased cohesin dose affects
157 both cell cycle and gene expression. In contrast, cell cycle gene expression is largely unaffected in
158 *stag2b* homozygous mutants, which are also viable and fertile. However *stag2b* mutants show a
159 unique narrowing of the midline mesodermal domain that forms the notochord.

160 Therefore, although both *rad21* and *stag2* cohesin mutants show deficiencies in mesoderm
161 derived from NMPs and MPCs, the underlying molecular mechanisms are remarkably dissimilar.
162 Rad21 deficiency blocks NMP differentiation leading to lack of mesodermal derivatives, while
163 loss of *stag2* mutants causes NMPs to downregulate Wnt signaling leading to epithelial to
164 mesenchyme (EMT) defects. Changes in phenotype and gene expression unique to *stag2* mutants
165 are rescued by stimulation of Wnt signaling by GSK3 inhibition, to which Rad21-deficient
166 embryos are impervious.

167

168 **Results**

169

170 **Combined loss of zebrafish Stag1b and Stag2b phenocopies the null cohesin mutation *rad21***
171 Zebrafish have four Stag paralogues: Stag1a, Stag1b, Stag2a, and Stag2b. Individual *stag* mutant
172 lines (except *stag2a*) were previously generated and are homozygous viable (Ketharnathan et al.,
173 2020). To determine which paralogs are crucial for zebrafish development, we analysed the
174 consequences of combining *stag1a* and *stag2b* as well as *stag1b* and *stag2b* mutants.

175 *stag1b*$^{-/-}$; *stag2b*$^{-/-}$ double mutant embryos were developmentally delayed compared to
176 wild type, and by ~48 hours post-fertilization (hpf), mutant embryos had arrested in development
177 presenting with small heads, pericardial oedema, upwards bending tails, and no blood circulation
178 (Fig. 1B compared with A). This phenotype resembles *rad21*$^{-/-}$ mutant embryos, which die due to
179 mitotic catastrophe (Fig. 1C) (Horsfield et al., 2007). In contrast, *stag1a*$^{-/-}$; *stag2b*$^{-/-}$ embryos
180 developed normally and are homozygous viable and fertile, although a small proportion (~5%) of
181 *stag1a*$^{-/-}$; *stag2b*$^{-/-}$ embryos displayed hemorrhaging above the notochord at 48 hpf (Fig. S1A).

182 The gene encoding haematopoietic and neuronal transcription factor Runx1 is expressed in
183 the anterior lateral plate mesoderm (ALM), the posterior lateral plate mesoderm (PLM), and in
184 Rohon-Beard (RB) neurons in early zebrafish development (Fig. 1D) (Kalev-Zylinska et al.,
185 2002). Rad21 is required for *runx1* expression in the PLM (Horsfield et al., 2007). We previously
186 found that *runx1* expression is normal in individual *stag* mutants (Ketharnathan et al., 2020).
187 However, we observed loss of *runx1* expression in the PLM of the *stag1b*$^{-/-}$; *stag2b*$^{-/-}$ embryos and
188 retained *runx1* expression in the ALM and RB neurons (Fig. 1E). This resembles changes in
189 *runx1* expression in the *rad21*$^{-/-}$ mutant (Fig. 1F) (Horsfield et al., 2007), and is consistent with a
190 requirement for an intact cohesin complex for *runx1* expression in the PLM. In contrast, *runx1*
191 expression was normal in *stag1a*$^{-/-}$; *stag2b*$^{-/-}$ embryos (Fig. S1B).

192 We next examined the morphology of mitotic cells in *stag1b*$^{-/-}$; *stag2b*$^{-/-}$ embryos at 48 hpf
193 (Fig. 1G-J). In contrast to wild type embryos (Fig. 1G, I) condensed chromosomes were
194 disorganized and abnormally distributed in *stag1b*$^{-/-}$; *stag2b*$^{-/-}$ embryos, (Fig. 1H). Lagging
195 chromosomes failed to properly separate during anaphase, resulting in some chromosomes
196 remaining in cell centers (Fig. 1J). These findings suggest that cells in *stag1b*$^{-/-}$; *stag2b*$^{-/-}$ mutants
197 lack functional cohesin by 48 hpf, leading to a mitotic blockade. Individual *stag* mutants
198 (Ketharnathan et al., 2020) as well as the *stag1a*$^{-/-}$*; stag2b*$^{-/-}$ double mutant, are homozygous
199 viable. However, loss of both, *stag1b* and *stag2b* is embryonic lethal and phenocopies the
200 previously described *rad21*$^{-/-}$ mutant (Horsfield et al., 2007). Our results suggest that Stag1b and
201 Stag2b cannot be compensated by Stag1a and Stag2a proteins in zebrafish.

202

203 **Cell division proceeds normally in early stage cohesin mutant embryos**

204 Loss of cohesin in *rad21*$^{-/-}$ homozygotes or *stag1b*$^{-/-}$; *stag2b*$^{-/-}$ double mutants has different
205 effects on *runx1* expression compared with viable mutations in *stag* genes. Therefore, we were
206 curious to know whether cell cycle effects owing to cohesin deficiency could be responsible for

207 gene expression changes, including *runx1.* Mitotic catastrophe occurs in embryos lacking
208 functional cohesin at 48 hpf (Fig. 1H,J) (Horsfield et al., 2007). However, 16-somite *rad21*$^{-/-}$
209 homozygotes have sufficient maternally-deposited cohesin to continue growth for another 24
210 hours (Horsfield et al., 2007). We chose to compare cell cycle progression in *stag2b*$^{-/-}$
211 homozygotes with *rad21*$^{-/-}$ homozygotes and heterozygotes to determine if these mutants alter cell
212 cycle progression during early embryogenesis, the stage when *runx1* expression is disrupted.

213 Using BrdU incorporation to mark S phases and phosphorylated histone H3 (phH_3)
214 staining to mark G2/M cells, we found that S phase proceeds normally in cohesin mutants (Fig.
215 2A-F). Moreover, the presence of cells double-positive for BrdU and phH_3 indicated that cells
216 progressed from S to M phase in cohesin mutant tailbuds (Fig. 2D-F). Flow cytometry showed no
217 significant differences in the proportions of cells in G1 (2n), S (2-4n), and G2/M (4n) phases
218 between cohesin deficient tailbuds and wild type controls (Figs. 2G, S2) with the exception of
219 *rad21*$^{-/-}$ embryos, which had significantly reduced numbers of cells in S phase (Fig. S2F,
220 p=0.0317 Mann-Whitney U test). We conclude that cell division proceeds essentially normally,
221 although *rad21*$^{-/-}$ embryos had slightly reduced numbers of cells in S phase. This is consistent
222 with previous findings that even when cohesin complex quantity is substantially reduced, there
223 remains sufficient cohesin to progress through the cell cycle during early embryogenesis
224 (Horsfield et al., 2007; Schuster et al., 2015).
225

226 **Cohesin complex quantity and composition affect tailbud gene transcription differently**

227 The zebrafish embryonic tailbud contains neuromesodermal progenitors (NMPs) and
228 midline progenitor cells (MPCs) as well as their neural and mesodermal derivatives (Fig. 3A) and
229 therefore represents an ideal model to study changes in developmental gene transcription and cell
230 fate decisions. Because Stag2 (rather than Stag1) is most likely to be involved in tissue-specific
231 gene transcription (Casa et al., 2020; Cuadrado and Losada, 2020; Kojic et al., 2018; Viny et al.,
232 2019), we compared *stag2b* homozygous mutants with *rad21* homozygotes and heterozygotes to
233 determine how the type of cohesin subunit mutation affects transcription in tailbuds. We
234 performed bulk RNA-seq on 4 pools of 80 excised tailbuds from wild type, *rad21*$^{-/-}$, *rad21*$^{+/-}$, and
235 *stag2b*$^{-/-}$ embryos stage-matched at 16 somites. Principal Component Analysis (PCA) separated
236 samples into distinct groups based on their genotype (Fig. 3B). PC1 accounts for 61% of the
237 variance and separated samples into two groups: homozygous and heterozygous *rad21* mutants vs
238 wildtype and *stag2b*$^{-/-}$ mutants. PC2 accounts for an additional 11% of the variance and separated
239 *rad21* homozygotes from heterozygotes, and *stag2b* mutants from wild type.

240 Normalized transcript counts of the cohesin subunits in the different genotypes showed
241 that *rad21* mutation is associated with reduced transcript counts of the other cohesin core
242 subunits, *smc1a* and *smc3*, and increased transcript counts of *stag* subunits. In contrast,
243 transcription of core subunits was unaffected or increased in *stag2b* mutants, and *stag1b* transcript
244 counts increased (Fig. S3). The findings are consistent with the idea that *rad21* mutation reduces
245 cohesin quantity while *stag2b* mutation alters cohesin composition. Differential gene expression
246 analysis (Love et al., 2014) revealed that 7250 genes are dysregulated in *rad21* homozygotes (Fig.
247 3C), 5144 in *rad21* heterozygotes (Fig. 3D) and 2054 in *stag2b* homozygotes (Fig. 3E). Notably,
248 survivable changes in cohesin dose ($rad21^{+/-}$) and composition ($stag2b^{-/-}$) strongly affect
249 transcription in the tailbud, indicating that normal levels and subunit makeup of the cohesin
250 complex are important for gene expression.

251 Of the shared significantly dysregulated genes in cohesin mutant tailbuds, 311 were
252 upregulated and 312 were downregulated in all cohesin deficient tailbuds (Fig. S4A,B), with the
253 highest overlap between $rad21^{+/-}$ and $rad21^{-/-}$. Pathway enrichment analysis using Metascape
254 (Zhou et al., 2019) showed that muscle organ development and energy metabolism were
255 upregulated in all three genotypes, with the highest similarity between $rad21^{+/-}$ and $rad21^{-/-}$ (Fig.
256 S4C). Of the downregulated gene pathways, none were conserved across all three genotypes, and
257 more pathways were shared between $stag2b^{+/-}$ and $rad21^{-/-}$ than with $rad21^{+/-}$. A significant
258 number of terms were unique to $rad21^{-/-}$ tailbuds including regulation of cell fate specification,
259 suggesting possible dysregulation of tailbud progenitor differentiation (Fig. S4D). Pathway
260 enrichment analysis of significantly downregulated genes in $rad21^{-/-}$ tailbuds using Reactome (Yu
261 and He, 2016) revealed 26 significantly affected pathways, with top hits associated with mitosis
262 and DNA damage repair (Fig. S5). The most affected pathway was cell cycle control, with 165
263 genes significantly downregulated in $rad21^{-/-}$. Although this pathway did not reach a significance
264 threshold in other mutants, 79 cell cycle genes were significantly downregulated in $rad21^{+/-}$ and
265 13 in $stag2b^{-/-}$ (Data S1-3).

266 The number of shared dysregulated genes between genotypes suggests that transcriptional
267 changes in $rad21^{+/-}$ mutants more closely resemble $rad21^{-/-}$ than $stag2b^{-/-}$ mutants (Fig. S4A-D).
268 Transcriptional changes reflect genotype rather than viability through to adulthood: $rad21^{+/-}$ and
269 $stag2b^{-/-}$ mutants are viable and $rad21^{-/-}$ mutants are not. Additionally, the results suggest that,
270 consistent with the small effect on cell cycle progression in *rad21* mutants during early
271 embryogenesis (Fig. 2), strong transcriptional changes relate to the expression of cell cycle genes
272 in this genotype.

273

274 **Subunit-specific effects of cohesin deficiency on transcription in tailbud cell populations**

275 To assess how cohesin deficiency versus composition affects cell fate decisions in the
276 tailbud we used the bulk RNA-seq data to quantify the expression of genes that mark progenitor
277 cells and their derivatives. We used *rad21*$^{-/-}$ as a genotype that represents cohesin deficiency and
278 *stag2b*$^{-/-}$ as a genotype that corresponds to altered cohesin composition.

279 Tailbud NMPs give rise to mesoderm and neuronal fates, while MPCs give rise to
280 floorplate, notochord and hypochord (Fig. 3A). In *rad21*$^{-/-}$ tailbuds, genes marking NMPs were
281 upregulated, neural genes were dysregulated (both up- and downregulated) and genes marking all
282 mesoderm fates were downregulated, including lateral mesoderm that may not derive from NMPs
283 (Fig. 4A). In contrast, NMP and mesoderm marker genes were more subtly affected in *stag2b*$^{-/-}$
284 tailbuds, with non-significant downregulation of *tbxta*, *tbxtb*, *tbx16* and *msgn1* (see Fig. S6 for
285 additional expression data). *stag2b*$^{-/-}$ tailbuds had increased expression of genes that mark mature
286 somites and decreased expression of genes marking lateral mesoderm (Fig. 4B).

287 Genes that mark MPCs were expressed normally in *rad21*$^{-/-}$. However, genes encoding
288 notochord-specific collagens were upregulated, and some notochord markers were significantly
289 dysregulated (Fig. 4C). In contrast, genes expressed in MPCs and midline tissues derived from
290 MPCs were significantly downregulated in *stag2b*$^{-/-}$ tailbuds (Fig. 4D). The results suggest that
291 *rad21* deficiency causes a block in NMP differentiation, while *stag2b* mutation either affects the
292 composition of mesoderm, or mesoderm gene expression, in tailbuds. Moreover, *rad21* mutation
293 had little effect on midline progenitors (with some effect on MPC derivatives) while *stag2b*
294 mutation reduced transcription of genes expressed in MPCs and all derivatives.

295

296 ***rad21* and *stag2b* mutants have different tailbud phenotypes**

297 We next investigated if gene expression changes reflect gross developmental changes in
298 the tailbud in *rad21* and *stag2b* mutants by imaging tailbud cell populations. NMPs are marked
299 by *sox2* and *tbxta* co-expression (Fig. S7). MPCs, a thin band of cells at the end of the notochord,
300 also co-express *sox2* and *tbxta*. Mesoderm progenitors express *tbxta* but not *sox2*, and
301 differentiate into paraxial mesoderm, labelled by *tbx16* expression. *sox2* expression alone labels
302 neural progenitors, lateral mesoderm, the floor plate, and the hypochord (Fig. 3A; Fig. S7)
303 (Steventon and Martin, 2022; Thisse et al., 2001). We used HCR RNA-FISH to visualize the
304 distribution of *sox2*, *tbxta*, and *tbx16* transcripts in cohesin deficient tailbuds (Fig. 5).

305 In *rad21*$^{-/-}$ homozygotes (Fig. 5B compared with A) and in *stag1b*$^{-/-}$; *stag2b*$^{-/-}$ double
306 mutants (Fig. 5F compared with D) *sox2* expression was expanded at the posterior wall of the
307 tailbud, and the zone of *sox2* expression extended into mesoderm progenitors accompanied by a
308 reduction of *tbxta* expression in these cells. Expression of *tbx16* was restricted to a smaller area
309 than in wild type. Approximately two-thirds of heterozygous *rad21*$^{+/-}$ embryos displayed similar
310 expression changes, resembling homozygotes (Fig. 5C compared with A). Like *rad21* mutants,
311 *stag2b*$^{-/-}$ mutants had expanded *sox2* expression in the posterior wall of the tailbud (Fig. 5E
312 compared with D). However, in *stag2b*$^{-/-}$, *tbx16* expression appeared normal, while the notochord,
313 visualized by *tbxta* expression, was narrower and did not widen at the posterior end where MPCs
314 reside. Ectopic expression of *sox2* was also observed in this region (Fig. 5E).

315 We measured the thickness of the notochord (as defined by *tbxta* expression) in wild-type
316 and cohesin-deficient embryos (Fig. 6A,B) and confirmed that notochords were significantly
317 narrower ($p \leq 0.0001$) in *stag2b*$^{-/-}$ embryos (Fig. 6C). In contrast, notochord width in *rad21*
318 homozygous and heterozygous embryos was similar to wild type (Fig. 6C). Surprisingly, *stag1b*$^{-/-}$
319 ; *stag2b*$^{-/-}$ double mutant embryos had notochords that were normal width (Fig. 5F; Fig. 6C).
320 Therefore, the narrow notochord phenotype was unique to *stag2b*$^{-/-}$, suggesting that the loss of
321 Stag2b impacts MPC differentiation. Because the narrow notochord phenotype is absent in rad21
322 mutants and stag1b$^{-/-}$; *stag2b*$^{-/-}$ double mutants, complete cohesin loss is likely epistatic to the
323 narrowed notochord in *stag2b*$^{-/-}$ mutants.

324 Altogether, our results indicate that different cohesin mutations have different effects on
325 cell populations in the tailbud. Loss of cohesin quantity in *rad21* mutants and *stag1b*$^{-/-}$; *stag2b*$^{-/-}$
326 double mutants caused reduction of *tbx16* and expansion of *sox2* expression, consistent with lack
327 of mesoderm induction. In contrast, *stag2b* mutation (which alters cohesin composition) leads to a
328 narrower notochord.

329

330 **Altered cell populations in *stag2b*$^{-/-}$ tailbuds likely result from downregulated Wnt signaling**
331 **in NMPs**

332 It is possible that altered cohesin composition through *stag2b* mutation has unique
333 transcriptional effects on cell fate in the tailbud. We chose to investigate this possibility further
334 using single-cell RNA-sequencing of *stag2b*$^{-/-}$ tailbuds compared with wild type at the 16-somite
335 stage.

336 We integrated the single-cell RNA-seq datasets from wild type and *stag2b*$^{-/-}$ tailbuds and
337 annotated clusters representing major cell types based on their gene expression profiles (Fig.
338 S8A). All cell types were present in both wild type and *stag2b*$^{-/-}$ samples (Fig. 7A), with minor
339 changes in clusters reflecting altered expression of cell type-specific markers observed in the bulk
340 RNA-seq analysis (Fig. 4C,D). Cell numbers were in slightly different proportions for some
341 clusters (Fig. 7B), with the biggest change being increased numbers of 'anterior paraxial
342 mesoderm 1' cells in *stag2b*$^{-/-}$ tailbuds compared to wild type (Fig. 7B).

343 Curious as to why the anterior paraxial mesoderm would separate into two clusters despite
344 expressing the same lineage genes (Fig. S8A) we further quantified and plotted the expression of
345 cell cycle status genes (Fig. S9A-C). This analysis showed that 'anterior paraxial mesoderm 1' is
346 distinct from 'anterior paraxial mesoderm 2' by having more cells in G0/G1 and S phase, while
347 all cells are in either G2/M (>90%) or S phase in 'anterior paraxial mesoderm 2' (Fig. S9D-F). In
348 *stag2b*$^{-/-}$ 'anterior paraxial mesoderm 1' cells were more likely to be in G0/G1 phase with fewer
349 cells in S phase compared to wild type. This finding suggests that gene expression or cell
350 population changes in *stag2b* mutants may not be entirely independent of the cell cycle.

351 NMPs give rise to paraxial mesoderm and to determine if this process is disturbed in *stag2b*$^{-/-}$
352 embryos we performed pseudo-bulk differential gene expression and GSEA (Gene Set
353 Enrichment Analysis) on NMPs and mesoderm clusters (Fig. S10, Data S4). GSEA showed that
354 NMP clusters had downregulated pathways related to stem cell development (GO:0048864),
355 mesenchymal cell development (GO:0048864) and determination of left/right symmetry
356 (GO:0043534) (Fig. 7C, Data S4, S5). These processes are in part positively regulated by
357 canonical Wnt signalling.

358 We plotted expression levels of Wnt pathway genes across various cell types comparing
359 genotypes (Data S6) and found significant downregulation of Wnt signatures in the NMPs and
360 anterior paraxial mesoderm clusters (Fig. 7D, Data S7). *tbxta* and Wnt ligands (*wnt8a*, *wnt8a-1*
361 and *wnt11f2*) were downregulated in the NMPs, as well as Wnt-responsive genes *vox* and *eve1*
362 (Fig. 7E). In addition there was significant downregulation of *fgf8a* in NMPs suggesting that
363 epithelial-to-mesenchymal transition (EMT) may be deficient in *stag2b*$^{-/-}$ embryos (Goto et al.,
364 2017). We also noticed decreased expression of *bmp4* and *id3*; other studies have shown that loss
365 of these genes skews endothelial cell fate towards paraxial mesoderm (Row et al., 2018).
366 Together the results suggest that decreased Wnt signalling and a diversion from endothelial to
367 paraxial mesoderm fate could account for the increased paraxial mesoderm population in *stag2b*$^{-/-}$
368 embryos.

369 Consistent with the narrower notochord phenotype observed in *stag2b*$^{-/-}$ embryos, single cell
370 RNA-seq revealed altered gene expression and cell composition in the notochord cluster (Fig. 7F,
371 Fig. S11A, Data S8). Altered GSEA pathways in *stag2b*$^{-/-}$ notochords showed upregulation of
372 Hedgehog signalling, cell adhesion and adherens junctions pathways compared with wild type
373 (Fig. S11B Data S9). Interestingly we observed an increase in hypochord cells in *stag2b*$^{-/-}$ tailbuds
374 (Fig. S11C-E, Data S8), which could be occurring at the expense of notochord. Reduced
375 notochord and increased paraxial mesoderm cell numbers in *stag2b*$^{-/-}$ could additionally be caused
376 by downregulation of *noto* in *stag2b*$^{-/-}$ notochords (Fig. 7F). Consistent with this idea, loss of *noto*
377 causes a switch in cell fates from notochord to paraxial mesoderm in mouse (Yamanaka et al.,
378 2007) and zebrafish (Halpern et al., 1995).

379 In summary, our results suggest that paraxial and axial midline tissue formation from the NMPs is
380 dysregulated in *stag2b*$^{-/-}$ tailbuds. One explanation for these observed changes could be
381 downregulation of Wnt signalling in NMPs that give rise to these tissues.
382

383 **Wnt stimulation rescues transcription in *stag2b* but not *rad21* mutant tailbuds**

384 We next determined whether Wnt stimulation could restore transcription in cohesin
385 deficient tailbuds. We performed RNA-seq on tailbuds of embryos treated from shield stage to
386 16-somite stage with the Wnt agonist, BIO (6-bromoindirubin-3'-oxime), which is a GSK3
387 inhibitor. Subsequently, we conducted interaction analysis (combined effect of genotype and
388 treatment) (Love et al., 2014) to identify genes exhibiting differential responses to Wnt
389 stimulation in cohesin deficient tailbuds compared to wild type. Heatmaps were used to display
390 clustering of the differentially responsive genes (Fig. 8A-C).

391 In *rad21*$^{-/-}$ and *rad21*$^{+/-}$, the genotype had a stronger effect on clustering of differentially
392 responsive genes than BIO treatment. Genes identified as responding differently to BIO treatment
393 in *rad21* mutants compared to wild type (395 in *rad21*$^{-/-}$ and 467 in *rad21*$^{+/-}$) cluster together in
394 the dendrograms regardless of BIO treatment (Fig. 8A,B). Primarily, expression of these genes
395 differs from wild type by being strongly responsive to BIO in wild type, and much less responsive
396 to BIO with homozygous or heterozygous *rad21* mutation. In *stag2b*$^{-/-}$ much more complex
397 interactions were observed between the genotype and BIO treatment. In the dendrograms of genes
398 differentially responsive to BIO (539 genes), untreated *stag2b*$^{-/-}$ gene sets cluster with BIO-treated
399 wild type, and BIO-treated *stag2b*$^{-/-}$ gene sets cluster with untreated wild type (Fig. 8C). This
400 suggests that there is an altered baseline of Wnt signaling in *stag2b*$^{-/-}$ and also that BIO
401 stimulation normalizes the expression of select dysregulated genes in *stag2b*$^{-/-}$ tailbuds.

402 A notable example of a gene with expression that is rescued by BIO in *stag2b*$^{-/-}$ but not in
403 *rad21* mutants is *vimr2*, a marker of EMT and mesoderm formation in the tailbud (Goto et al.,
404 2017). Expression of *vimr2* was strongly downregulated in all cohesin mutant tailbuds (Fig. 8D,
405 Fig. S12). While BIO treatment had a minimal effect on *vimr2* transcript counts in wild type and
406 *rad21* mutants, it restored *vimr2* levels in *stag2b*$^{-/-}$ tailbuds to wild type (Fig. 8D). Interestingly,
407 our single cell RNA-seq data show that *vimr2* is expressed in NMPs and is significantly
408 downregulated in *stag2b*$^{-/-}$ mutants (Fig. 9AB). This finding raises the possibility that EMT
409 anomalies marked by downregulated *vimr2* could be responsible for changes in mesoderm
410 induction in *stag2b*$^{-/-}$.
411

412 **Wnt stimulation rescues notochord width in *stag2b*$^{-/-}$ tailbuds**

413 If BIO stimulation can restore transcription in *stag2b*$^{-/-}$ tailbuds, we reasoned that it may
414 also rescue the narrower notochord phenotype in *stag2b*$^{-/-}$ embryos. Using HCR with probes for
415 *sox2*, *tbxta*, and *tbx16*, we quantified the thickness of the notochord (*tbxta*) in wild-type and
416 *stag2b*$^{-/-}$ embryos both with and without BIO treatment (Fig. 10, Fig. S13). While Wnt stimulation
417 modestly increased notochord width in wild type ($p \leq 0.01$), it significantly increased the width in
418 *stag2b*$^{-/-}$ embryos ($p \leq 0.0001$) (Fig. 10E). When we compared the notochord width in wild type
419 embryos to that in *stag2b*$^{-/-}$ embryos treated with BIO, the difference was statistically
420 insignificant. Therefore, Wnt stimulation rescues the notochord phenotype in *stag2b*$^{-/-}$ embryos.

421 In summary, our findings suggest that Wnt stimulation with BIO normalizes gene
422 expression and phenotype in *stag2b*$^{-/-}$ embryos, while gene expression in *rad21* mutants is
423 unresponsive to Wnt stimulation.
424

425 **Discussion**

426 Germline mutations in subunits of the cohesin complex or its regulators are implicated in
427 developmental disorders known as cohesinopathies, and somatic mutations are now known to
428 cause a variety of cancers. Such cohesin mutations are invariably partial loss-of-function rather
429 than null alleles, because of the essential cell cycle role of cohesin. To date, few studies have
430 compared the developmental consequences of reducing the overall amount of cohesin versus
431 altering its composition. Upon deficiency of Stag2, the Stag1 subunit will compensate in the
432 cohesin complex thereby altering cohesin composition. Upon deficiency of Rad21, the overall
433 quantity of cohesin complexes on DNA decreases. In this study, we took advantage of *stag* and

434 *rad21* gene mutations in zebrafish to show that cohesin composition versus quantity lead to
435 strikingly different consequences for gene transcription and cell differentiation.

436 Rad21 is an essential subunit in the cohesin complex. Using a zebrafish point mutation
437 *rad21*^{nz171} that progressively reduces *rad21* transcript from heterozygotes to homozygotes, we
438 show that Rad21 deficiency dose-dependently correlates with downregulation of core cohesin
439 subunits, a transcriptional dysregulation signature enriched in cell cycle genes, and a block in
440 mesoderm induction in the tailbud. Although *rad21* heterozygotes are viable and fertile,
441 developmental anomalies in the tailbud are more similar between heterozygotes and *rad21*
442 homozygotes (which die by 48 hpf) than homozygous viable *stag2b*$^{-/-}$. Additional mutation of
443 *stag1b* on top of *stag2b* resulted in loss of viability and a phenotype resembling *rad21* mutants.

444 It is not clear whether the consequences of Rad21 deficiency are related to cell cycle
445 effects. A recent study showed that blocking the cell cycle in zebrafish does not affect the
446 development of cell types (Kukreja et al., 2023), although it does affect the numbers of presomitic
447 mesoderm cells and erythrocytes (Kukreja et al., 2023). Consistent with cell cycle effects, we
448 observed that *rad21* mutation impacts mesoderm differentiation and our previous work has shown
449 that erythropoiesis is downregulated in cohesin mutants (Horsfield et al., 2007; Ketharnathan et
450 al., 2020). However, cell cycle impairment is unlikely to account for all the defects associated
451 with the reduction in cohesin dose, and it does not explain the transcriptional and phenotypic
452 changes observed in a viable, fertile *rad21* heterozygotes. Our previous work has shown that
453 Rad21 deficiency has remarkably specific transcriptional and developmental consequences, for
454 example, cell-type specific loss of *runx1* expression (Horsfield et al., 2007; Marsman et al., 2014;
455 Schuster et al., 2015). Consistent with a non-cell cycle related transcriptional role for Rad21,
456 complete removal of Rad21 interferes with transcription in post-mitotic neurons, which is rescued
457 upon restoring Rad21 (Weiss et al., 2021). While homozygous mutation in *stag2b* has no
458 statistically significant effect on the cell cycle determined by cytometry, single cell RNA-seq
459 showed an increase in G0/G1 and reduction in S phases in the paraxial mesoderm cell cluster 1 in
460 *stag2b*$^{-/-}$ mutants, indicating cell cycle downregulation. It is difficult to determine whether this
461 reduced cell cycle progression in *stag2b*$^{-/-}$ mutants is intrinsic, or caused by altered signalling
462 pathways such as Wnt, or a combination of these.

463 Although *stag2b*$^{-/-}$ mutants are viable and fertile, changes in mesoderm differentiation are
464 apparent at tailbud stages. An increase in paraxial mesoderm of *stag2b*$^{-/-}$ mutants was detected by
465 both bulk and single cell RNA-seq. Cells over-represented in paraxial mesoderm of *stag2b*$^{-/-}$
466 mutants have exited the cell cycle and are more differentiated, since they primarily reside in the

467 anterior paraxial mesoderm region where somites start to form. We speculate that these
468 characteristics reflect an inability of *stag2b*$^{-/-}$ mutants to maintain mesoderm progenitors, thereby
469 lowering the proportion of immature cells relative to mature cells. Consistent with this
470 interpretation, in *stag2b*$^{-/-}$ mutants have fewer of the less mature posterior paraxial mesoderm
471 cells and *stag2b*$^{-/-}$ NMPs have downregulated stem cell pathways.

472 *stag2b*$^{-/-}$ mutants have a unique, narrow notochord phenotype. An increase in anterior
473 paraxial mesoderm in *stag2b*$^{-/-}$ mutants could in turn affect cell fate decisions in the notochord
474 through increased Notch signalling. The single cell RNA-seq data shows an increase in
475 hypochord cell numbers at the expense of notochord in *stag2b*$^{-/-}$ mutants, possibly linked to
476 increased Notch signalling from higher numbers of paraxial mesoderm cells (Latimer and Appel,
477 2006). Reduced Wnt signalling might also directly lead to a decrease in *noto* expression affecting
478 notochord identity. Cells that lose notochord identity were previously shown to end up in the
479 paraxial mesoderm (Halpern et al., 1995; Yamanaka et al., 2007). Finally, EMT driver *vimr2* is
480 downregulated in NMPs and cell adhesion pathways are increased in the notochord, indicating
481 that the ability of cells to undergo EMT and contribute to various developing tissues might be
482 altered in *stag2b*$^{-/-}$ mutants. Strikingly, our rescue experiments suggest that driving Wnt signaling
483 can compensate for transcriptional and phenotypic changes in the *stag2b*$^{-/-}$ mutants. The
484 noticeably narrower notochord in *stag2b*$^{-/-}$ mutants is rescued by stimulation of Wnt signaling via
485 inhibition of GSK3. Moreover, transcription in *stag2b*$^{-/-}$ mutant tailbuds is rescued to wild type
486 upon GSK3 inhibition.

487 We and others have previously reported dysregulated Wnt signaling upon cohesin
488 mutation (Chin et al., 2020; Grazioli et al., 2021; Mazzola et al., 2019; Medina et al., 2016;
489 Pileggi et al., 2021; Schuster et al., 2015) but the directionality of Wnt signaling disturbance
490 remains unclear. We have shown stabilization of β-catenin and both up- and downregulation of
491 components of the Wnt signaling pathway, indicating that the effects of cohesin deficiency on
492 Wnt are likely to be complex (Chin et al., 2020). Interestingly, GSK3α inhibition was shown to
493 stabilize cohesin on chromatin, promoting continued loop extrusion (Park et al., 2023). Stabilized
494 loop extrusion is dependent on cohesin as it was eliminated with knockdown of Rad21 (Park et
495 al., 2023). The compound (BIO) we used to stimulate Wnt inhibits GSK3 and does not distinguish
496 between α and β forms. It is possible that GSK3 inhibition was able to rescue transcription and
497 phenotypes in *stag2b*$^{-/-}$ mutant tailbuds but not in *rad21* mutants because a reduction in Rad21
498 reduces the number of complexes that can be stabilized, whereas Stag1b compensates for loss of
499 Stag2b in those complexes.

500 Stag1-containing cohesin resides primarily at CTCF sites that demarcate contact domains
501 that are invariant between tissues. Stag2-containing cohesin resides at CTCF and non-CTCF sites,
502 where it is thought to regulate tissue-specific transcription (Casa et al., 2020; Cuadrado and
503 Losada, 2020; Kojic et al., 2018; Viny et al., 2019). Viny et al (2019) showed that Stag1-cohesin
504 cannot fully substitute for Stag2 cohesin in haematopoietic stem cells. It is possible that Stag1-
505 containing cohesin has different properties in loop extrusion than Stag2 cohesin (Alonso-Gil et
506 al., 2023; Cuadrado and Losada, 2020) and likely that developmental gene transcription in *stag2b*$^{-}$
507 $^{/-}$ mutant tailbuds is altered because of compensation by Stag1b. We do not believe the other Stag
508 orthologs are major contributors to development in zebrafish owing to the lethality of *stag1b*$^{-/-}$;
509 *stag2b*$^{-/-}$ double mutants.

510 Stag proteins may have functions that are independent of the cohesin complex. For
511 example a recent study found that upon RAD21 depletion, STAG proteins remain bound to
512 chromatin, interact with CTCF, and cluster in 3D (Porter et al., 2023). STAG proteins interact
513 with RNA and R-loops even in the absence of cohesin. Drosophila's SA cohesin subunit
514 (equivalent to Stag2) is differentially enriched at enhancers and promoters near origins of
515 replication where it is proposed to recruit cohesin (Pherson et al., 2019). In contrast, RAD21
516 appears to be key for stable binding of cohesin at CTCF sites. A recent study showed that when
517 RAD21 is cleaved, cohesin is released from DNA including at CTCF sites and loops at these
518 elements are lost (Liu and Dekker, 2022). Interestingly CTCF-independent cohesin-anchored
519 loops within chromatin domains persisted despite RAD21 cleavage. The different molecular and
520 structural behaviour of RAD21 and STAG proteins is consistent with the diverse developmental
521 consequences we observed upon germline mutation in these genes.

522 It is possible that some of the molecular basis for developmental abnormalities is shared
523 between NIPBL deficiency and STAG2 mutation. A study describing single cell RNA sequencing
524 of early-stage mouse embryos with one deleted copy of *Nipbl* showed that these embryos also
525 experience changes in mesoderm fate and have altered mesoderm cell populations (Chea et al.,
526 2024). Nipbl loss altered the regulation of genes involved in EMT, which parallels our findings in
527 *stag2b* mutant zebrafish embryos. This raises the possibility that abnormal regulation of
528 mesoderm fate could be a conserved feature of the cohesinopathies.

529 Alternate transcriptional and developmental consequences with Stag2 and Rad21
530 deficiency have implications for the amelioration of cohesinopathies where Wnt agonists have
531 been explored as potential therapeutic agents for individuals with CdLS (Grazioli et al., 2021),
532 and additionally, for the treatment of cohesin-mutant cancers (Chin et al., 2020). Our results

533 suggest that reducing cohesin dose has very different consequences to altering cohesin
534 composition. This indicates that mutations in core cohesin subunits need to be considered
535 differently to mutations in alternate cohesin subunits or cohesin regulators when developing
536 therapeutics.
537
538 **Materials and Methods**
539
540 **Zebrafish Husbandry**

541 Wild type (WIK) (Rauch et al., 1997), *stag1a*^{nz204} (Ketharnathan et al., 2020), *stag1b*^{nz205}
542 (Ketharnathan et al., 2020), *stag2b*^{nz207} (Ketharnathan et al., 2020) and *rad21*^{nz171} *(Horsfield et al.,*
543 *2007)* zebrafish lines were maintained at 28 °C according to established husbandry methods
544 (Westerfield, 1995). Zebrafish were housed in the Otago Zebrafish Facility (Department of
545 Pathology, University of Otago, Dunedin, New Zealand). All animal work was performed in
546 accordance with the Otago Zebrafish Facility Standard Operating Procedures (AUP 21-110) and
547 under Environmental Risk Management Authority approval numbers GMC005627, GMD100922
548 and GMC001366. For all experiments, embryos were developed at 22 or 28 °C.
549
550 **Whole-mount *in situ* hybridisations (WISH) and hybridization chain reaction (HCR) RNA-**
551 **FISH**

552 WISH for *runx1* was performed using 0.5 ng/µL of riboprobe as previously described
553 (Kalev-Zylinska et al., 2002). Probes for *sox2*, *tbxta*, and *tbx16* and HCR reagents were purchased
554 from Molecular Instruments, Inc (USA). HCR was performed according to the manufacturer's
555 protocol for zebrafish embryos.
556
557 **Flow cytometry**

558 Embryos at the 16-somite stage were fixed in methanol (García-Castro et al., 2021) and
559 tailbuds were dissected (n=30). *rad21* heterozygotes and homozygotes were identified by
560 genotyping the heads of individual embryos using genomic DNA extraction (Meeker et al., 2007)
561 followed by a custom TaqMan assay. Cells were filtered through a 40 µm cell strainer and nuclei
562 were stained with DRAQ5 (#ab108410, Abcam) at 5 µM final concentration on ice for 45 min in
563 the dark. Cell cycle profiles of three independent replicates for each genotype were obtained using
564 a BD FACS Aria III (BD Biosciences). Data analysis and plots were generated using Cytoflow
565 (Teague, 2022).

566

567 **BrdU incorporation**

568 For cell cycle analyses dechorionated embryos were incubated in 10 mM BrdU in
569 Ringer's solution for 30 minutes on ice, rinsed three times with Ringer's solution and incubated
570 for 30 minutes or 2 hours at 28 °C. Embryos were fixed with 4% PFA overnight at 4 °C,
571 dehydrated in methanol and stored at -20 °C in 100% MeOH.

572 For staining, embryos were rehydrated in a series of 5-minute washes with PBST/MeOH.
573 Embryos older than 24 hpf were treated with 10 µg/ml proteinase K for 10 minutes, followed by
574 three 5-minute washes in PBST and post-fixation in 4% PFA for 20 minutes at room temperature.
575 Samples were then rinsed three times with sterile distilled water. For BrdU staining, embryos
576 were rinsed twice in 2 N HCl and incubated in 2 N HCl for 1 hour at room temperature to
577 denature DNA and expose the BrdU epitope. Alternatively, for antigen exposure, embryos were
578 treated with acetone for 20 minutes on ice.

579

580 **Immunohistochemistry**

581 Samples were rinsed twice with sterile distilled water and washed twice with PBST for 5
582 minutes. Embryos were incubated in blocking solution (0.2% Roche block, 10% FBS, 1% DMSO
583 in PBST) for 30 minutes, followed by a 2-day incubation with primary antibodies at 4 °C.
584 Primary antibodies used are as follows: anti-phH3 (#3377, Cell Signaling Technology; 1:1000),
585 anti-α-tubulin (#T6199, Sigma-Aldrich; 1:500), and anti-BrdU (#B35141, Thermo Fisher
586 Scientific). Antibodies were washed off with three 10-minute washes in PBST and two 10-minute
587 washes in 1% FBS in PBST. Embryos were then incubated with secondary antibodies (1:1000) in
588 1% FBS in PBST at 4 °C for 2 days in the dark. Secondary antibodies used for
589 immunofluorescence were goat anti-mouse Alexa Fluor 488 (1:1000, #A11001, Thermo Fisher
590 Scientific), chicken anti-rabbit Alexa Fluor 647(1:1000, #A21443, Thermo Fisher Scientific). On
591 the second day, Hoechst 33342 (1 µg/ml) (Thermo Fisher Scientific; 1:1000) was added. Embryos
592 were washed five times for 10 minutes with PBST and imaged.

593

594 **Microscopy**

595 Fixed embryos were immersed in 70% glycerol to obtain bright field images. Live
596 embryos were anaesthetized with MS-222 (200 mg/L) and embedded in 3% methylcellulose.
597 Bright-field images were captured using the Leica M205FA epifluorescence microscope equipped
598 with a DFC490 camera and Leica Applications Suite software (Leica Microsystems, Germany).

599 For confocal microscopy embryos were mounted in 1% low melting agarose (w/v).
600 Confocal images were acquired using a Nikon C2 confocal microscope as Z-stacks of the optical
601 sections. The images were processed using NIS-Elements Denoise.ai Software. Maximum
602 intensity projections were used for the figures.
603
604 **BIO treatment**

605 30 µM 6-bromoindirubin-3'-oxime (BIO) solution was diluted to 2.5 µM BIO in E3
606 medium. Embryos were sorted into 50 embryos per plate and treated with 2.5 µM BIO from 4 hpf
607 until tailbud dissection at the 16-somite stage.
608
609 **Tailbud bulk RNA sequencing (RNA-seq) and analyses**

610 Tailbuds were dissected from stage-matched embryos at 16 somites (16-18 hpf) as
611 illustrated in Figure 3A. For RNA-seq, tailbuds were individually lysed in 3 µL of RLT + BME
612 (Qiagen RNeasy) and stored in separate PCR tubes at -80 °C to await genotyping of heads (for
613 *rad21*$^{-/-}$ and *rad21*$^{+/-}$). Total RNA was extracted from the pools of 80 tailbuds per sample using
614 the RNeasy Micro kit (74104; Qiagen, Germany). Quality and quantity of RNA were assessed
615 using Qubit 4.0 Fluorometer (Thermo Fisher Scientific, USA), Agilent RNA 6000 Nano Kit on
616 2100 Bioanalyzer (Agilent Technologies, Netherlands) and NanoPhotometer NP80 Touch
617 (Implen GmbH, Germany).

618 Libraries were prepared from 250 ng of total RNA using the TruSeq Stranded mRNA
619 Library Prep kit (Illumina, USA) and TruSeq RNA CD Index Plate (Illumina, USA) for sample
620 multiplexing. The concentration of the libraries was quantified using a Qubit 4.0 Fluorometer
621 (Thermo Fisher Scientific, USA), and the mean fragment size was assessed using the DNA High
622 Sensitivity KIT on a 2100 Bioanalyzer (Agilent Technologies, Netherlands). 4 nM pooled
623 libraries was sequenced on NovaSeq S1 flow cell by Livestock Improvement Corporation Ltd.
624 (New Zealand).

625 RNA-seq reads were trimmed using Cutadapt (Martin, 2011), and aligned to the reference
626 genome (GRCz11.98, genome-build-accession NCBI:GCA_000002035.4) with HISAT2 (Kim et
627 al., 2019) and SAMtools (Danecek et al., 2021). FeatureCounts (Liao et al., 2014) was used to
628 generate fragment count matrices. DESeq2 (Love et al., 2014) was used to perform differential
629 gene expression analysis, and multi-testing correction was done using the Benjamini-Hochberg
630 procedure. The false discovery rate (FDR) threshold was set at 5%. Pathway analysis was
631 performed using Reactome (Yu and He, 2016) and Metascape (Zhou et al., 2019). Genotype-

632 specific BIO treatment effects were tested by adding an interaction term (modelling the
633 interaction between treatment and genotype) at the experimental design stage prior to calling
634 differential genes. R package ggplot2 was used for data visualisation (Wickham et al., 2016).
635

636 **Single cell RNA sequencing**

637 Stage matched 16-somite embryos (wild type or *stag2b*$^{-/-}$) were dechorionated using
638 pronase (20 mg/ml in E3) and deyolked in calcium-free Ringer's solution (116 mM NaCl, 2.6
639 mM KCl, 5 mM HEPES, pH 7.0). Tailbud tissue was dissected using a small needle and pooled
640 (n=30). Tailbuds were then incubated with collagenase (2 mg/ml in 0.05% trypsin, 1 mM EDTA,
641 pH 8.0, PBS) in 1.3 ml volume at 28 °C for 15 mins with intermittent pipetting to achieve a single
642 cell suspension. The reaction was stopped by adding 200 µL of a stop solution (30% calf serum, 6
643 mM $CaCl_2$, PBS). Cells were centrifuged at 500 x *g* for 5 min and re-suspended in 1 ml of
644 resuspension solution (1% Calf Serum, 0.8 mM $CaCl_2$, 50 U/ml Penicillin, 0.05 mg/ml
645 Streptomycin). After centrifuging cells were resuspended in 700 µL of resuspension buffer and
646 filtered through a 40 µm cell strainer and kept on ice. Single cell suspensions were processed at
647 the Genomics High Throughput Facility (UC Irvine, USA) according to the manufacturers
648 protocol for the 10X Chromium single cell platform (10X Genomics, Pleasanton, CA. USA)
649 specifically the Chromium Single Cell 3' Library and Gel Bead Kit v3 (Cat.No. PN-1000128).
650 Libraries were sequenced on a HiSeq2500 platform (Illumina) yielding 1,024,641,718 reads for
651 wild type and 1,166,072,985 reads for the *stag2b*$^{-/-}$ sample.
652

653 **Single cell RNA sequencing data analysis**

654 Single-cell RNA-seq FASTQ files were demultiplexed using Cellranger (v7.1.0) (Zheng et
655 al., 2017), mapped to the Danio rerio.GRCz11 (danRer11) transcriptome (v4.3.2) (Lawson et al.,
656 2020) including intronic reads. We obtained an estimated number of 18,704 cells (wild type) and
657 26,360 cells (*stag2b*$^{-/-}$). In the wild type sample mean reads per cell were 54,782, median UMI
658 (unique molecular identifier) were 15,696 and 3,651 genes detected per cell.

659 In *stag2b*$^{-/-}$ we detected 44,236 mean reads per cell, 13,922 median UMIs and 3,561
660 median genes per cell. The total number of genes detected was 25,879 for wild type and 26,168
661 for *stag2b*$^{-/-}$. 98.2% and 98.1% of reads (wild type/ *stag2b*$^{-/-}$) had valid barcodes with Q30 of
662 94.4% for both samples. 92% and 91.5% of the reads mapped confidently to the zebrafish
663 genome. Downstream anlaysis was performed using scvi-tools (v0.20.3) (Gayoso et al., 2022) and
664 scanpy (v1.9.3) (Wolf et al., 2018). After filtering of empty cells doublet removal was performed

using Solo (Bernstein et al., 2020). scanpy was used to filter out cells with less than 200 genes, genes detected in less than 3 cells, cells exceeding gene counts of 98% of the median. We also filtered out cells having more than 15% mitochondrial reads (indicating cellular stress).

Filtering and QC steps resulted in 15,298 wild type cells and 21,278 *stag2b*$^{-/-}$cells. The data were then normalised to 10,000 UMIs per cell using the function (scanpy.pp.normalize). Integration of the two datasets was performed using scvi-tools function scVI model (Lopez et al., 2018). Specifically we set up the model using the following command: scvi.model.SCVI.setup_anndata (anndata, layer = "counts", categorical_covariate_keys=["Genotype"], continuous_covariate_keys=['pct_counts_mt', 'total_counts']). We then trained the model and obtained the latent space with model.get_latent_representation(). The neighbourhood graph as well as the UMAP (Uniform Manifold Approximation and Projection) plot was determined using scanpy functions using the scVI latent space as input. Leiden clustering (resolution 0.4) resulted in 13 clusters. Significant differentially expressed genes (DEGs) between clusters and genotypes were determined using the the model calculated above using scVI as well as the Wilcoxon rank-sum test (Benjamini-Hochberg correction).

A list of differentially expressed genes for clusters of interest between wild type and *stag2b*$^{-/-}$ tailbuds can be found in Data S4. GSEApy (Fang et al., 2023) was used for gene set enrichment analysis on DEGs identified by pseudobulk analysis and Data S4. We used the function (scanpy.tl.score_genes) for a list of genes obtained from wiki pathways: Canonical_Wnt_pathway (WP1349). The function (scanpy.pp.scale) was used to scale the data to unit variance and zero mean and the Wilcoxon rank-sum test (Benjamini-Hochberg correction) was used to determine significance for the data shown in Fig. 7. Cell cycle phase was determined by using Scanpy's (Wolf et al., 2018) (sc.tl.score_genes_cell_cycle) function and https://github.com/scverse/scanpy_usage/blob/master/180209_cell_cycle/data/regev_lab_cell_cycle_genes.txt as gene list. Each cell is assigned a score based on the expression of marker genes for S and G2/M phase, if neither gene set score is above a threshold the cell is classified as G0/G1.

Data availability

Bulk RNA-seq data are available at GEO under accession number GSE247246.
Single cell RNA-seq data generated in this study are available at GEO under accession number GSE248952.

Acknowledgments

The authors would like to thank Noel Jhinku and Dr Doug Mackie for the expert management of the zebrafish facility and Dr Robert Woolley for the confocal microscopy advice. The authors are grateful to Dr Ben Martin and Dr Christian Mosimann for helpful advice and discussions.

Funding

Royal Society of NZ Marsden Fund grant MFP-UOO2013 (JAH, GG).

Author contributions

Conceptualization: AAL, JAH

Methodology: AAL, MM, GG, DT, BM, JA, JAH

Investigation: AAL, MM, GG, DT, SK, JAH

Visualization: AAL, MM, GG, DT, SK, TFS, JAH

Supervision: GG, JA, TFS, JAH

Writing—original draft: AAL, MM, JAH

Writing—review & editing: AAL, MM, GG, DT, SK, BM, TFS, JA, JAH

Competing interests:

No competing interest declared

References

Adane, B., Alexe, G., Seong, B. K. A., Lu, D., Hwang, E. E., Hnisz, D., Lareau, C. A., Ross, L., Lin, S. and Cruz, F. S. D. (2021). STAG2 loss rewires oncogenic and developmental programs to promote metastasis in Ewing sarcoma. *Cancer cell* **39**, 827-844. e810.

Alonso-Gil, D., Cuadrado, A., Giménez-Llorente, D., Rodríguez-Corsino, M. and Losada, A. (2023). Different NIPBL requirements of cohesin-STAG1 and cohesin-STAG2. *Nature Communications* **14**, 1326.

Bailey, M. L., Tieu, D., Habsid, A., Tong, A. H. Y., Chan, K., Moffat, J. and Hieter, P. (2021). Paralogous synthetic lethality underlies genetic dependencies of the cancer-mutated gene STAG2. *Life Science Alliance* **4**.

Ball, A. R., Jr., Chen, Y. Y. and Yokomori, K. (2014). Mechanisms of cohesin-mediated gene regulation and lessons learned from cohesinopathies. *Biochim Biophys Acta* **1839**, 191-202.

733 **Bernstein, N. J., Fong, N. L., Lam, I., Roy, M. A., Hendrickson, D. G. and Kelley, D. R.** (2020).
734 Solo: doublet identification in single-cell RNA-seq via semi-supervised deep learning. *Cell*
735 *systems* **11**, 95-101. e105.

736 **Casa, V., Gines, M. M., Gusmao, E. G., Slotman, J. A., Zirkel, A., Josipovic, N., Oole, E., van**
737 **IJcken, W. F., Houtsmuller, A. B. and Papantonis, A.** (2020). Redundant and specific
738 roles of cohesin STAG subunits in chromatin looping and transcriptional control. *Genome*
739 *Research* **30**, 515-527.

740 **Chea, S., Kreger, J., Lopez-Burks, M. E., MacLean, A. L., Lander, A. D. and Calof, A. L.**
741 (2024). Gastrulation-stage gene expression in Nipbl+/− mouse embryos foreshadows the
742 development of syndromic birth defects. *Science Advances* **10**, eadl4239.

743 **Cheng, H., Zhang, N. and Pati, D.** (2020). Cohesin subunit RAD21: From biology to disease.
744 *Gene* **758**, 144966.

745 **Cheng, N., Li, G., Kanchwala, M., Evers, B. M., Xing, C. and Yu, H.** (2022). STAG2 promotes
746 the myelination transcriptional program in oligodendrocytes. *Elife* **11**, e77848.

747 **Cheung, K. and Upton, J.** (2015). Cornelia de Lange Syndrome. *J Hand Surg Am* **40**, 2501-2503.

748 **Chin, C. V., Antony, J., Ketharnathan, S., Labudina, A., Gimenez, G., Parsons, K. M., He, J.,**
749 **George, A. J., Pallotta, M. M., Musio, A., et al.** (2020). Cohesin mutations are synthetic
750 lethal with stimulation of WNT signaling. *Elife* **9**, e61405.

751 **Ciosk, R., Shirayama, M., Shevchenko, A., Tanaka, T., Toth, A., Shevchenko, A. and**
752 **Nasmyth, K.** (2000). Cohesin's binding to chromosomes depends on a separate complex
753 consisting of Scc2 and Scc4 proteins. *Molecular cell* **5**, 243-254.

754 **Cratsenberg, D. M., Winningham, P. J. and Starr, L. J.** (2021). Second reported individual with
755 a partial STAG2 deletion: Middle interhemispheric variant holoprosencephaly in STAG2-
756 related cohesinopathy. *Clinical Dysmorphology* **30**, 159-163.

757 **Cuadrado, A. and Losada, A.** (2020). Specialized functions of cohesins STAG1 and STAG2 in
758 3D genome architecture. *Curr Opin Genet Dev* **61**, 9-16.

759 **Cukrov, D., Newman, T. A., Leask, M., Leeke, B., Sarogni, P., Patimo, A., Kline, A. D.,**
760 **Krantz, I. D., Horsfield, J. A. and Musio, A.** (2018). Antioxidant treatment ameliorates
761 phenotypic features of SMC1A-mutated Cornelia de Lange syndrome in vitro and in vivo.
762 *Human molecular genetics* **27**, 3002-3011.

763 **Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A.,**
764 **Keane, T., McCarthy, S. A. and Davies, R. M.** (2021). Twelve years of SAMtools and
765 BCFtools. *Gigascience* **10**, giab008.

766 **Davidson, I. F., Bauer, B., Goetz, D., Tang, W., Wutz, G. and Peters, J.-M.** (2019). DNA loop
767 extrusion by human cohesin. *Science* **366**, 1338-1345.

768 **Deardorff, M. A., Bando, M., Nakato, R., Watrin, E., Itoh, T., Minamino, M., Saitoh, K.,**
769 **Komata, M., Katou, Y., Clark, D., et al.** (2012a). HDAC8 mutations in Cornelia de Lange
770 syndrome affect the cohesin acetylation cycle. *Nature* **489**, 313-317.

771 **Deardorff, M. A., Kaur, M., Yaeger, D., Rampuria, A., Korolev, S., Pie, J., Gil-Rodriguez, C.,**
772 **Arnedo, M., Loeys, B., Kline, A. D., et al.** (2007). Mutations in cohesin complex members
773 SMC3 and SMC1A cause a mild variant of cornelia de Lange syndrome with predominant
774 mental retardation. *Am J Hum Genet* **80**, 485-494.

775 **Deardorff, M. A., Noon, S. E. and Krantz, I. D.** (2020). Cornelia de Lange syndrome.

Deardorff, M. A., Wilde, J. J., Albrecht, M., Dickinson, E., Tennstedt, S., Braunholz, D., Mönnich, M., Yan, Y., Xu, W. and Gil-Rodríguez, M. C. (2012b). RAD21 mutations cause a human cohesinopathy. *The American Journal of Human Genetics* **90**, 1014-1027.

Dixon, J. R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J. S. and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376-380.

Dorsett, D. and Krantz, I. D. (2009). On the molecular etiology of Cornelia de Lange syndrome. *Annals of the New York Academy of Sciences* **1151**, 22-37.

Dorsett, D. and Strom, L. (2012). The ancient and evolving roles of cohesin in gene expression and DNA repair. *Curr Biol* **22**, R240-250.

Fang, Z., Liu, X. and Peltz, G. (2023). GSEApy: a comprehensive package for performing gene set enrichment analysis in Python. *Bioinformatics* **39**, btac757.

Freyberger, F., Kokotović, T., Krnjak, G., Frković, S. H. and Nagy, V. (2021). Expanding the known phenotype of Mullegama–Klein–Martinez syndrome in male patients. *Human Genome Variation* **8**, 37.

Friman, E. T., Flyamer, I. M., Marenduzzo, D., Boyle, S. and Bickmore, W. A. (2023). Ultra-long-range interactions between active regulatory elements. *Genome Research* **33**, 1269-1283.

Fudenberg, G., Imakaev, M., Lu, C., Goloborodko, A., Abdennur, N. and Mirny, L. A. (2016). Formation of chromosomal domains by loop extrusion. *Cell reports* **15**, 2038-2049.

García-Castro, H., Kenny, N. J., Iglesias, M., Álvarez-Campos, P., Mason, V., Elek, A., Schönauer, A., Sleight, V. A., Neiro, J. and Aboobaker, A. (2021). ACME dissociation: a versatile cell fixation-dissociation method for single-cell transcriptomics. *Genome biology* **22**, 1-34.

Gayoso, A., Lopez, R., Xing, G., Boyeau, P., Valiollah Pour Amiri, V., Hong, J., Wu, K., Jayasuriya, M., Mehlman, E. and Langevin, M. (2022). A Python library for probabilistic analysis of single-cell omics data. *Nature biotechnology* **40**, 163-166.

Gerlich, D., Koch, B., Dupeux, F., Peters, J.-M. and Ellenberg, J. (2006). Live-cell imaging reveals a stable cohesin-chromatin interaction after but not before DNA replication. *Current biology* **16**, 1571-1578.

Goel, V. Y., Huseyin, M. K. and Hansen, A. S. (2023). Region Capture Micro-C reveals coalescence of enhancers and promoters into nested microcompartments. *Nature Genetics*, 1-9.

Gokce-Samar, Z., de Bellescize, J., Arzimanoglou, A., Putoux, A., Chatron, N., Lesca, G. and Des Portes, V. (2022). STAG2 microduplication in a patient with eyelid myoclonia and absences and a review of EMA-related reported genes. *European Journal of Medical Genetics*, 104636.

Goto, H., Kimmey, S. C., Row, R. H., Matus, D. Q. and Martin, B. L. (2017). FGF and canonical Wnt signaling cooperate to induce paraxial mesoderm from tailbud neuromesodermal progenitors through regulation of a two-step epithelial to mesenchymal transition. *Development* **144**, 1412-1424.

Grazioli, P., Parodi, C., Mariani, M., Bottai, D., Di Fede, E., Zulueta, A., Avagliano, L., Cereda, A., Tenconi, R. and Wierzba, J. (2021). Lithium as a possible therapeutic strategy for Cornelia de Lange syndrome. *Cell Death Discovery* **7**, 34.

820 **Gruber, S., Haering, C. H. and Nasmyth, K.** (2003). Chromosomal cohesin forms a ring. *Cell*
821 **112**, 765-777.

822 **Halpern, M. E., Thisse, C., Ho, R. K., Thisse, B., Riggleman, B., Trevarrow, B., Weinberg, E.**
823 **S., Postlethwait, J. H. and Kimmel, C. B.** (1995). Cell-autonomous shift from axial to
824 paraxial mesodermal development in zebrafish floating head mutants. *Development* **121**,
825 4257-4264.

826 **Hnisz, D., Weintraub, A. S., Day, D. S., Valton, A.-L., Bak, R. O., Li, C. H., Goldmann, J.,**
827 **Lajoie, B. R., Fan, Z. P. and Sigova, A. A.** (2016). Activation of proto-oncogenes by
828 disruption of chromosome neighborhoods. *Science* **351**, 1454-1458.

829 **Horsfield, J. A.** (2023). Full circle: a brief history of cohesin and the regulation of gene expression.
830 *The FEBS journal* **290**, 1670-1687.

831 **Horsfield, J. A., Anagnostou, S. H., Hu, J. K., Cho, K. H., Geisler, R., Lieschke, G., Crosier,**
832 **K. E. and Crosier, P. S.** (2007). Cohesin-dependent regulation of Runx genes.
833 *Development* **134**, 2639-2649.

834 **Horsfield, J. A., Print, C. G. and Monnich, M.** (2012). Diverse developmental disorders from the
835 one ring: distinct molecular pathways underlie the cohesinopathies. *Front Genet* **3**, 171.

836 **Hou, W., Li, Y., Zhang, J., Xia, Y., Wang, X., Chen, H. and Lou, H.** (2022). Cohesin in DNA
837 damage response and double-strand break repair. *Critical Reviews in Biochemistry and*
838 *Molecular Biology* **57**, 333-350.

839 **Huis, P. J., Herzog, F., Ladurner, R., Davidson, I. F., Piric, S., Kreidl, E., Bhaskara, V.,**
840 **Aebersold, R. and Peters, J.-M.** (2014). Characterization of a DNA exit gate in the human
841 cohesin ring. *Science* **346**, 968-972.

842 **Ireland, M., Donnai, D. and Burn, J.** (1993). Brachmann-de Lange syndrome. Delineation of the
843 clinical phenotype. *American journal of medical genetics* **47**, 959-964.

844 **Jackson, L., Kline, A., Barr, M. and Koch, S. D.** (1993). de Lange syndrome: a clinical review
845 of 310 individuals. *American journal of medical genetics* **47**, 940-946.

846 **Kagey, M. H., Newman, J. J., Bilodeau, S., Zhan, Y., Orlando, D. A., van Berkum, N. L.,**
847 **Ebmeier, C. C., Goossens, J., Rahl, P. B. and Levine, S. S.** (2010). Mediator and cohesin
848 connect gene expression and chromatin architecture. *Nature* **467**, 430-435.

849 **Kalev-Zylinska, M. L., Horsfield, J. A., Flores, M. V. C., Postlethwait, J. H., Vitas, M. R.,**
850 **Baas, A. M., Crosier, P. S. and Crosier, K. E.** (2002). Runx1 is required for zebrafish
851 blood and vessel development and expression of a human RUNX1-CBF2T1 transgene
852 advances a model for studies of leukemogenesis.

853 **Kane, L., Williamson, I., Flyamer, I. M., Kumar, Y., Hill, R. E., Lettice, L. A. and Bickmore,**
854 **W. A.** (2022). Cohesin is required for long-range enhancer action at the Shh locus. *Nature*
855 *structural & molecular biology* **29**, 891-897.

856 **Ketharnathan, S., Labudina, A. and Horsfield, J. A.** (2020). Cohesin Components Stag1 and
857 Stag2 Differentially Influence Haematopoietic Mesoderm Development in Zebrafish
858 Embryos. *Front Cell Dev Biol* **8**, 617545.

859 **Kim, D., Paggi, J. M., Park, C., Bennett, C. and Salzberg, S. L.** (2019). Graph-based genome
860 alignment and genotyping with HISAT2 and HISAT-genotype. *Nature biotechnology* **37**,
861 907-915.

862 **Kojic, A., Cuadrado, A., De Koninck, M., Giménez-Llorente, D., Rodríguez-Corsino, M.,**
863 **Gómez-López, G., Le Dily, F., Marti-Renom, M. A. and Losada, A.** (2018). Distinct

roles of cohesin-SA1 and cohesin-SA2 in 3D chromosome organization. *Nature structural & molecular biology* **25**, 496-504.

Krab, L. C., Marcos-Alcalde, I., Assaf, M., Balasubramanian, M., Andersen, J. B., Bisgaard, A.-M., Fitzpatrick, D. R., Gudmundsson, S., Huisman, S. A. and Kalayci, T. (2020). Delineation of phenotypes and genotypes related to cohesin structural protein RAD21. *Human genetics* **139**, 575-592.

Krantz, I. D., McCallum, J., DeScipio, C., Kaur, M., Gillis, L. A., Yaeger, D., Jukofsky, L., Wasserman, N., Bottani, A. and Morris, C. A. (2004). Cornelia de Lange syndrome is caused by mutations in NIPBL, the human homolog of Drosophila melanogaster Nipped-B. *Nature genetics* **36**, 631-635.

Krijger, P. H. and de Laat, W. (2016). Regulation of disease-associated gene expression in the 3D genome. *Nat Rev Mol Cell Biol* **17**, 771-782.

Kruszka, P., Berger, S. I., Casa, V., Dekker, M. R., Gaesser, J., Weiss, K., Martinez, A. F., Murdock, D. R., Louie, R. J., Prijoles, E. J., et al. (2019). Cohesin complex-associated holoprosencephaly. *Brain* **142**, 2631-2643.

Kueng, S., Hegemann, B., Peters, B. H., Lipp, J. J., Schleiffer, A., Mechtler, K. and Peters, J.-M. (2006). Wapl controls the dynamic association of cohesin with chromatin. *Cell* **127**, 955-967.

Kukreja, K., Patel, N., Megason, S. G. and Klein, A. M. (2023). Global decoupling of cell differentiation from cell division in early embryo development. *bioRxiv*, 2023.2007. 2029.551123.

Latimer, A. J. and Appel, B. (2006). Notch signaling regulates midline cell specification and proliferation in zebrafish. *Developmental biology* **298**, 392-402.

Lawson, N. D., Li, R., Shin, M., Grosse, A., Yukselen, O., Stone, O. A., Kucukural, A. and Zhu, L. (2020). An improved zebrafish transcriptome annotation for sensitive and comprehensive detection of cell type-specific genes. *Elife* **9**, e55792.

Liao, Y., Smyth, G. K. and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930.

Liu, J. and Krantz, I. D. (2008). Cohesin and human disease. *Annu. Rev. Genomics Hum. Genet.* **9**, 303-320.

Liu, J., Zhang, Z., Bando, M., Itoh, T., Deardorff, M. A., Clark, D., Kaur, M., Tandy, S., Kondoh, T. and Rappaport, E. (2009). Transcriptional dysregulation in NIPBL and cohesin mutant human cells. *PLoS biology* **7**, e1000119.

Liu, Y. and Dekker, J. (2022). CTCF–CTCF loops and intra-TAD interactions show differential dependence on cohesin ring integrity. *Nature cell biology* **24**, 1516-1527.

Lopez, R., Regier, J., Cole, M. B., Jordan, M. I. and Yosef, N. (2018). Deep generative modeling for single-cell transcriptomics. *Nature methods* **15**, 1053-1058.

Love, M. I., Huber, W. and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology* **15**, 1-21.

Marsman, J. and Horsfield, J. A. (2012). Long distance relationships: enhancer-promoter communication and dynamic gene transcription. *Biochim Biophys Acta* **1819**, 1217-1227.

905 **Marsman, J., O'Neill, A. C., Kao, B. R., Rhodes, J. M., Meier, M., Antony, J., Monnich, M.**
906 **and Horsfield, J. A.** (2014). Cohesin and CTCF differentially regulate spatiotemporal
907 runx1 expression during zebrafish development. *Biochim Biophys Acta* **1839**, 50-61.

908 **Martin, M.** (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads.
909 *EMBnet. journal* **17**, 10-12.

910 **Mayerova, N., Cipak, L. and Gregan, J.** (2020). Cohesin Biology: From Passive Rings to
911 Molecular Motors. *Trends Genet* **36**, 387-389.

912 **Mazzola, M., Deflorian, G., Pezzotta, A., Ferrari, L., Fazio, G., Bresciani, E., Saitta, C.,**
913 **Ferrari, L., Fumagalli, M. and Parma, M.** (2019). NIPBL: a new player in myeloid cell
914 differentiation. *Haematologica* **104**, 1332.

915 **McNairn, A. J. and Gerton, J. L.** (2008). The chromosome glue gets a little stickier. *Trends in*
916 *genetics* **24**, 382-389.

917 **Medina, M. A., Ugarte, G. D., Vargas, M. F., Avila, M. E., Necunir, D., Elorza, A. A.,**
918 **Gutierrez, S. E. and De Ferrari, G. V.** (2016). Alternative RUNX1 Promoter Regulation
919 by Wnt/beta-Catenin Signaling in Leukemia Cells and Human Hematopoietic Progenitors.
920 *J Cell Physiol* **231**, 1460-1467.

921 **Meeker, N. D., Hutchinson, S. A., Ho, L. and Trede, N. S.** (2007). Method for isolation of PCR-
922 ready genomic DNA from zebrafish tissues. *Biotechniques* **43**, 610-614.

923 **Meier, M., Grant, J., Dowdle, A., Thomas, A., Gerton, J., Collas, P., O'Sullivan, J. M. and**
924 **Horsfield, J. A.** (2018). Cohesin facilitates zygotic genome activation in zebrafish.
925 *Development* **145**, dev156521.

926 **Mullegama, S. V., Klein, S. D., Mulatinho, M. V., Senaratne, T. N., Singh, K., Center, U. C.**
927 **G., Nguyen, D. C., Gallant, N. M., Strom, S. P. and Ghahremani, S.** (2017). De novo
928 loss-of-function variants in STAG2 are associated with developmental delay, microcephaly,
929 and congenital anomalies. *American journal of medical genetics Part A* **173**, 1319-1327.

930 **Mullegama, S. V., Klein, S. D., Signer, R. H., Center, U. C. G., Vilain, E. and Martinez-Agosto,**
931 **J. A.** (2019). Mutations in STAG2 cause an X-linked cohesinopathy associated with
932 undergrowth, developmental delay, and dysmorphia: Expanding the phenotype in males.
933 *Molecular genetics & genomic medicine* **7**, e00501.

934 **Mungan, Z., Akyüz, F., Bugra, Z., Yönal, O., Öztürk, S., Acar, A. and Çevikbas, U.** (2003).
935 Familial visceral myopathy with pseudo-obstruction, megaduodenum, Barrett's esophagus,
936 and cardiac abnormalities. *Official journal of the American College of Gastroenterology|*
937 *ACG* **98**, 2556-2560.

938 **Muto, A., Calof, A. L., Lander, A. D. and Schilling, T. F.** (2011). Multifactorial origins of heart
939 and gut defects in nipbl-deficient zebrafish, a model of Cornelia de Lange Syndrome. *PLoS*
940 *biology* **9**, e1001181.

941 **Muto, A. and Schilling, T. F.** (2017). Zebrafish as a model to study cohesin and cohesinopathies.
942 *Cohesin and Condensin: Methods and Protocols*, 177-196.

943 **Nagasaka, K., Davidson, I. F., Stocsits, R. R., Tang, W., Wutz, G., Batty, P., Panarotto, M.,**
944 **Litos, G., Schleiffer, A. and Gerlich, D. W.** (2023). Cohesin mediates DNA loop extrusion
945 and sister chromatid cohesion by distinct mechanisms. *Molecular Cell* **83**, 3049-3063.
946 e3046.

947 **Nasmyth, K. and Haering, C. H.** (2009). Cohesin: its roles and mechanisms. *Annu Rev Genet* **43**,
948 525-558.

949 **Nishiyama, T.** (2019). Cohesion and cohesin-dependent chromatin organization. *Current opinion*
950 *in cell biology* **58**, 8-14.

951 **Ochi, Y., Kon, A., Sakata, T., Nakagawa, M. M., Nakazawa, N., Kakuta, M., Kataoka, K.,**
952 **Koseki, H., Nakayama, M., Morishita, D., et al.** (2020). Combined Cohesin-RUNX1
953 Deficiency Synergistically Perturbs Chromatin Looping and Causes Myelodysplastic
954 Syndromes. *Cancer Discov* **10**, 836-853.

955 **Opitz, J. M. and Reynolds, J. F.** (1985). The Brachmann-de Lange syndrome. *American journal*
956 *of medical genetics* **22**, 89-102.

957 **Parelho, V., Hadjur, S., Spivakov, M., Leleu, M., Sauer, S., Gregson, H. C., Jarmuz, A.,**
958 **Canzonetta, C., Webster, Z. and Nesterova, T.** (2008). Cohesins functionally associate
959 with CTCF on mammalian chromosome arms. *Cell* **132**, 422-433.

960 **Park, D. S., Nguyen, S. C., Isenhart, R., Shah, P. P., Kim, W., Barnett, R. J., Chandra, A.,**
961 **Luppino, J. M., Harke, J. and Wai, M.** (2023). High-throughput Oligopaint screen
962 identifies druggable 3D genome regulators. *Nature*, 1-9.

963 **Percival, S. M., Thomas, H. R., Amsterdam, A., Carroll, A. J., Lees, J. A., Yost, H. J. and**
964 **Parant, J. M.** (2015). Variations in dysfunction of sister chromatid cohesion in esco2
965 mutant zebrafish reflect the phenotypic diversity of Roberts syndrome. *Dis Model Mech* **8**,
966 941-955.

967 **Pherson, M., Misulovin, Z., Gause, M. and Dorsett, D.** (2019). Cohesin occupancy and
968 composition at enhancers and promoters are linked to DNA replication origin proximity in
969 Drosophila. *Genome research* **29**, 602-612.

970 **Piché, J., Van Vliet, P. P., Pucéat, M. and Andelfinger, G.** (2019). The expanding phenotypes
971 of cohesinopathies: one ring to rule them all! *Cell Cycle* **18**, 2828-2848.

972 **Pileggi, S., La Vecchia, M., Colombo, E. A., Fontana, L., Colapietro, P., Rovina, D., Morotti,**
973 **A., Tabano, S., Porta, G., Alcalay, M., et al.** (2021). Cohesin Mutations Induce Chromatin
974 Conformation Perturbation of the H19/IGF2 Imprinted Region and Gene Expression
975 Dysregulation in Cornelia de Lange Syndrome Cell Lines. *Biomolecules* **11**, 1622.

976 **Porter, H., Li, Y., Neguembor, M. V., Beltran, M., Varsally, W., Martin, L., Cornejo, M. T.,**
977 **Pezic, D., Bhamra, A., Surinova, S., et al.** (2023). Cohesin-independent STAG proteins
978 interact with RNA and R-loops and promote complex loading. *Elife* **12**, e79386.

979 **Rao, S. S., Huntley, M. H., Durand, N. C., Stamenova, E. K., Bochkov, I. D., Robinson, J. T.,**
980 **Sanborn, A. L., Machol, I., Omer, A. D. and Lander, E. S.** (2014). A 3D map of the
981 human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**,
982 1665-1680.

983 **Rauch, G.-J., Granato, M. and Haffter, P.** (1997). A polymorphic zebrafish line for genetic
984 mapping using SSLPs on high-percentage agarose gels. *Technical Tips Online* **2**, 148-150.

985 **Remeseiro, S., Cuadrado, A., Gomez-Lopez, G., Pisano, D. G. and Losada, A.** (2012). A unique
986 role of cohesin-SA1 in gene regulation and development. *EMBO J* **31**, 2090-2102.

987 **Row, R. H., Pegg, A., Kinney, B. A., Farr 3rd, G. H., Maves, L., Lowell, S., Wilson, V. and**
988 **Martin, B. L.** (2018). BMP and FGF signaling interact to pattern mesoderm by controlling
989 basic helix-loop-helix transcription factor activity. *Elife* **7**, e31018.

990 **Row, R. H., Tsotras, S. R., Goto, H. and Martin, B. L.** (2016). The zebrafish tailbud contains
991 two independent populations of midline progenitor cells that maintain long-term germ layer
992 plasticity and differentiate in response to local signaling cues. *Development* **143**, 244-254.

993 **Sanborn, A. L., Rao, S. S., Huang, S.-C., Durand, N. C., Huntley, M. H., Jewett, A. I.,**
994 **Bochkov, I. D., Chinnappan, D., Cutkosky, A. and Li, J.** (2015). Chromatin extrusion
995 explains key features of loop and domain formation in wild-type and engineered genomes.
996 *Proceedings of the National Academy of Sciences* **112**, E6456-E6465.

997 **Schmidt, D., Schwalie, P. C., Ross-Innes, C. S., Hurtado, A., Brown, G. D., Carroll, J. S.,**
998 **Flicek, P. and Odom, D. T.** (2010). A CTCF-independent role for cohesin in tissue-specific
999 transcription. *Genome research* **20**, 578-588.

1000 **Schmidt, J., Dreha-Kulaczewski, S., Zafeiriou, M.-P., Schreiber, M.-K., Wilken, B., Funke,**
1001 **R., Neuhofer, C. M., Altmüller, J., Thiele, H. and Nürnberg, P.** (2022). Somatic
1002 mosaicism in STAG2-associated cohesinopathies: Expansion of the genotypic and
1003 phenotypic spectrum. *Frontiers in Cell and Developmental Biology* **10**, 1025332.

1004 **Schuster, K., Leeke, B., Meier, M., Wang, Y., Newman, T., Burgess, S. and Horsfield, J. A.**
1005 (2015). A neural crest origin for cohesinopathy heart defects. *Hum Mol Genet* **24**, 7005-
1006 7016.

1007 **Sjögren, C. and Nasmyth, K.** (2001). Sister chromatid cohesion is required for postreplicative
1008 double-strand break repair in Saccharomyces cerevisiae. *Current Biology* **11**, 991-995.

1009 **Soardi, F. C., Machado-Silva, A., Linhares, N. D., Zheng, G., Qu, Q., Pena, H. B., Martins, T.**
1010 **M., Vieira, H. G., Pereira, N. B. and Melo-Minardi, R. C.** (2017). Familial STAG2
1011 germline mutation defines a new human cohesinopathy. *NPJ genomic medicine* **2**, 7.

1012 **Steventon, B. and Martin, B.** (2022). A fishy tail: Insights into the cell and molecular biology of
1013 neuromesodermal cells from zebrafish embryos.

1014 **Sun, Y., Xu, X., Zhao, W., Zhang, Y., Chen, K., Li, Y., Wang, X., Zhang, M., Xue, B. and Yu,**
1015 **W.** (2023). RAD21 is the core subunit of the cohesin complex involved in directing genome
1016 organization. *Genome Biology* **24**, 1-27.

1017 **Teague, B.** (2022). Cytoflow: A python toolbox for flow cytometry. *bioRxiv*, 2022.2007.
1018 2022.501078.

1019 **Thisse, B., Pflumio, S., Fürthauer, M., Loppin, B., Heyer, V., Degrave, A., Woehl, R., Lux, A.,**
1020 **Steffan, T. and Charbonnier, X.** (2001). Expression of the zebrafish genome during
1021 embryogenesis. 2001. *ZFIN Direct Data Submission (http://zfin. org)*.

1022 **Tonkin, E. T., Wang, T.-J., Lisgo, S., Bamshad, M. J. and Strachan, T.** (2004). NIPBL,
1023 encoding a homolog of fungal Scc2-type sister chromatid cohesion proteins and fly Nipped-
1024 B, is mutated in Cornelia de Lange syndrome. *Nature genetics* **36**, 636-641.

1025 **Viny, A. D., Bowman, R. L., Liu, Y., Lavallee, V. P., Eisman, S. E., Xiao, W., Durham, B. H.,**
1026 **Navitski, A., Park, J., Braunstein, S., et al.** (2019). Cohesin Members Stag1 and Stag2
1027 Display Distinct Roles in Chromatin Accessibility and Topological Control of HSC Self-
1028 Renewal and Differentiation. *Cell Stem Cell* **25**, 682-696 e688.

1029 **Watrin, E. and Peters, J.-M.** (2006). Cohesin and DNA damage repair. *Experimental cell research*
1030 **312**, 2687-2693.

1031 **Weiss, F. D., Calderon, L., Wang, Y.-F., Georgieva, R., Guo, Y., Cvetesic, N., Kaur, M.,**
1032 **Dharmalingam, G., Krantz, I. D. and Lenhard, B.** (2021). Neuronal genes deregulated
1033 in Cornelia de Lange Syndrome respond to removal and re-expression of cohesin. *Nature*
1034 *Communications* **12**, 2919.

1035 **Wendt, K. S., Yoshida, K., Itoh, T., Bando, M., Koch, B., Schirghuber, E., Tsutsumi, S.,**
1036 **Nagae, G., Ishihara, K. and Mishiro, T.** (2008). Cohesin mediates transcriptional
1037 insulation by CCCTC-binding factor. *Nature* **451**, 796-801.

1038 **Westerfield, M.** (1995). *The zebrafish book: a guide for the laboratory use of zebrafish*
1039 *(Brachydanio rerio)*: University of Oregon press.

1040 **Wickham, H., Chang, W. and Wickham, M. H.** (2016). Package 'ggplot2'. *Create elegant data*
1041 *visualisations using the grammar of graphics. Version* **2**, 1-189.

1042 **Wolf, F. A., Angerer, P. and Theis, F. J.** (2018). SCANPY: large-scale single-cell gene expression
1043 data analysis. *Genome biology* **19**, 1-5.

1044 **Yamanaka, Y., Tamplin, O. J., Beckers, A., Gossler, A. and Rossant, J.** (2007). Live imaging
1045 and genetic analysis of mouse notochord formation reveals regional morphogenetic
1046 mechanisms. *Developmental cell* **13**, 884-896.

1047 **Yu, G. and He, Q.-Y.** (2016). ReactomePA: an R/Bioconductor package for reactome pathway
1048 analysis and visualization. *Molecular BioSystems* **12**, 477-479.

1049 **Yuan, B., Neira, J., Pehlivan, D., Santiago-Sim, T., Song, X., Rosenfeld, J., Posey, J. E., Patel,**
1050 **V., Jin, W. and Adam, M. P.** (2019). Clinical exome sequencing reveals locus
1051 heterogeneity and phenotypic variability of cohesinopathies. *Genetics in Medicine* **21**, 663-
1052 675.

1053 **Zhang, B. N., Chan, T. C. Y., Tam, P. O. S., Liu, Y., Pang, C. P., Jhanji, V., Chen, L. J. and**
1054 **Chu, W. K.** (2019). A cohesin subunit variant identified from a peripheral sclerocornea
1055 pedigree. *Disease Markers* **2019**.

1056 **Zheng, G. X., Terry, J. M., Belgrader, P., Ryvkin, P., Bent, Z. W., Wilson, R., Ziraldo, S. B.,**
1057 **Wheeler, T. D., McDermott, G. P. and Zhu, J.** (2017). Massively parallel digital
1058 transcriptional profiling of single cells. *Nature communications* **8**, 14049.

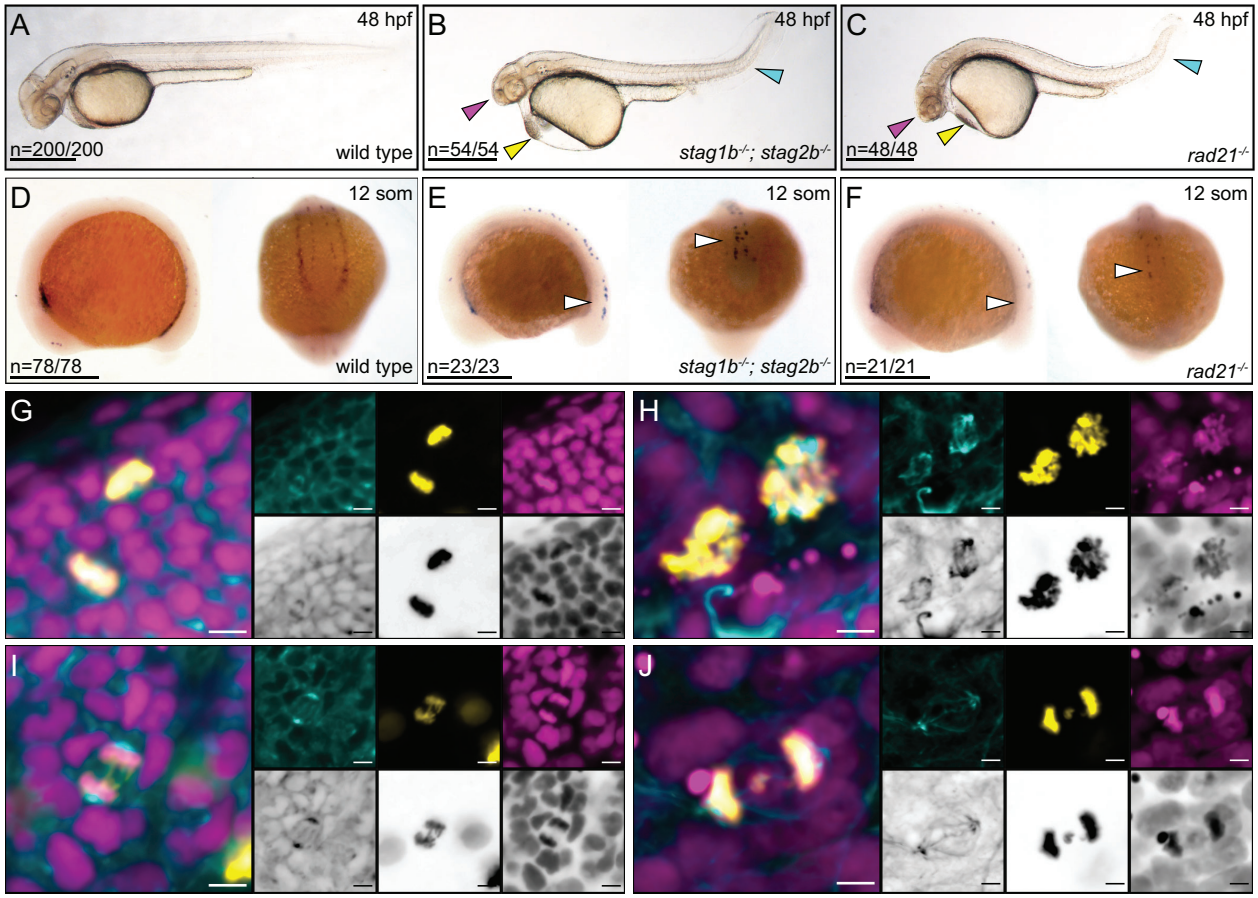
1059 **Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A. H., Tanaseichuk, O., Benner, C.**
1060 **and Chanda, S. K.** (2019). Metascape provides a biologist-oriented resource for the
1061 analysis of systems-level datasets. *Nature communications* **10**, 1523.

1062 **Zhu, Z. and Wang, X.** (2019). Roles of cohesin in chromosome architecture and gene expression.
1063 In *Seminars in Cell & developmental biology*, pp. 187-193: Elsevier.

1064

1065

1066 **Figures**



1067

1068 **Fig. 1. Combined loss of zebrafish Stag1b and Stag2b phenocopies the null cohesin mutation *rad21*.**
1069 **(A-C)** Lateral views of representative wild type (A), *stag1b*$^{-/-}$; *stag2b*$^{-/-}$ (B) and *rad21*$^{-/-}$ (C) embryos at 48
1070 hours post-fertilization (hpf). Arrows indicate developmental anomalies: magenta for a small head, yellow
1071 for pericardial oedema, and cyan for a kinked tail. Scale bars are 500 µm. **(D-F)** Expression of *runx1* at 12
1072 somites in wild type (D), *stag1b*$^{-/-}$; *stag2b*$^{-/-}$ (E) and *rad21*$^{-/-}$ (F) embryos. Lateral and posterior views are
1073 shown. White arrows indicate the loss of *runx1* expression in PLM. Scale bars are 500 µm. The numbers in
1074 the lower left hand corner indicate the number of embryos with similar expression patterns. **(G-J)** Confocal
1075 images of cell cycle progression in wild type (G,I) and *stag1b*$^{-/-}$; *stag2b*$^{-/-}$ (H,J) embryos at 48 hpf stained
1076 with anti-α-tubulin (cyan), anti-phH3 (yellow) antibodies and Hoechst (magenta). Images are maximum
1077 intensity projections of 3 (0.15 µm) optical sections taken from the tail region of 48 hpf embryos. Scale bars
1078 are 5 µm.
1079
1080
1081
1082

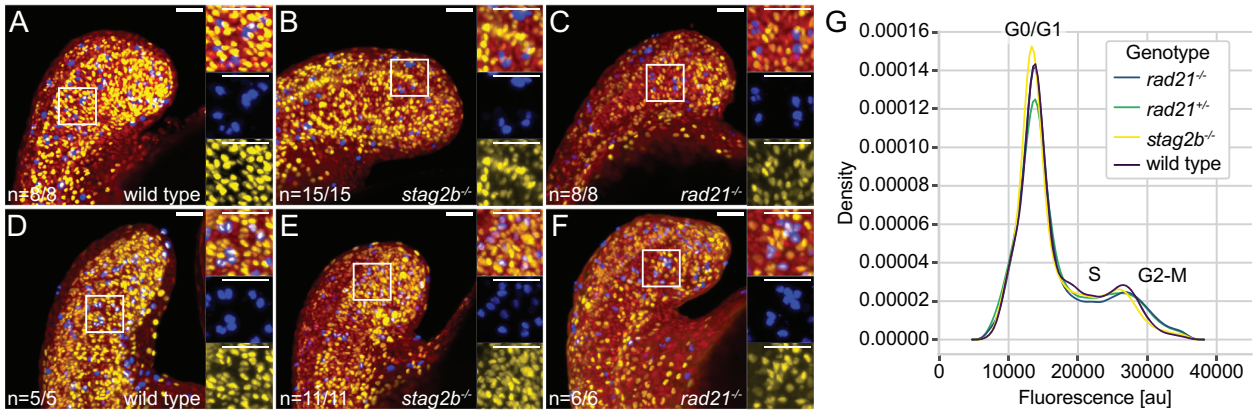


Fig. 2. The cell cycle is not blocked in cohesin mutants at the 16-somite stage. (A-F) Confocal images showing S phases and M phases in wild type (A, D), *stag2b*$^{-/-}$ (B, E), and *rad21*$^{-/-}$ (C, F) tailbuds at ~16 hpf. S phases are detected with anti-BrdU (yellow) and M phases with anti-phH3 (blue) antibodies; nuclei are stained with Hoechst (red). BrdU incorporation was measured after incubation for 30 minutes (A-C) or 2 hours (D-F). Zoomed-in images of a selected area indicated by the box are shown. Images are maximum intensity projections of 33 (4.8 µm) optical sections. Scale bars are 40 µm. **(G)** Density plot (*y*-axis) showing the average signal of 3 replicates per genotype over fluorescence signal (DNA stain DRAQ5, *x*-axis; artificial units).

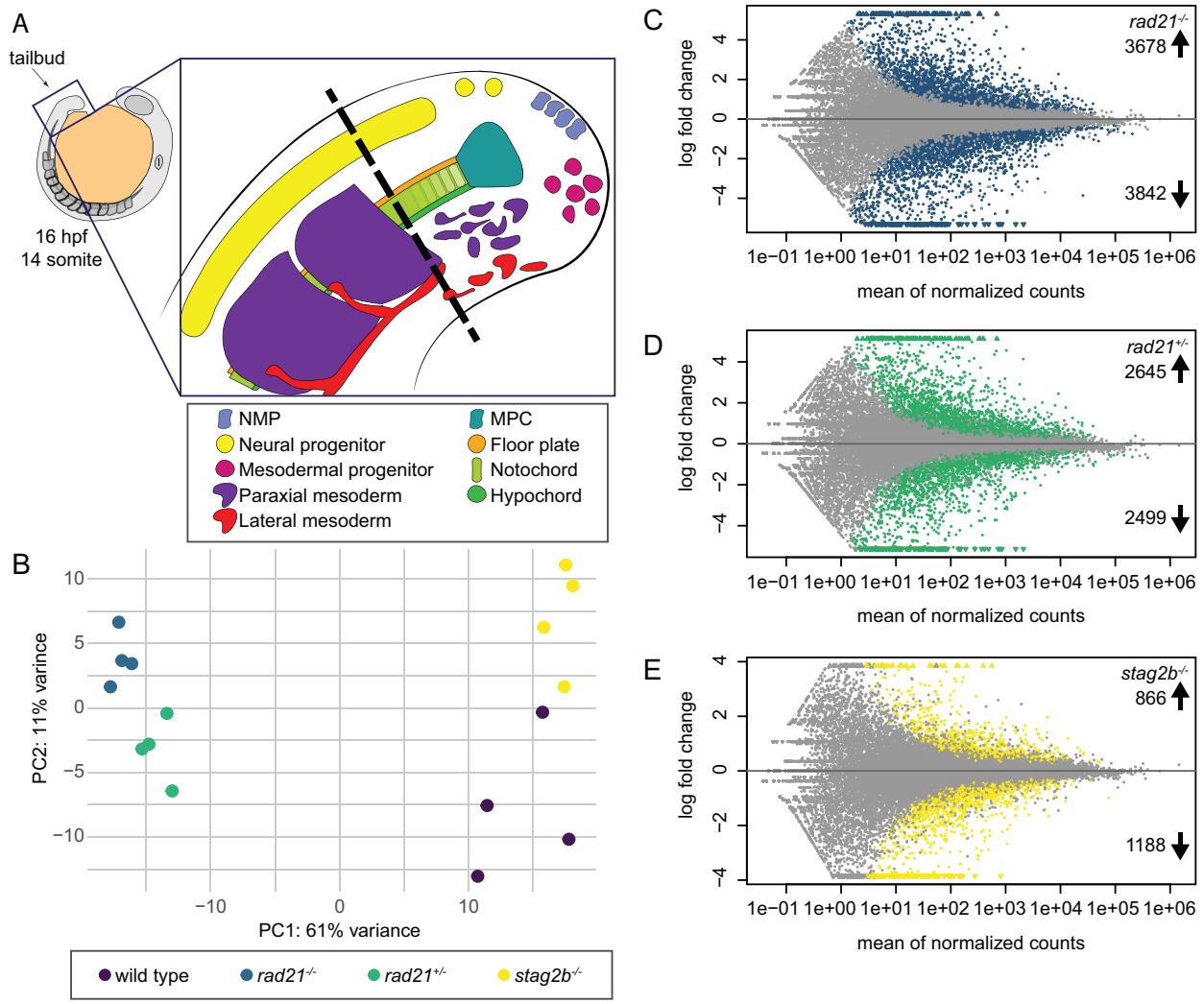


Fig. 3. Bulk RNA sequencing analyses of Rad21- and Stag2b-deficient tailbuds. (A) Schematic representation of progenitor cells and specialized tissues in the zebrafish tailbud. The zebrafish tailbud consists of two pools of bipotent progenitors: neuromesodermal progenitors (NMPs) and midline progenitor cells (MPCs). The dashed line shows the location of tailbud excision for RNA-seq. **(B)** Principal Component Analysis of gene expression in wild type and cohesin deficient tailbuds at the 16-somite stage. Genotypes are distinguished by colour: wild type samples are displayed in purple, *rad21*$^{-/-}$ in blue, *rad21*$^{+/-}$ in green, and *stag2b*$^{-/-}$ in yellow. **(C-E)** The MA (M (log ratio) and A (mean average) scales) plots display changes in gene expression in *rad21*$^{-/-}$ (C), *rad21*$^{+/-}$ (D), and *stag2b*$^{-/-}$ (E) compared to the wild type tailbuds. Each dot represents a gene, with colored dots indicating those with significant (5% false discovery rate, FDR) changes in expression. 7520 genes were dysregulated in *rad21*$^{-/-}$ tailbuds (3678 up- and 3842 downregulated), while 5144 genes were dysregulated in *rad21*$^{+/-}$ tailbuds (2645 up- and 2499 downregulated). In contrast, *stag2b*$^{-/-}$ tailbuds had substantially fewer dysregulated genes (2054: 866 up- and 1188 downregulated).

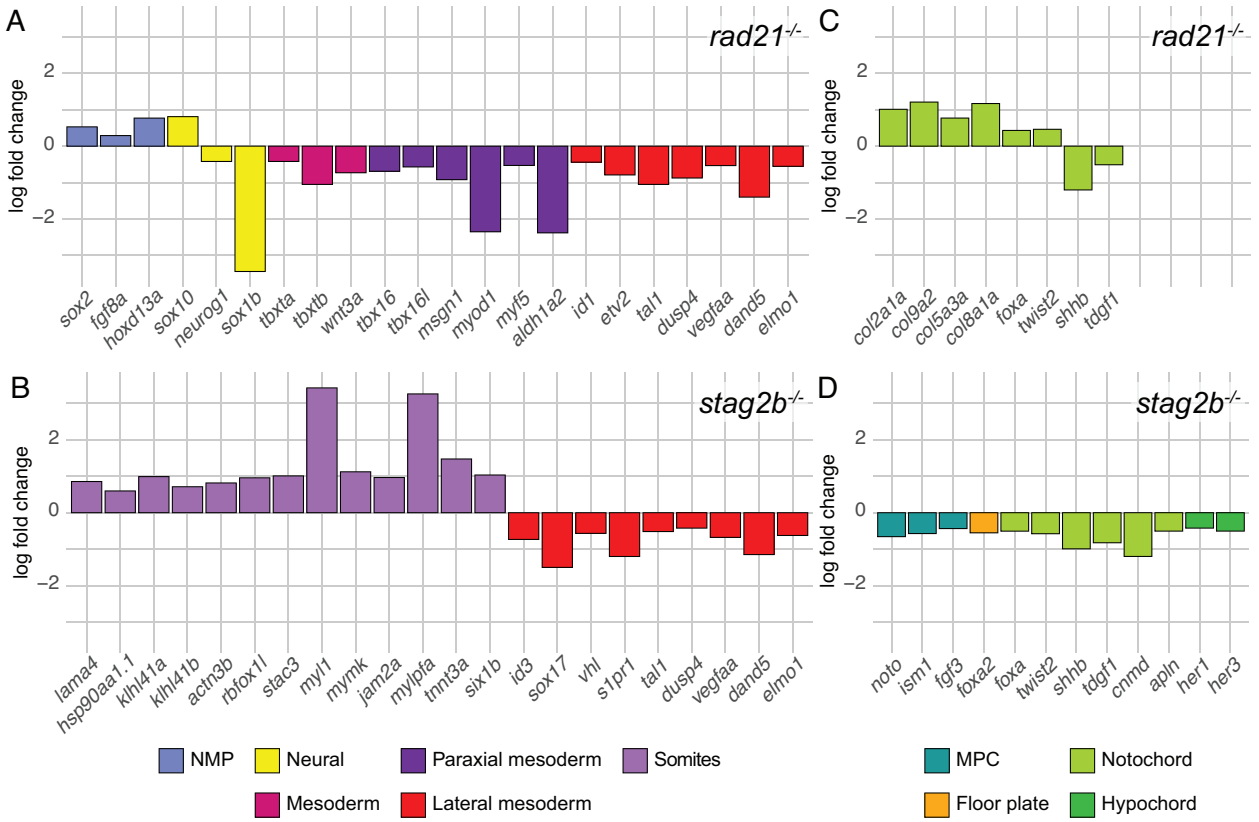


Fig. 4. Expression of genes that mark progenitor cells and their derivatives in *rad21* and *stag2b* homozygous mutant tailbuds. (A-D) The bar graphs display log2 fold changes significantly (5% FDR) dysregulated marker genes in *rad21*$^{-/-}$ (A, C) and *stag2b*$^{-/-}$ (B, D) tailbuds compared to wild type. The different categories of marker genes are represented by different colors as specified in the key.

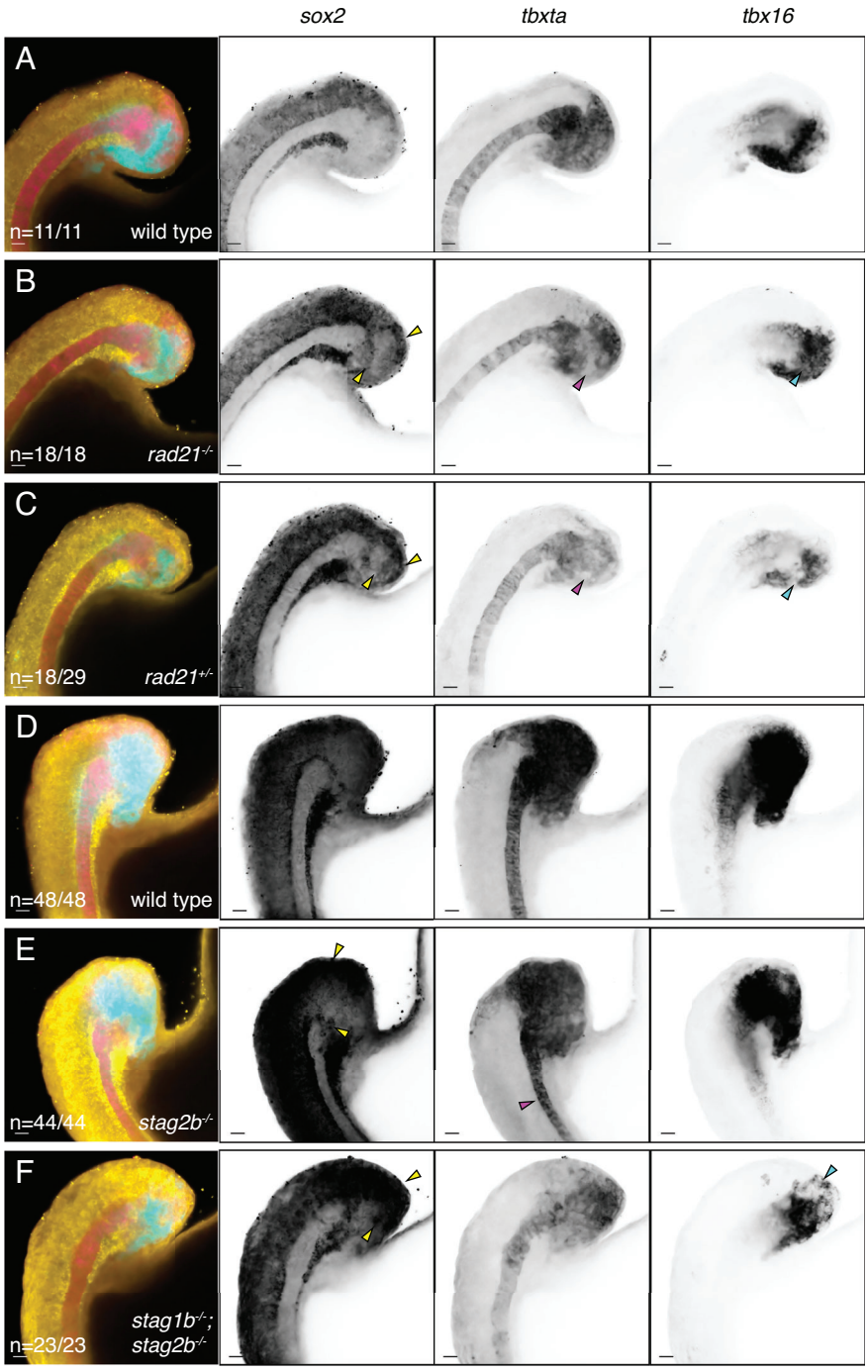


Fig. 5. Distribution of *sox2*, *tbxta*, and *tbx16* transcripts in cohesin-deficient tailbuds. (A, D) wild type, **(B)** *rad21*$^{-/-}$, **(C)** *rad21*$^{+/-}$, **(E)** *stag2b*$^{-/-}$, and **(F)** *stag1b*$^{-/-}$; *stag2b*$^{-/-}$ zebrafish tailbuds at the 16-somite stage showing expression of *sox2* (yellow), *tbxta* (magenta), and *tbx16* (cyan). Increased *sox2* expression in the NMP region and ectopic expression of *sox2* in the mesoderm is indicated by yellow arrows (B, C, E, F). Pink arrows point to the loss of *tbxta* expression in the region of mesodermal induction (B, C) and the narrow notochord (E), while cyan arrows indicate a decrease in *tbx16* expression. Images are maximum intensity projections of 3 (4.8 µm) optical sections. Scale bars are 20 µm. The number of embryos with each expression pattern out of the total analyzed is noted at the bottom left of the merged panels.

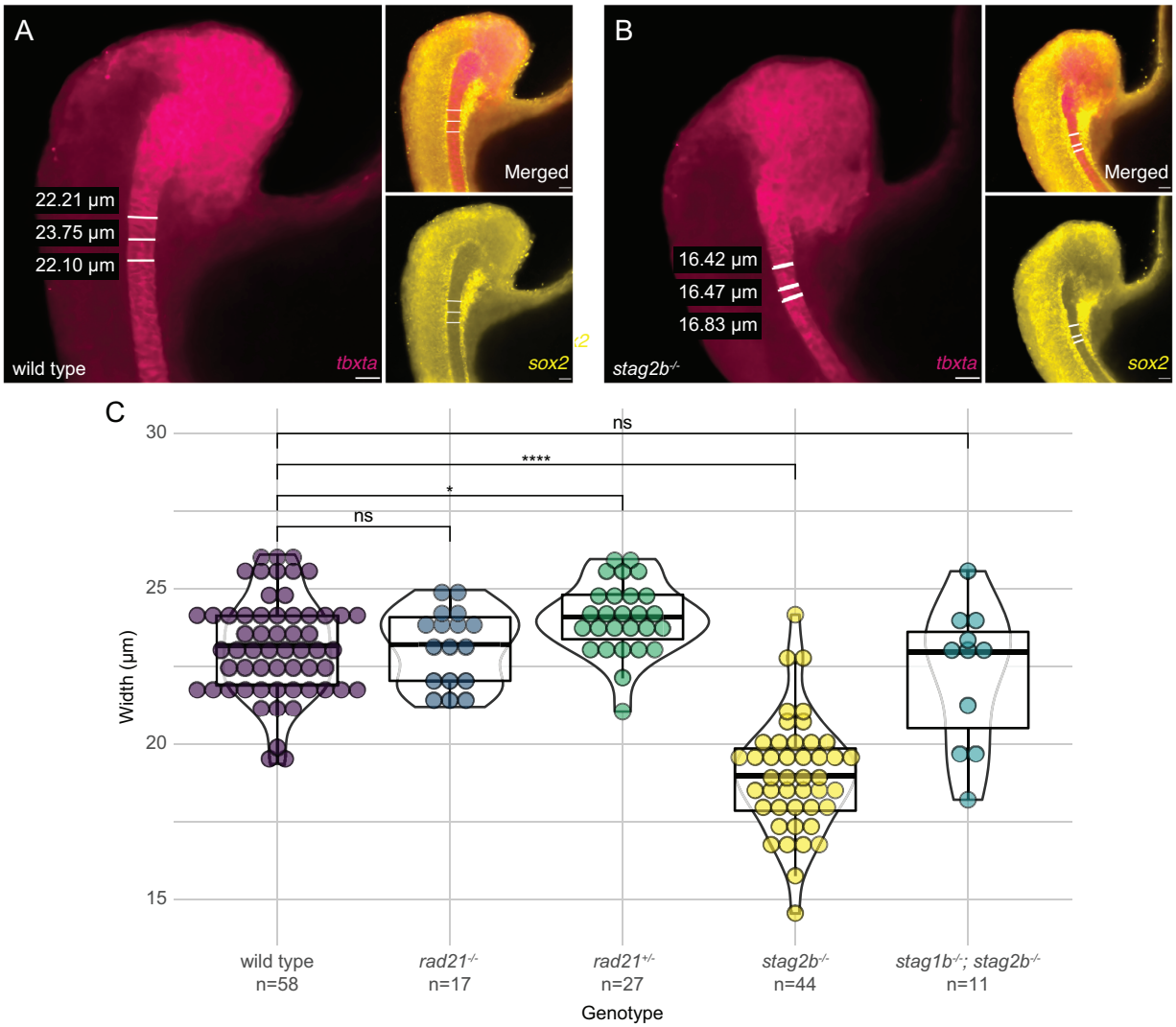


Fig. 6. Narrower notochords in *stag2b* mutants are rescued by additional *stag1b* mutation. (A, B) Example of notochord width measurement using *tbxta* expression and absence of *sox2* expression. The scale bar is 20 µm. **(C)** Violin plots with overlaid box plots visualizing measurements of notochord width. The genotype and the number of embryos measured in each group are indicated on the x-axis. Significance was determined using an unpaired t-test: * $p < 0.05$, **** $p < 0.0001$.

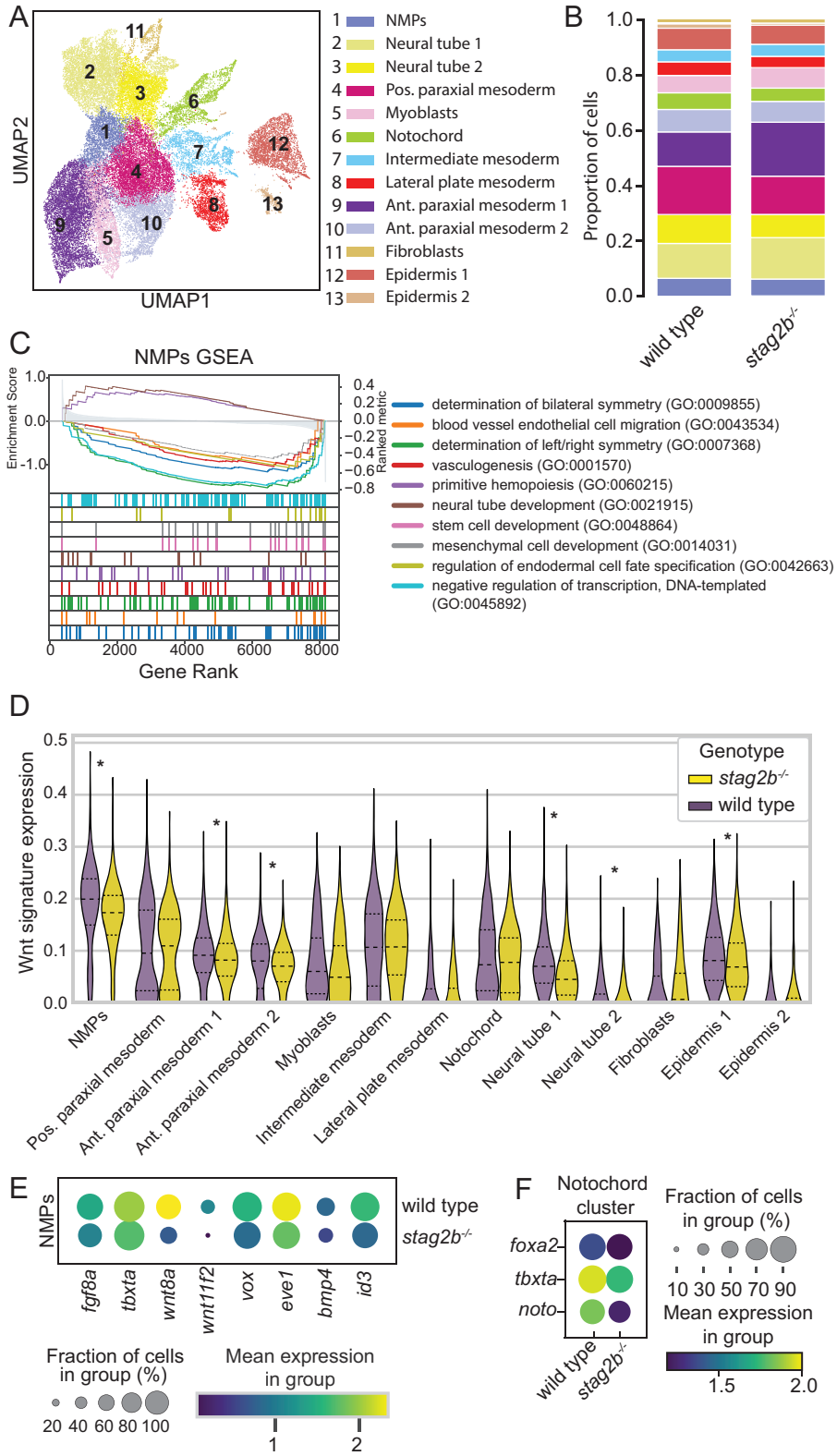
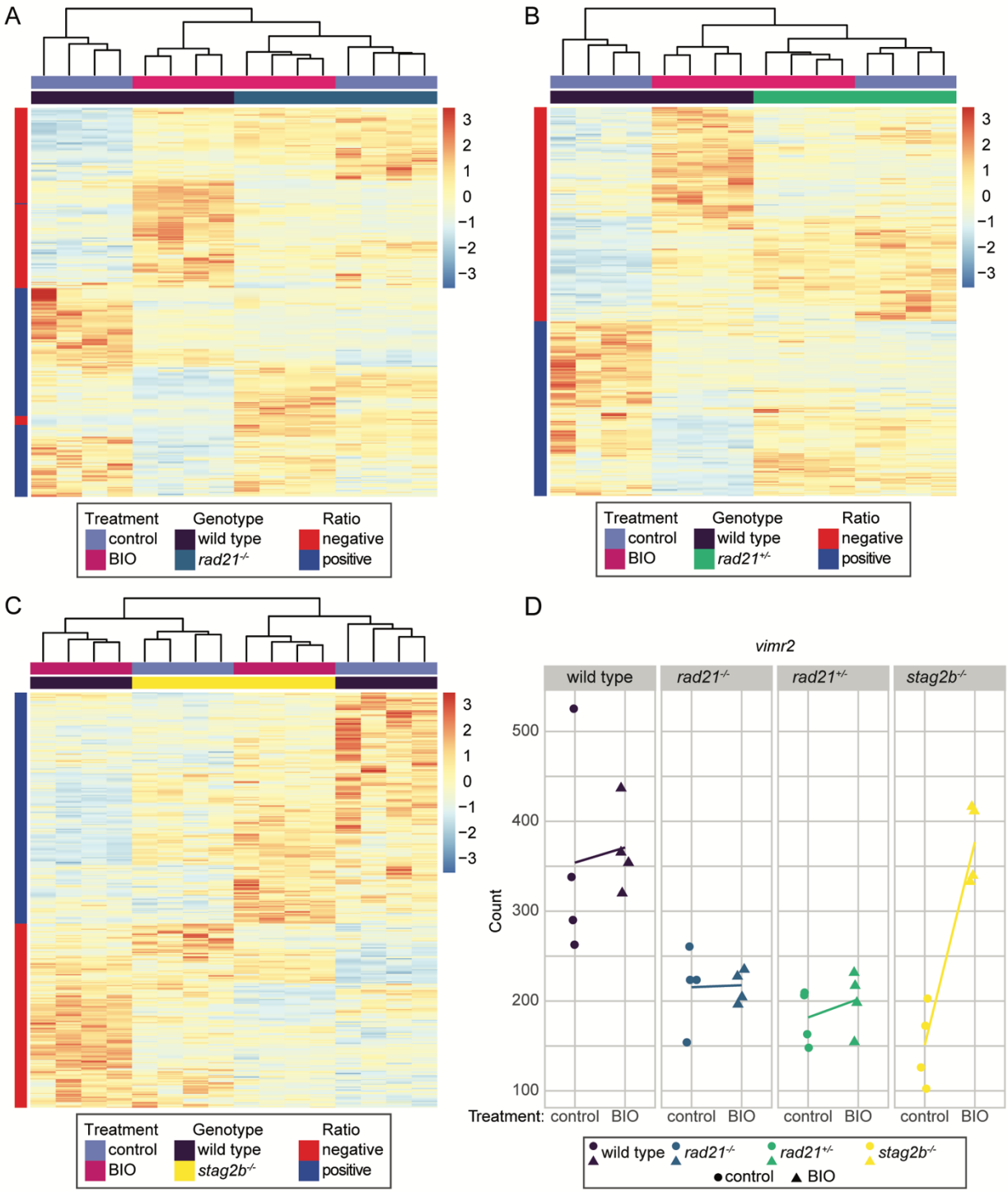


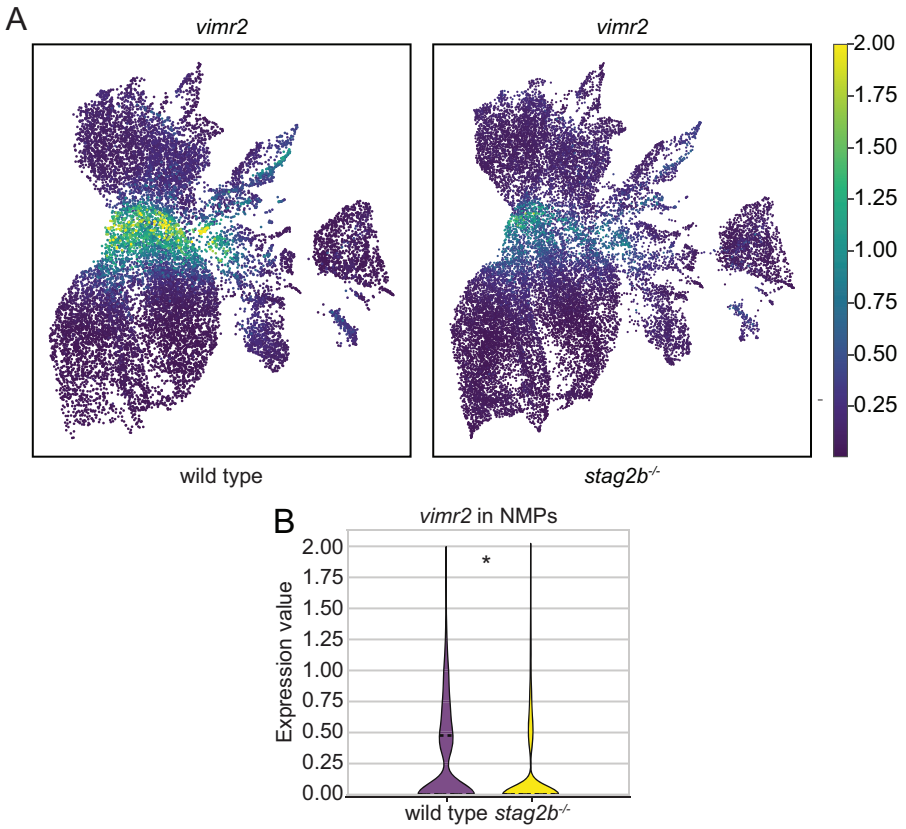
Fig. 7. Single cell RNA-seq of tailbuds from embryos at the 16-somite stage shows disruption of Wnt signaling in *stag2b*$^{-/-}$ NMPs. (A) Uniform manifold approximation and projection (UMAP) dimensional reduction of two integrated datasets of wild type (15,298 cells) and *stag2b*$^{-/-}$ (21,278 cells) tailbud samples (total 36,576 cells) with clustering of the major cell types. **(B)** Stacked bar graph showing cell type proportions inwild type and *stag2b*$^{-/}$ **(C)** Gene set enrichments for genes ranked by Z score for differential expression between wild type and *stag2b*$^{-/}$ NMP cluster. **(D)** Violin plot of Wnt gene expression signature (log-normalized) among different cell types in *stag2b*$^{-/-}$ (yellow) and wild type (purple) embryos. Wilcoxon rank-sum test with 5% FDR. **(E)** Dot plot showing the mean expression of differentially expressed genes

1142 part of FGF, Wnt and BMP pathways in the NMP cluster in wild type and *stag2b*$^{-/-}$ embryos. (**F**) Dotplot of
1143 differentially expressed genes in the notochord cluster.
1144



1145
1146 **Fig. 8. Wnt stimulation normalizes gene expression in *stag2b*$^{-/-}$ but not in *rad21*$^{-/-}$ or *rad21*$^{+/-}$ tailbuds.**
1147 Embryos were treated from shield stage with 2.5 µM BIO, then tailbuds were collected at 16 somites. 4
1148 replicate pools of 80 tailbuds were used per condition for RNA-seq. The heatmaps display expression levels
1149 of the genes that responded differently to BIO stimulation in cohesin mutant genotypes compared with wild
1150 type as determined by an interaction analysis. Heatmaps display results from 4 replicates of **(A)** *rad21*$^{-/-}$, **(B)**
1151 *rad21*$^{+/-}$, and **(C)** *stag2b*$^{-/-}$ versus wild type. Red and blue colors indicate upregulation and downregulation,
1152 respectively, compared to the mean expression. **(D)**, *vimr2* expression is rescued by BIO stimulation in
1153 *stag2b*$^{-/-}$ but not in *rad21*$^{-/-}$ or *rad21*$^{+/-}$. Dot plots illustrate the transcript counts of *vimr2* in wild type (purple),

1154 *rad21*$^{-/-}$ (blue), *rad21*$^{+/-}$ (green), and *stag2b*$^{-/-}$ (yellow). The x-axis indicates the treatment status, and the y-
1155 axis represents the normalised counts. Lines connect the means of the counts for each sample group.
1156



1157
1158 **Fig. 9. *stag2b* mutation affects *vimr2* expression in NMPs. (A)** Expression of *vimr2* in UMAP
1159 representation in wild type and *stag2b*$^{-/-}$ tailbuds at 16 somites. **(B)** Violin plot showing downregulation of
1160 *vimr2* expression in the NMPs in *stag2b*$^{-/-}$. * $p < 0.005$, Wilcoxon rank-sum test.
1161

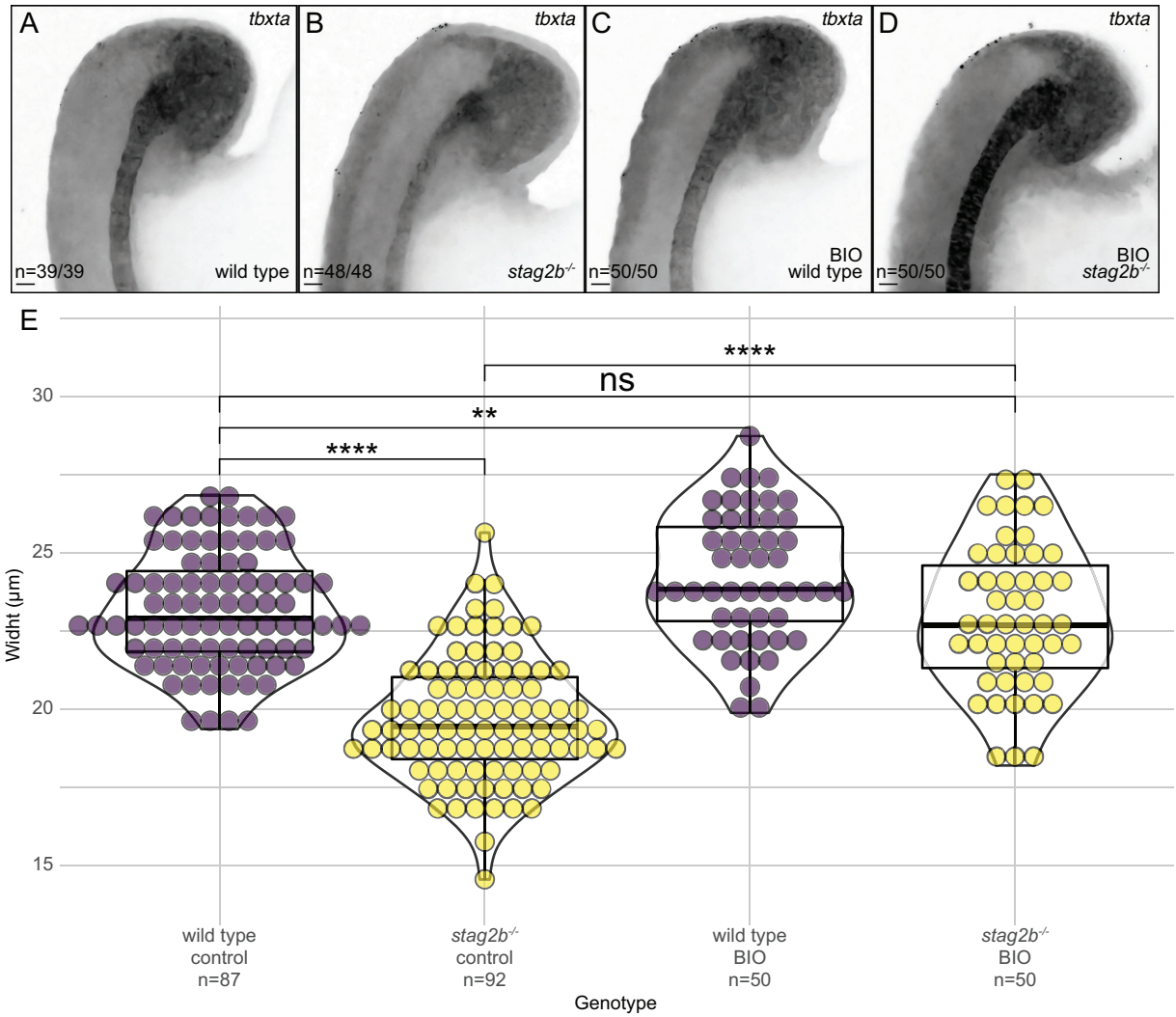


Fig. 10. Wnt stimulation rescues notochord width in *stag2b*$^{-/-}$. (A-D) Expression pattern of *tbxta* in wild type (A, C) and *stag2b*$^{-/-}$ (B, D) zebrafish tailbuds with (C, D) and without (A, B) Wnt stimulation. Images are maximum intensity projections of 3 (4.8 µm) optical sections. Scale bars are 20 µm. The number of embryos with each expression pattern out of the total analyzed is noted. **(E)** Violin plots with overlaid box plots visualizing measurements of notochord width. The x-axis indicates the genotype, treatment status and the number of embryos measured in each group. Significance was determined using an unpaired t-test: ** $p < 0.01$, **** $p < 0.0001$.

Supplementary Information

Cohesin complex quantity versus composition independently affect early development in zebrafish

Anastasia Labudina *et al.*

*Corresponding author. Email: julia.horsfield@otago.ac.nz

This PDF file includes:

Figs. S1 to S13

Other data for this manuscript include the following:

Excel table containing Data S1 to S9

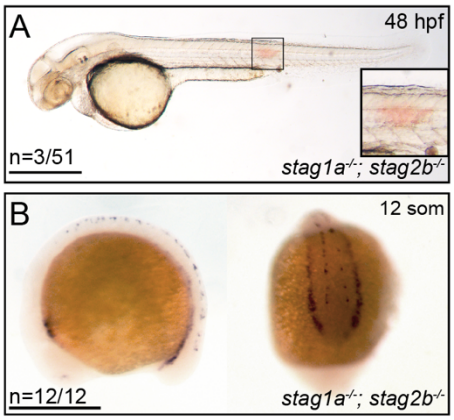


Fig. S1. Trunk haemorrhaging and *runx1* expression in *stag1a*$^{-/-}$*; stag2b*$^{-/-}$ mutant embryos.
(A) Lateral views of representative and *stag1a*$^{-/-}$*; stag2b*$^{-/-}$ embryos at 48 hpf. The boxed region (inset) outlines a trunk haemorrhage representative of those observed in around 5% of the *stag1a*$^{-/-}$*; stag2b*$^{-/-}$ double mutant embryos. Scale bars are 500 µm.
(B) Normal expression of *runx1* at 12 somites in *stag1a*$^{-/-}$; *stag2b*$^{-/-}$ embryos. Lateral (left) and posterior (right) views are shown. Scale bars are 500 µm. The numbers in the lower left hand corner indicate the number of embryos with the expression pattern shown.

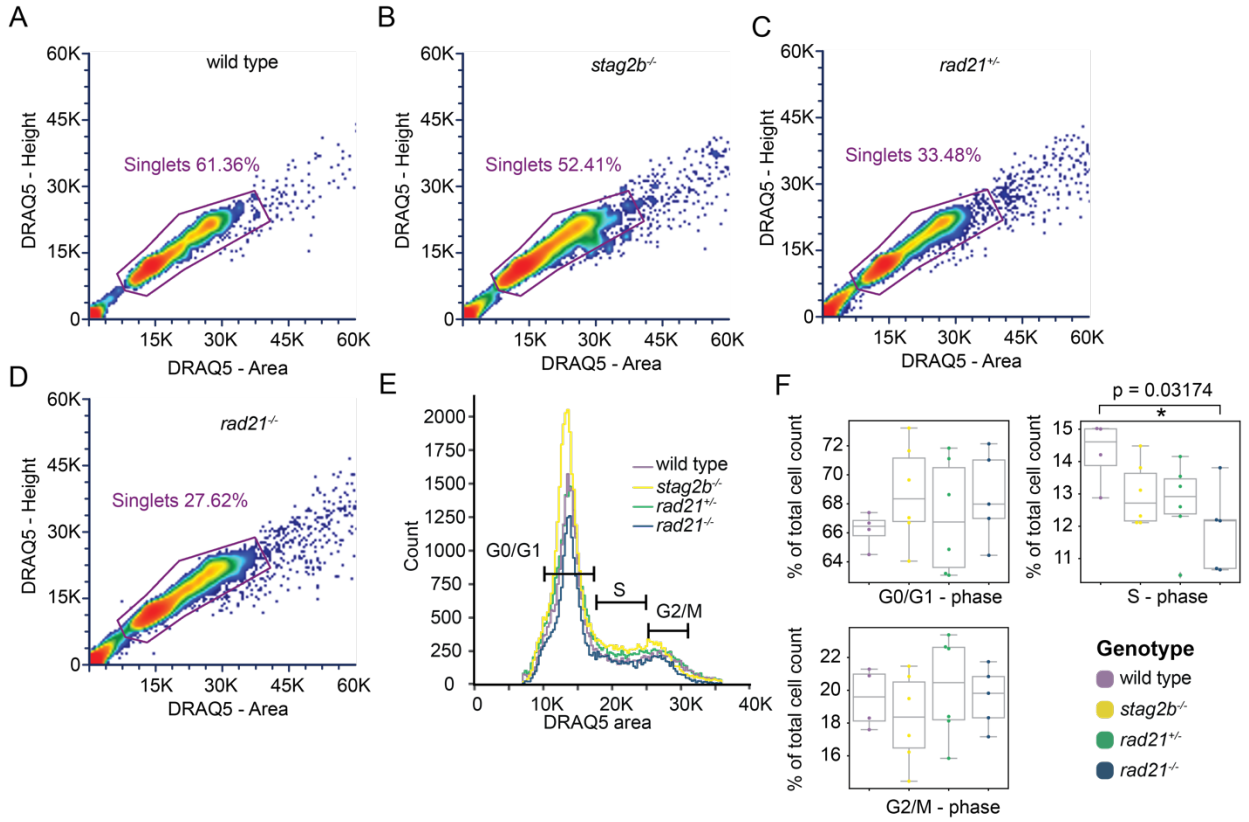


Fig. S2. Flow cytometry analyses show that *rad21*$^{-/-}$ tailbuds have fewer cells in S phase, while other cohesin-mutant genotypes resemble wild type. (A-D) Pseudocolour dot plot of flow cytometry data showing density of cells across DNA content shown as area (*x*-axis) and height of the signal (*y*-axis) of DRAQ5 (labelling DNA) for wild type (A), *stag2b*$^{-/-}$ (B), *rad21*$^{+/-}$ (C) and *rad21*$^{-/-}$ (D) tailbud cells. **(E)** Count density plot of wildtype (purple), *stag2b*$^{-/-}$ (yellow), *rad21*$^{+/-}$ (green) and *rad21*$^{-/-}$ (blue) combined replicate samples (n>4). **(F)** Box and swarm plots showing the proportions of cell cycle phases as G0/1, S and G2/M based on DNA content. Boxes show quartiles, whiskers show 1.5 inter-quartile ranges of the lower and upper quartile. *P < 0.05 Mann-Whitney U test.

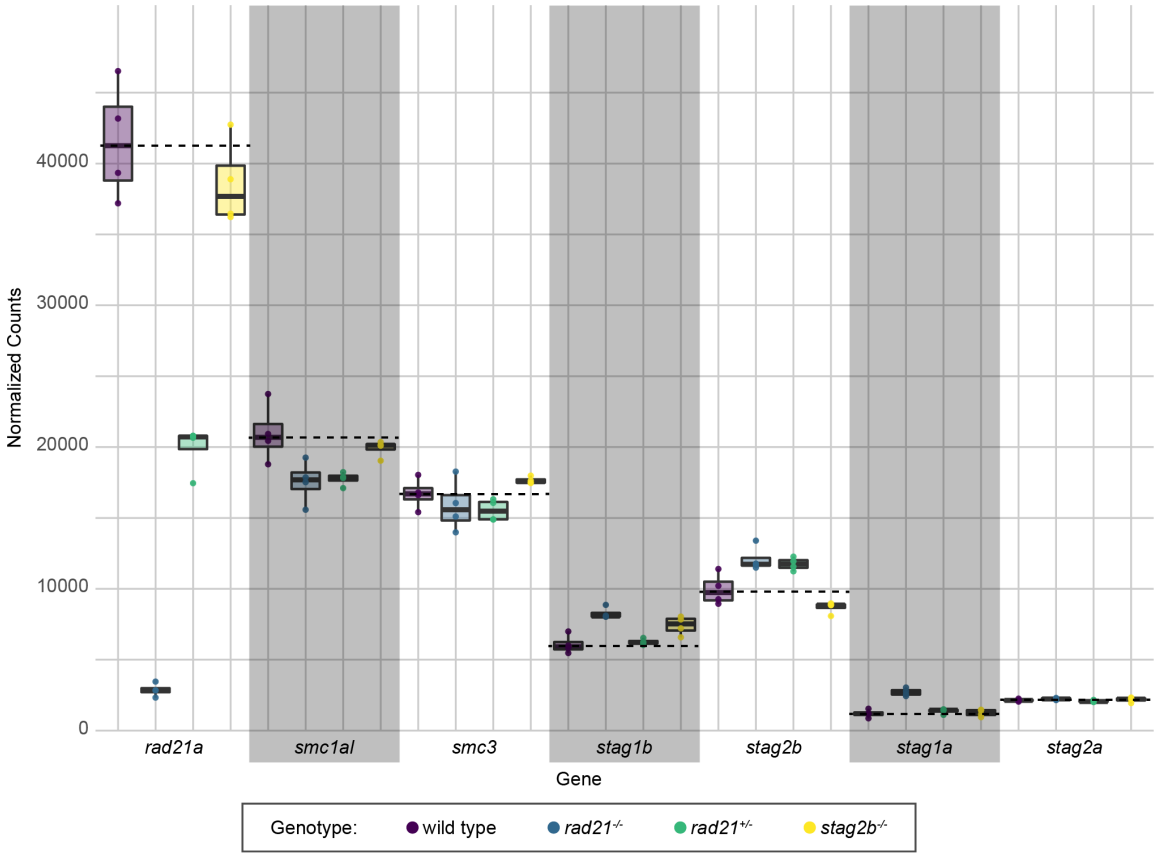


Fig. S3. Transcript counts of cohesin subunits in the tailbud.
Normalised transcript counts of *rad21a*, *smc1al*, *smc3*, *stag1b*, *stag2b*, *stag1a* and *stag2a* taken from RNA-seq of tailbuds and visualised with box plots defining the distribution of expression levels across these genes in the samples. Genotypes are distinguished by colour: wild type samples are displayed in purple, *rad21*$^{-/-}$ in blue, *rad21*$^{+/-}$ in green, and *stag2b*$^{-/-}$ samples in yellow. For each panel, the dotted line indicates the level of expression in wild type.

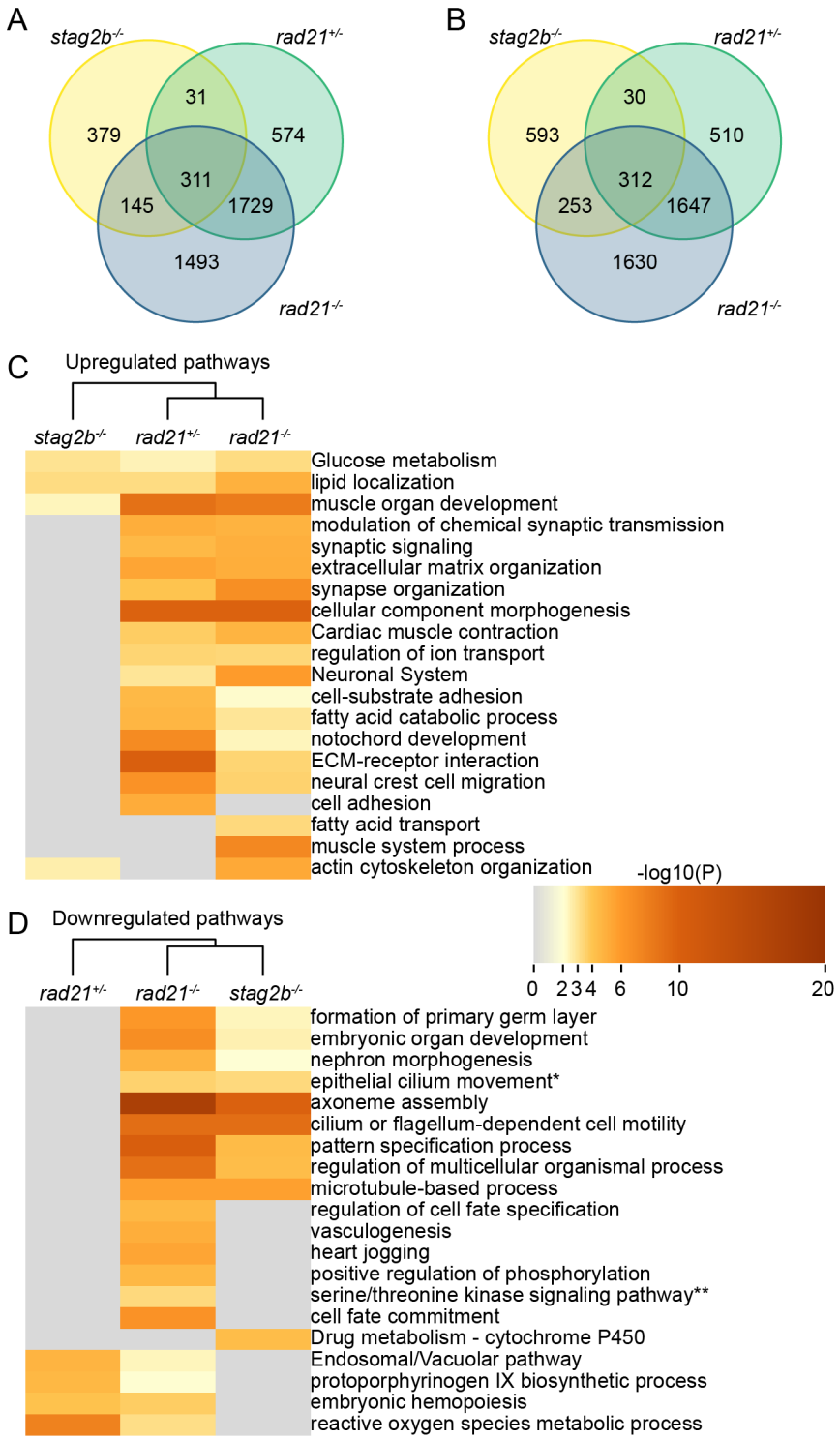


Fig. S4. Overlap of dysregulated genes and pathway enrichment in cohesin mutant tailbuds.

(A, B) The Venn diagrams depict the overlap of significantly upregulated (A) and downregulated (B) genes in cohesin deficient tailbuds. **(C, D)** Metascape heat maps displaying the top 20 terms enriched among significantly upregulated (C) and downregulated (D) genes in cohesin-deficient tailbuds. Corresponding *p*-values are indicated by the colour scale.

*Epithelial cilium movement involved in extracellular fluid movement
**Regulation of transmembrane receptor protein serine/threonine kinase signalling pathway

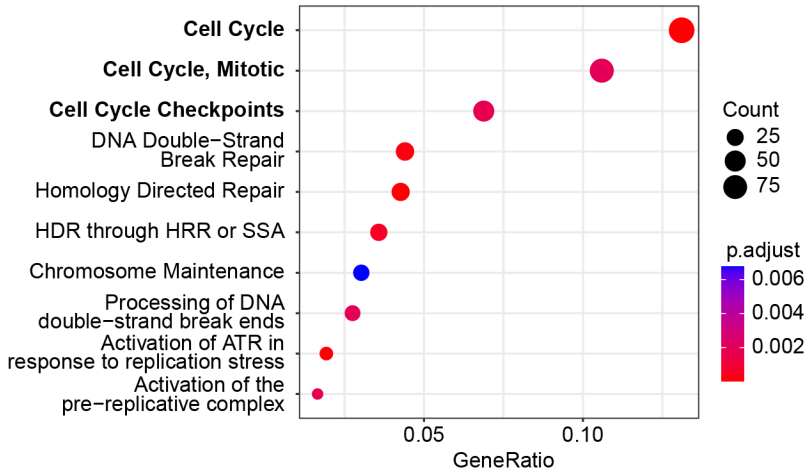


Fig. S5. Reactome analyses of downregulated pathways in *rad21*$^{-/-}$ mutant tailbuds. The dot plot shows the top 10 enriched Reactome pathways (out of 26) among the significantly downregulated genes in *rad21*$^{-/-}$ tailbuds. The size of each dot indicates the number of genes affected in the pathway, and the dot colour represents the adjusted p-value (Padj). Source data available in Supplementary Data 1-3.

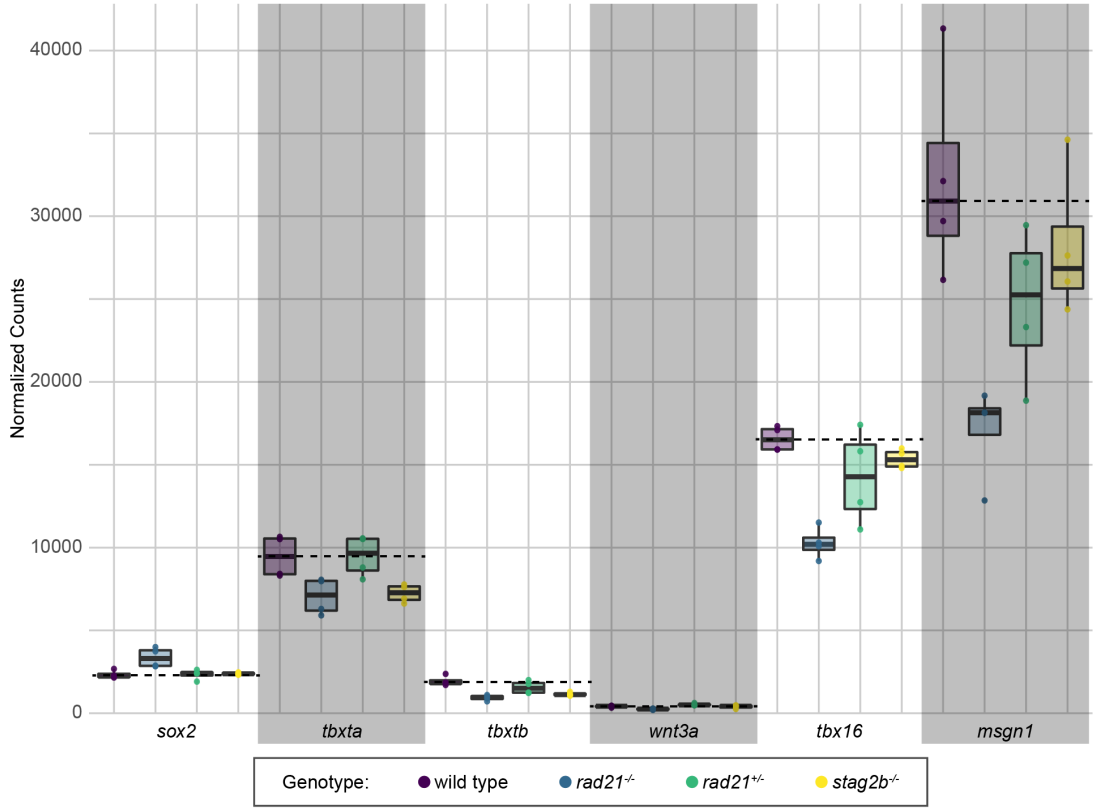


Fig. S6. Expression levels of *sox2*, *tbxta*, *tbxtb*, *wnt3a*, *tbx16* and *msgn1* in cohesin mutants. Normalised transcript counts of *sox2*, *tbxta*, *tbxtb*, *wnt3a*, *tbx16* and *msgn1* taken from tailbud RNA-seq data with 4 replicates visualised as box plots defining the distribution of expression levels across these genes in the samples. Genotypes are distinguished by colour: wild type samples are displayed in purple, *rad21*$^{-/-}$ in blue, *rad21*$^{+/-}$ in green, and *stag2b*$^{-/-}$ samples in yellow.

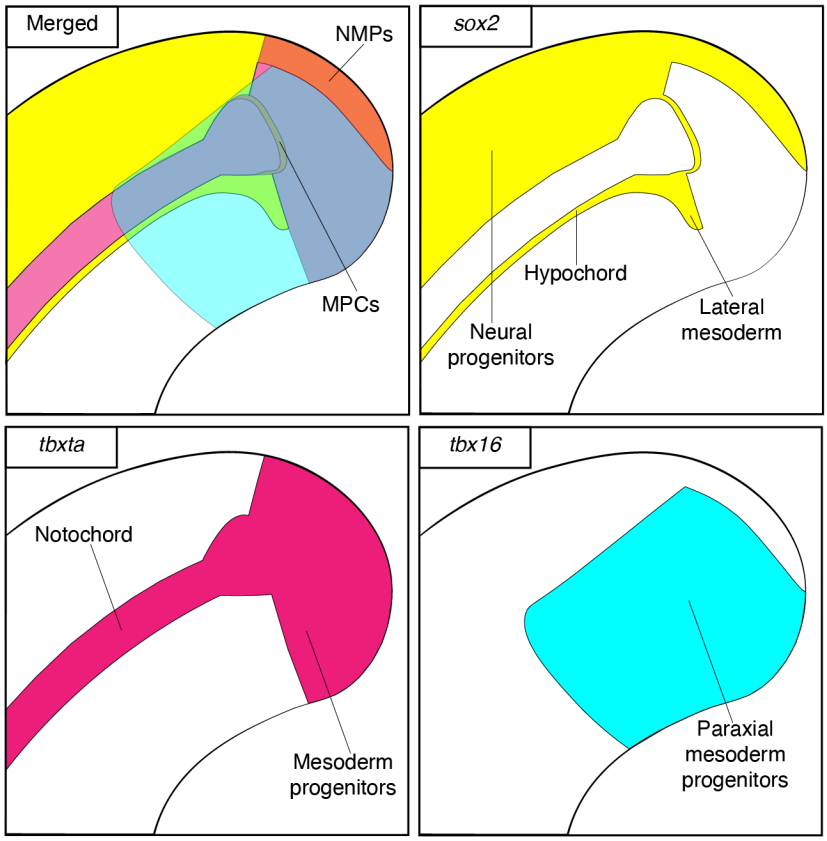


Fig. S7. Expression pattern of marker genes in the tailbud at the 16-somite stage. Schematic depicts the regions in the tailbud and progenitor types where expression of *sox2* (yellow), *tbxta* (magenta) and *tbx16* (cyan) expression is expected.

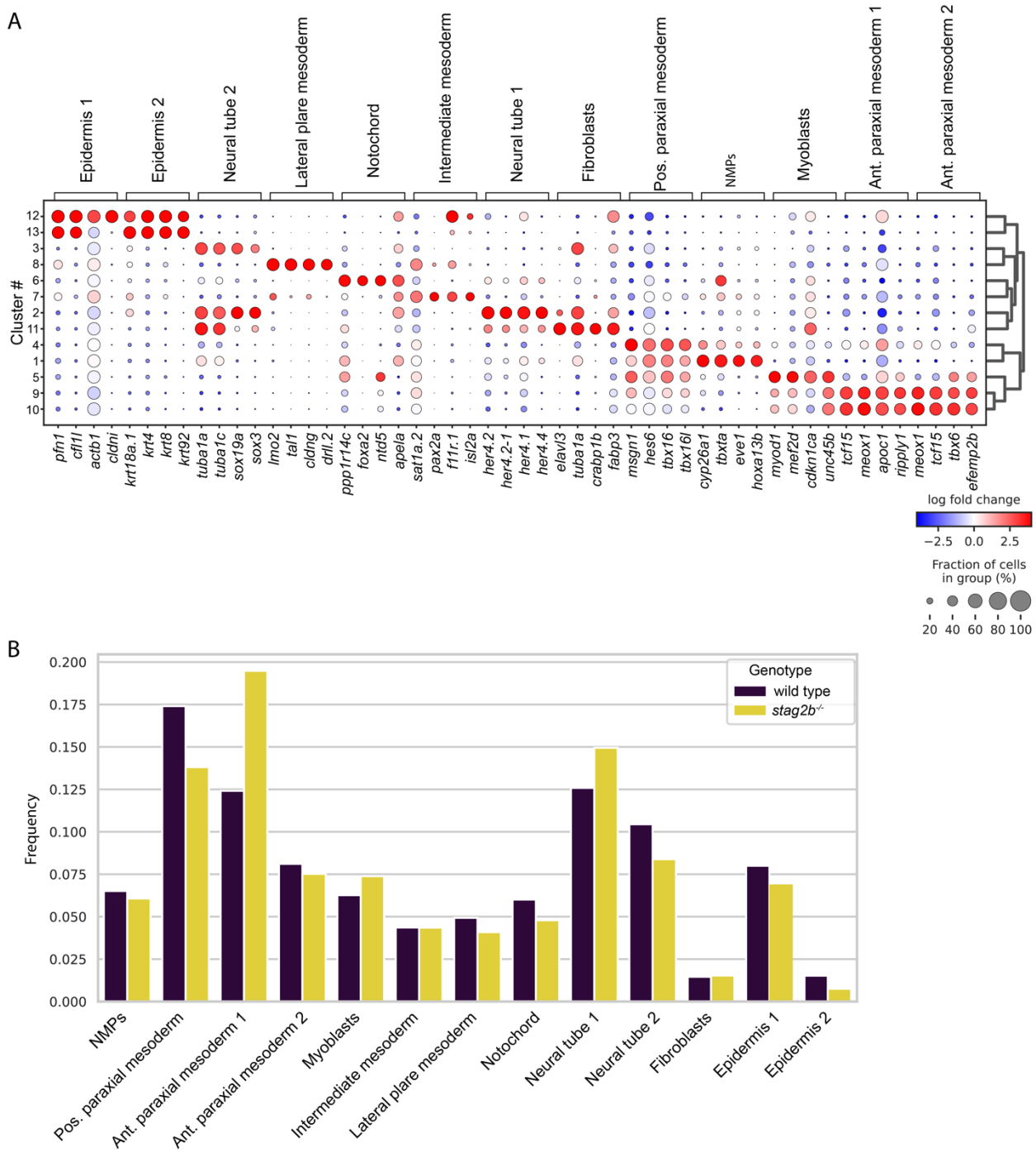


Fig. S8. Cell population analysis of single cell RNA sequencing of *stag2b*$^{-/-}$ mutant tailbuds compared with wild type.

(A) Dot plot depicting the expression of the top 4 marker genes per cell population cluster identified (see Fig. 7A). The dot size scales with the fraction of cells expressing the gene, and the dot colour indicates the log fold change between the clusters. **(B)** Differences in the proportion of cell types between wild type and *stag2b*$^{-/-}$ tailbuds among clusters as shown in Fig. 7A.

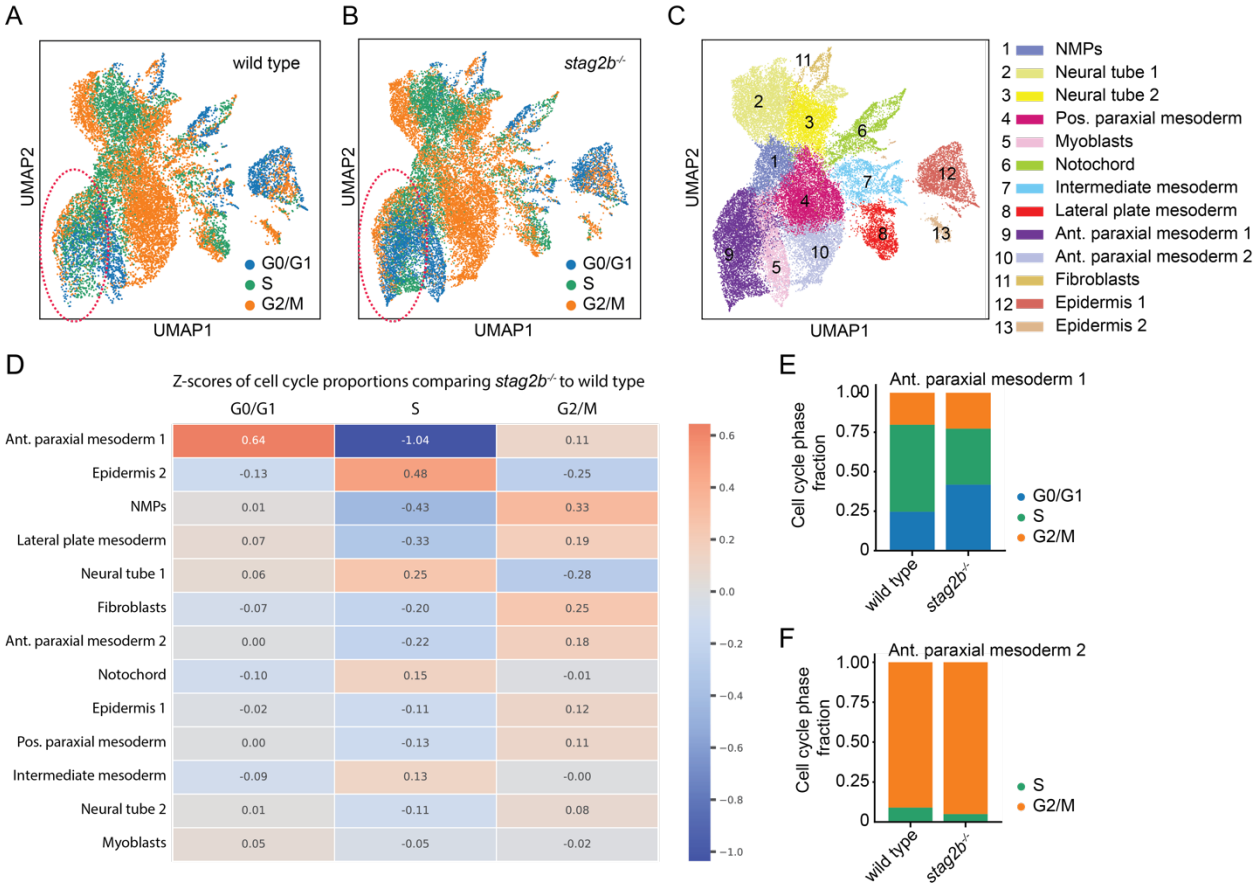


Fig. S9. Cell cycle phase analysis of single cell data from *stag2b*$^{-/-}$ and wild type tailbuds. **(A-B)** UMAP of wild type and *stag2b*$^{-/-}$ single cell data showing cell cycle phases in blue (G0/G1), green (S) and orange (G2/M). Dashed red outline indicates the anterior paraxial mesoderm cluster 1. All cluster annotations for the integrated dataset in **(C)**. **(D)** Ranked heatmap of z-scores of *stag2b*$^{-/-}$ cell clusters comparing cell cycle phase proportions. Red indicates an increase and blue a decrease in cell cycle phase compared to wild type. **(E)** Stacked bar plot of cell cycle phase fractions comparing anterior paraxial mesoderm 1 cluster in wild type with *stag2b*$^{-/-}$. **(F)** Stacked bar plot of cell cycle phase fractions comparing anterior paraxial mesoderm 2 cluster in wild type with *stag2b*$^{-/-}$.

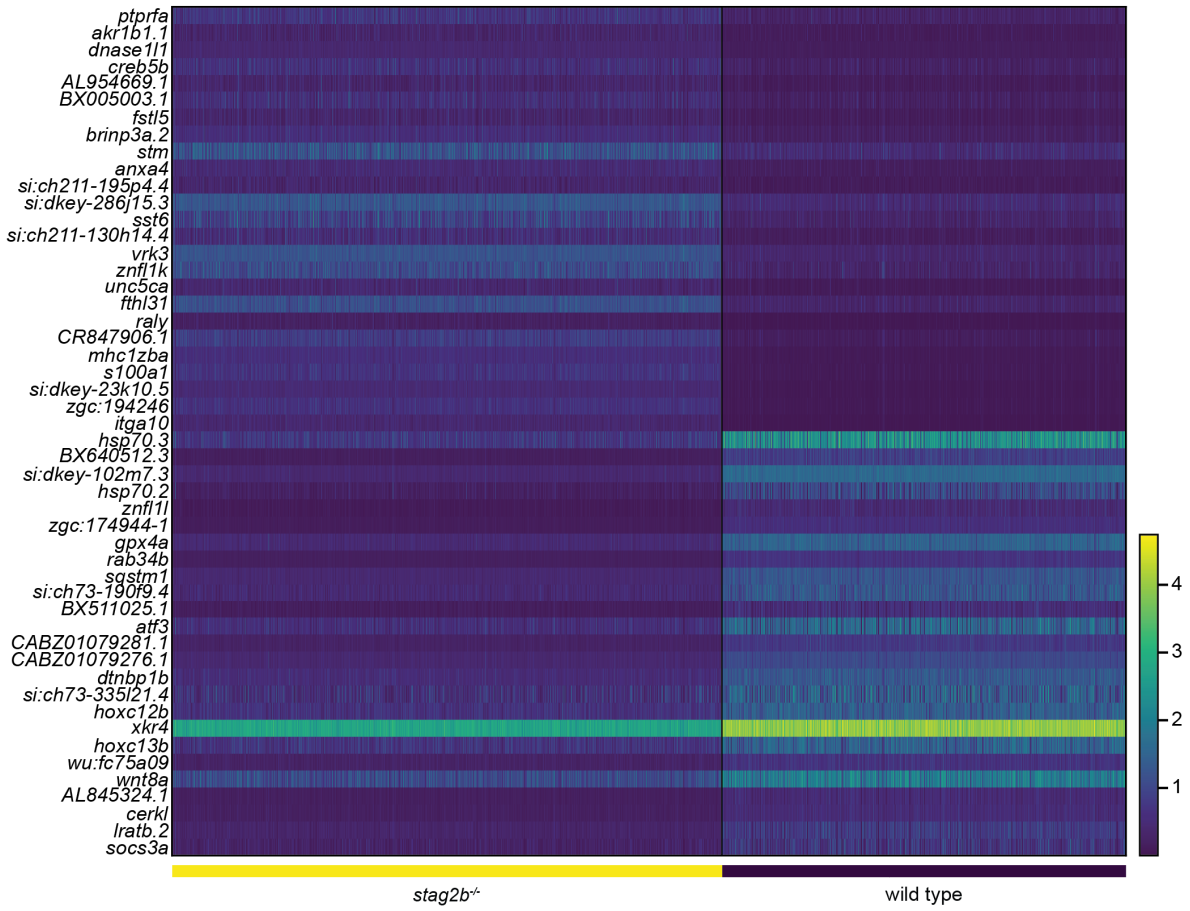


Fig. S10. Heat map showing the top 25 differentially up- and downregulated genes between wild type and *stag2b*$^{-/-}$ NMP single cell RNA-seq data. For a full list see Supplementary Data 4.

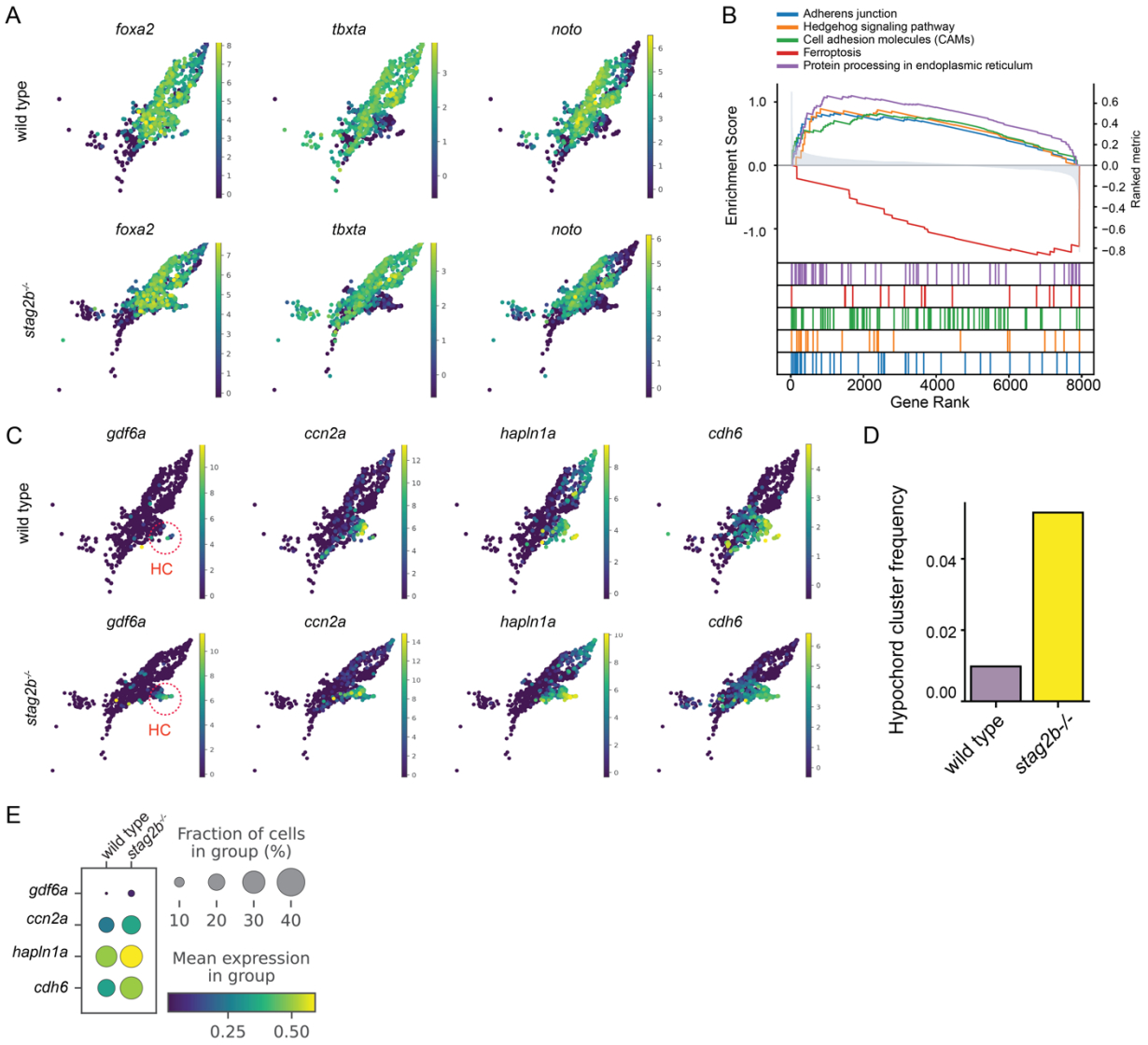


Fig. S11. Notochord single cell analysis shows upregulation of Hedgehog signalling with an expansion of hypochord cells in *stag2b*$^{-/-}$ tailbuds. **(A)** UMAP plot of notochord markers *fox2a, tbxta* and *noto* showing their expression in wild type (top) and *stag2b*$^{-/-}$ (bottom) notochord cluster. **(B)**. Gene set enrichment analysis of differentially expressed genes derived from pseudo bulk analysis comparing the notochord cluster in wild type and *stag2b*$^{-/-}$ tailbud single cell data. **(C)** UMAP of hypochord markers *gdf6a, ccn2a, hapln1a* and *cdh6* in wild type (top) and *stag2b*$^{-/-}$ (bottom) notochord cluster. **(D)** Bar plot of hypochord-marked cells as proportion of cells in the notochord cluster in wild type and *stag2b*$^{-/-}$ tailbud single data. **(E)** Dot plot of hypochord markers in the notochord cluster comparing wild type and *stag2b*$^{-/-}$ tailbuds.

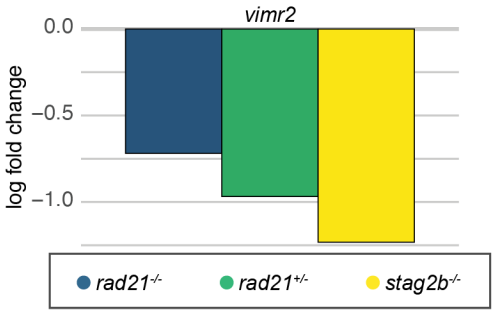


Fig. S12. *vimr2* expression in cohesin mutants. The bar graph displays the log2 fold change (5% FDR) for *vimr2* transcripts in *rad21*$^{-/-}$ (blue), *rad21*$^{+/-}$ (green) and *stag2b*$^{-/-}$ (yellow) tailbuds compared to wild type.

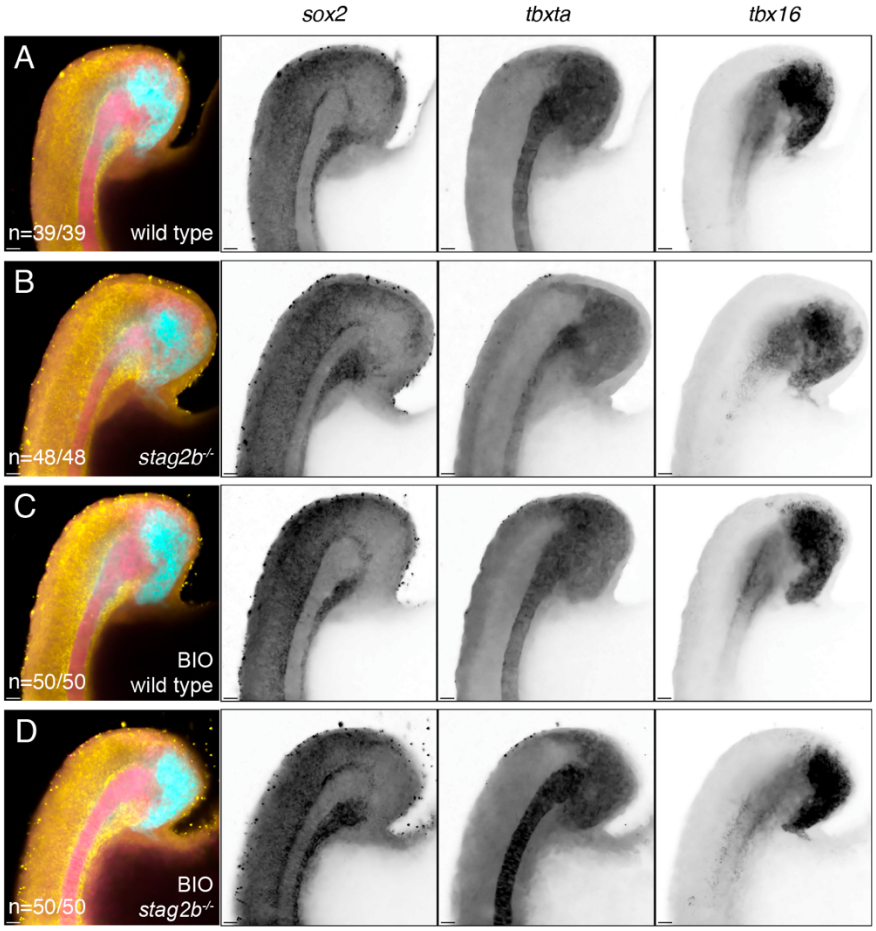


Fig. S13. Wnt stimulation rescues the notochord phenotype in *stag2b*$^{-/-}$ mutants.
(A-D) Expression patterns of *sox2*, *tbxta*, and *tbx16* in wild type (A, C) and *stag2b*$^{-/-}$ (B, D) zebrafish tailbuds with (C, D) and without (A, B) Wnt stimulation. Images are maximum intensity projections of 3 (4.8 µm) optical sections. Scale bars are 20 µm. The number of embryos with each expression pattern out of the total analysed is noted.

Data S1 (separate file): Cell cycle genes (Reactome PA) significantly (5% FDR) dysregulated in homozygous mutant *rad21* tailbuds.

Data S2 (separate file): Cell cycle genes (Reactome PA) significantly (5% FDR) dysregulated in heterozygous mutant *rad21* tailbuds.

Data S3 (separate file): Cell cycle genes (Reactome PA) significantly (5% FDR) dysregulated in homozygous mutant *stag2b* tailbuds.

Data S4 (separate file): Significantly (5% FDR) differentially expressed genes in the NMP subset comparing homozygous *stag2b*$^{-/-}$ and wild type tailbuds, related to Figure 7C,E and Figure S10.

Data S5 (separate file): Data S5: Gene set enrichment analysis results comparing NMPs of homozygous *stag2b*$^{-/-}$ and wild type tailbuds, related to Figure 7C.

Data S6 (separate file): Data S6: Gene list related to canonical Wnt signalling score, related to Figure 7D.

Data S7 (separate file): Wilcoxon rank sum test statistics for each cell cluster comparing the Wnt signaling score shown in Figure 7D among homozygous *stag2b*$^{-/-}$ and wild type genotypes.

Data S8 (separate file): Significantly (5% FDR) differentially expressed genes in the notochord subset comparing homozygous *stag2b*$^{-/-}$ and wild type tailbuds, related to Figure S11.

Data S9 (separate file): Gene set enrichment analysis results comparing the notochord cluster of homozygous *stag2b*$^{-/-}$ and wild type tailbuds, related to Figure S11.