Discovery of novel functions of minor ZGA gene on heart and vascular regulation in Golden pompano genome provides clues for anti-cancer therapeutic strategies

Honglin luo[1, 3 #], Yongde Zhang[1, #], Changmian Ji [2, #], Yongzhen Zhao[1, #], Jinxia Peng [1, #], Xiuli Chen[1], Lequn Li [3], Yin Huang[1], Qingyun Liu[1], Pingping He[1], Pengfei Feng[1], Chunling Yang[1], Pinyuan Wei[1], Haiyan Yu[2], Hongkun Zheng[2, *], Yong Lin[1, *], Xiaohan Chen[1, *]

[1]Guangxi Key Laboratory for Aquatic Genetic Breeding and Healthy Aquaculture, Guangxi Institute of Fishery Sciences, Nanning, 530021, China

[2] Biomarker Technologies, Beijing, 101300, China.

[3] Department of Hepatobiliary Surgery, Affiliated Tumor Hospital of Guangxi Medical University, Nanning 530021, PR China.

[#] These authors contributed equally to this work.

[*]Prime correspondence to: Xiaohan Chen. Hongkun zheng and YongLin are also corresponding authors

Running title: Golden pompano genome provides new insights into heart and vascular development and clues for anti-cancer therapeutic strategies

Corresponding authors addresses and emails:

NO. 8, Qingshan Road, Guangxi Key Laboratory for Aquatic Genetic Breeding and Healthy Aquaculture, Guangxi Institute of Fishery Sciences, Nanning, 530021, China.

Xiaohan Chen: Chnxhn@163.com; Hongkun zheng: zhenghk@biomarker.com.cn; Yonglin: linnn2005@126.com.

The development of heart- and vascular-network is essential for embryonic survival and organ function. However, whether the genes in the minor zygotic genome activation (ZGA) shaped by a biased-subgenome affect the ZGA and have functions on heart and vascular development are largely unknown. In this study, we discover a novel gene *nxh* from a group of hub-genes within minor ZGA shaped by a WGD-driven-subgenome of the *de novo* golden pompano genome. We find that loss of its homolog *nxhl* in zebrafish not only delays the ZGA activation, but causes pericardial oedema, caudal fin defects, and impairs the growth of intersegmental vessel and caudal vein plexus. Furthermore, loss of *nxhl* reduces the heart- and vascular-associated landmark gene network (including keystone gene *ptprb, hand2, tie2, hey2,* and *s1pr1*, etc.) via a hitherto unreported *nxhl*-nucleolin-*ptprb* pathway. We also find that Harbi1 expression in cancer tissue is significantly higher than para-cancer tissue, and silencing of Harbi1, the human homolog of *nxh*, suppress the angiogenesis of the HUVECs, and migration and EMT of hepatocellular carcinoma cells, indicating its potential as an anti-cancer therapeutic target. Our findings provide new insights into cardio- and vascular-linked functions of the genes in the WGD shaping minor ZGA during marine fish embryo development and clues for anti-cancer therapeutic strategies.

.

Heart and vascular development are essential for embryonic survival and organ function throughout life and must be precisely programed. Heart and vascular development are closely linked[1,2]. In zebrafish, heart development starts at 16 hpf[3], while vascular development begins at 24 hpf[4] in zebrafish. Both progresses are logically affected by the global zygotic genome activation (ZGA,3 hpf[5]) process. ZGA is a vital process for subsequent embryo development[6] (for a review of the mechanisms underlying ZGA, see ref[7]). Although the timing of ZGA and from one to another transcriptional activation vary in different vertebrate species (reviewed in refs[7,8]), studies using genomic approaches confirmed that a minor wave of ZGA (a low level, yet distinct, early activation of transcription) exists prior to the major wave of ZGA[9]. The latest study in mice demonstrated that transient inhibition of minor ZGA results in compromised development of major ZGA and subsequent stages[10], suggesting the minor ZGA may contribute to signaling the further activation of the genome. Therefore, no matter in fish or mammals, the genes affecting minor ZGA process might impact the heart and vascular development even the following cell or organ biological functions.

On the other hand, vessels are tightly associated with tumor or cancer progression. Vessel propagations facilitate the malignant of cancers[11]. Targeting the cancer angiogenesis or vessels has been shown to be an effective way to fight against cancers[12]. The molecules that can inhibit the vascular development have potential value on the anti-cancer therapy[13]. Considering the force of conservative evolution, we supposed that genes associated with vascular development in lower vertebrates, such as fish, might keep the corresponding core functions in the higher vertebrates, such as humen. We tried to find such connections between fish and humen.

It has been evidenced that teleosts underwent a third round of while genome duplication (WGD; Ts3R, about ~320 million years ago (Mya)[14,15], which is thought to have been a driver for the radiation and diversification of teleosts[16] and to be a cause of phenotypic innovation and complexity[17,18]. After WGD, many of the duplicated genes are lost. This loss often shows a subgenome-bias pattern[19,20]. The correlations between the WGD driven-subgenome-bias pattern and the emergence of evolutionary novelties have been demonstrated[21]. Considering the powerful force of WGD on gene functions, we believe that the genes in the WGD-driven-subgenome may play roles on heart and vascular development. However, whether the genes in the minor ZGA shaped by a biased-subgenome affect the ZGA and have functions on heart and vascular development are largely unknown. Moreover, whether the corresponding

82 human homolog genes in the minor ZGA shaped by a biased-subgenome in lower vertebrate function as
83 anti-cancer molecules by targeting angiogenesis in human is unclear. In this study, we tried to reveal the
84 functions of such genes based on the genome and embryonic transcriptome of a sea fish *Trachinotus*
85 *ovatus*.

86 *T. ovatus* (Linnaeus 1758, also called golden pompano) is a member of the family *Carangidae*
87 (Rafinesque, 1815)[22]. It is widely distributed and cultured in the Asia-Pacific region, and has significant
88 economic importance for offshore cage aquaculture in China and Southeast Asian countries[22]. Here, we
89 report the *de novo* chromosome level-assembly and annotation of golden pompano genome that
90 underwent a subgenomic pattern (least fractionated (LF), medium fractionated (MF), and complete
91 fractionated (Other)). We identified a minor ZGA prior to the major ZGA during its embryonic
92 development, and a novel gene EVM0008813 (designated as New XingHuo, *nxh*), which concurrently
93 belongs to the subgenome MF and the minor ZGA. We studied the function of *nxh* in the zebrafish model
94 by morpholino knockdown of its homolog *zgc:113227* (designated as New XingHuo-like, *nxhl*). For the
95 first time, we reveal that loss of *nxhl* delays the subsequent ZGA activation, causes heart and vascular
96 defects and reduces the angiogenesis via *nxhl*-nucleolin-*ptprb* network in zebrafish. This implies that *nxh*
97 may play a similar function in golden pompano development. Importantly, silence of its human homolog
98 Harbi1 significantly inhibits the migration, invasion and tube formation of HUVEC cells and the
99 proliferation and metastasis of hepatocellular carcinoma cell lines, showing an anti-cancer potential. Our
100 study provides a new insight on cardio-and vascular-linked functions of the genes in the minor ZGA of
101 teleosts during embryonic development and clues for anti-cancer therapeutic strategies.

102 .

103 **RESULTS**

104 **Genome sequencing, assembly and annotation**

105 We first estimated the genome size and GC content of golden pompano at 656.98 Mb and 41.54% based
106 on 21-Kmer (Figure S1).Then we sequenced and assembled the genome of a female individual through a
107 combination of three technologies: paired-end sequencing with the Illumina HiSeq platform,
108 single-molecule real time (SMRT) sequencing with the PacBio Sequel platform, and optical genome
109 mapping with the BioNano Genomics Saphyr System (Figure S2; TableS1, S2A), and an assembly
110 containing 1,490 scaffolds, with a scaffold N50 length of 21.02 Mb was built (Table S1C and S2A). Next,
111 we anchored and oriented the assembly sequences onto 24 pseudo-molecules, which accounted for

112 ~98.60% (636.61 Mb) of genome assembly (Figure S3; Table S1 and S2) according to the interaction
113 frequency mapping of 38.7-fold high-through chromosome conformation capture (Hi-C) data (Table S1d).
114 We finally obtained a genome assembly containing 645.62 Mb of genome data made of 1,536 scaffolds
115 (included 24 pseudochromosomes) with scaffold N50 size of 20.22 Mb (Table S2b). The quality
116 evaluation of the assembled genome turned out to reveal a high level of contiguity and connectivity for
117 the golden pompano genome and facilitated further analyses (Table S2, S3; Figure S3). By combing the
118 evidence of *de novo*, homologs and transcriptome approaches, we annotated 24,186 high confidence
119 protein-coding genes, 152 pseudogenes and 1,274 noncoding RNAs (Table S3, S4 S5 and S6), in
120 comparison to zebrafish (25,642 protein-coding genes). We visualized the genomic landscape of genes,
121 repetitive sequences, genome map markers, Hi-C data, and GC content of the golden pompano genome
122 by circus[23] (Figure S4). As expected, the repetitive elements accumulated in low gene density regions
123 (Figure S4 AB). The optical markers and Hi-C data are evenly distributed across the genome (Figure S4
124 CD), as well as GC content (Figure S4 E). Detailed genome data can be found in the supplementary
125 information (Table S7-S17).
126

127 **Subgenomic patterning suggests karyotypes-retained genes crucial for retaining embryo**
128 **development stability**

129 To examine the evolutionary position of the golden pompano, we reconstructed the evolutionary history
130 of teleost fishes. Phylogenies strongly supported the traditional classification that the golden pompano
131 belongs to *Carangidae*. Golden pompano is a sister of *Seriola dumerilie,* and diverged at approximate
132 59.99 Mya (Figure 1A). Gene family analyses showed that Perciformes and zebrafish share 9,412 families,
133 and the number of species-specific families in Perciformes is apparently lower than in zebrafish (Figure
134 S5). Assuming a constant rate of distribution of silent substitutions (dS)[24] of 1.5e-8, we estimated the
135 dates of Ts3R and Ss4R at 350 Mya and 96 Mya, respectively (Figure 1B). Genome collinearity
136 comparison among spotted gar karyotypes (preserved in golden pompano genome) showed that 894 pairs
137 of paralogous genes that were inherited from the Ts3R event (ohnologues) retained a double-conserved
138 synteny block in the golden pompano genome (Figure 1C, and Table S18), implying that teleost-ancestral
139 karyotypes are considerably conserved in post-Ts3R rediploidization with large fissions, fusions or
140 translocations (Figure 1D, and Table S19-S20). Next, we classified the Ts3R subgenomes according to the
141 integrity of gene as belonging to the LF, MF, and Other subgenome[25]. The component of
142 rediploidization-driven subgenomes (LF, MF, and Other) is unequally distributed among golden pompano

143 subgenomes (Figure 1E), suggesting an asymmetric retention of ancestral subgenomes in teleosts, which
144 is commomly observed in plants[26,27].

145 Although the gene expression evolution following the WGD event has debated in fishes[28,29], knowledge
146 on the rediploidization process and embryo development stability is lacking. We then compared the
147 genome-wide transcriptional levels of LF, MF, and Other karyotypes from whole-embryo development
148 stages (OSP to YAPS) (Figure 2A). Karyotypes-retained regions (LF and MF) showed comparable
149 expression levels during the embryo development, while karyotypes-loss regions (Other) were expressed
150 at significantly lower levels (signed-rank sum test, $P < 0.01$) (Figure S6A). The Ks/Ks values of
151 karyotypes-retained regions are significantly lower than those of karyotypes-loss regions (Figure S6B).
152 This observation indicated that karyotypes-loss genes evolved faster than did the karyotypes-retained
153 regions. We propose that karyotypes-retained genes are crucial for retaining embryo development stability
154 and that karyotypes-loss genes are more prone to contribute to genetic diversity. Detail descriptions about
155 subgenome can be found in supplementary information.

156

157 **Discovery of a minor ZGA gene that concurrently belongs to the karyotypes-retained genes during**
158 **embryo development**

159 To prove this idea, we collected 19 embryo stages of the golden pompano and sequenced the mRNAs to
160 figure out the gene expression patterns (Figure S7, and Table S21-S26). We found that all 57 of the
161 samples were separated into two components according to the expression profile (Figure 2 ABC). The
162 first 33 samples (from OSP to MGS) cluster into a clade and the residual 24 samples (from LGS to YAPS)
163 cluster into another. The genes in the first clade were clearly "silenced" compared with those of the
164 second clade, in which the gene levels show an explosive increase, which is reminiscent of the ZGA
165 process. We then identified the stages after MGS, namely from LGS to YAPS stage, are the major ZGA
166 (Figure 2 ABC, Figure S8 and Figure S7). This ZGA expression profile was verified by our qPCR results
167 (Figure S9). We also noticed that before LGS, a group of genes in three stages, HBS, EGS, and MGS, are
168 highly expressed in the first clade (Figure 2D). This unique gene expression profile might be the minor
169 wave of ZGA before the major wave of ZGA. The existence of minor ZGA was previously documented in
170 mammal embryo development[9,10], and it was suggested to contribute to signaling the massive activation
171 of the genome[30]. Thus, we asked what the functions of the genes in the minor ZGA are during the embryo
172 development of golden pompano.

173 Next, we clustered these genes in the minor ZGA using the WGCNA R package and we found that

174 most of them clustered into the purple_module and are co-expressed in a close network, indicating
175 regulatory roles for these genes before the major ZGA (Figure 2 EF). Then we randomly picked 8 genes
176 to verify their expression trends, and we found all of them to coincide in the expression profile of the
177 minor ZGA genes (Figure S9). We subsequently choose the mostly coincident gene to the minor ZGA
178 expression trend, EVM0008813 (designated as New XingHuo, *nxh*; Figure S10), as a representative gene
179 to figure out the function of minor ZGA during embryo development of golden pompano. We found that
180 *nxh* is one of the karyotypes-retained genes (MF). We then searched the Nr database in NCBI by BLASTp
181 and only one gene *zgc:113227* (designated as New XingHuo-like; *nxhl*) shares 54.7% similarity to *nxh* at
182 the amino acid level in zebrafish. The collinear analysis shows its homolog gene in zebrafish is also *nxhl*,
183 and both have the same functional domain DDE_Tnp_4 as the other seven genes have (Figure S10),
184 suggesting they may have similar biological functions during embryo development. What could the
185 function of *nxh* be?
186

187 **Depletion of *nxhl* delays the ZGA program in zebrafish**

188 Next, we investigated the functions of *nxhl*, the zebrafish homolog of *nxh*, in zebrafish. Since *nxhl* shares
189 high (54.7%) similarity to *nxh* at the amino acid level, we used the morpholino-based knockdown method
190 in zebrafish to study the function of *nxhl,* which may represent the function of *nxh* because they share the
191 DDE_Tnp_4 domain. We used two types of morpholinos (MOs) (splice blocking: *nxhl* ^{e1i1} MO and
192 translational blocking: *nxhl* ^{ATG} MO) to silence *nxhl*. We firstly investigated whether *nxhl* plays roles on
193 the ZGA progress. As the main wave of zygotic gene transcription is activated at 3 hpf [5](Figure 3 A (a)),
194 to determine the effects of *nxhl* depletion on zygotic transcription, we analyzed the transcriptome of
195 untreated embryos and embryos treated with *nxhl*-targeting MOs using RNA-seq. We identified
196 differentially expressed transcripts between wild type and *nxhl*-depleted embryos, and found that *nxhl*
197 depletion most robustly affected transcript levels post ZGA at 6 hpf but not 3 hpf (Figure 3 A (b) ,Table
198 S27).The principal component (PCA) plot in Figure 3 A (c) showed that *nxhl* depletion results in a big
199 difference on zygotic transcription at 6 hpf instead of 3 hpf. We next classified[31] the differentially
200 expressed transcripts into maternal, weakly maternal or zygotic transcripts in *nxhl*-depleted embryos
201 compared with wild type. We found that upon *nxhl* depletion, the differentially expressed maternal
202 transcripts relative to wild-type and zygotic transcripts under-represented at 6 hpf were more abundant
203 than that at 3 hpf, suggesting that ZGA was delayed (Figure 3 A (d), Table S27). We next plotted the
204 expression levels of differentially expressed transcripts identified from shield stage at all time points, and

205 found that transcripts in *nxhl*-depleted embryos significantly increase or decrease at 6 hpf instead of that
206 at 3 hpf with comparison to wild type (Figure 3 B (a and b)). We then detected the expressions of the
207 ZGA marker genes using quantitative PCR, which showed that six ZGA marker genes expressed more
208 robustly at 6 hpf rather than 3 hpf comparing to wild type, confirming the delay of ZGA program (Figure
209 3 B (c)). We also found that at the 6 hpf shield stage, transcripts reduced upon *nxhl* depletion are involved
210 in morphogenesis development, Wnt signalling, and oxidoreductase activity, while transcripts elevated
211 participate in protein binding and mitochondrial assembly (Figure 3 C, Table S28). Overall, our data
212 suggest that *nxhl* is essential to promotion of ZGA program and depletion of *nxhl* delays the ZGA in
213 zebrafish. This finding is consistent with previous research in the mouse[10] and highlights the important
214 function of *nxh* and *nxhl* during embryo development.
215
216 **Knockdown of *nxhl* causes morphological defects in the heart and caudal fin in zebrafish**
217 Next, we investigated whether loss of *nxhl* affects morphology development the zebrafish. We knocked
218 down *nxhl* as depiction in Figure 4 A. Heart and vascular development starts at 16 hpf[3] and 24 hpf[4] in
219 zebrafish, respectively. This means heart and vascular development occur after the major ZGA (3 hpf).
220 Since the loss of *nxhl* significantly suppresses the major ZGA in zebrafish, we speculated that *nxhl* may
221 affect the heart and vascular development. Then we checked the changes in the heart and vascular system
222 after the silencing of *nxhl.* We observed that injection of 4 ng of morpholino *nxhl* ^{e1i1} MO and *nxhl* ^{ATG}
223 MO both resulted in nearly identical phenotypes of pericardial oedema, body axis bending, and caudal fin
224 defects (Figure 4 B (a-k), Figure S11 A, and Figure S12) at 3 dpf (days post fertilization), suggesting that
225 the phenotype of *nxhl* knockdown is *nxhl* -specific (Figure 4 B). Considering the vascular system,
226 embryos injected with *nxhl* ^{e1i1} MO present thinner ISVs (yellow arrows) and ectopic sprouts (asterisk) of
227 dorsal aorta compared with controls, and the *nxhl* knockdown prevents the PAV formation, the precursor
228 to the lymphatic system. Moreover, heartbeat and circulation in the caudal vein (CV) is visible in the
229 control fish, but is abnormal in *nxhl*-MO-injected fish (Supplementary Movie1, Supplementary Movie2).
230 The efficiency of *nxhl* ^{e1i1} MO and *nxhl* ^{ATG} MO at 6 hpf and 3 dpf demonstrated that injection of 4 ng of
231 *nxhl* ^{e1i1} and *nxhl* ^{ATG} morpholino dramatically disrupted normal splicing of *nxhl* (Figure 4B (j)) ,
232 indicating high efficiency and specificity of the morpholino knockdown of *nxhl*. We also recorded a high
233 percentage of embryos with defects (81.55%, n=103 embryos in *nxhl* ^{e1i1} MO and 100%, n=106 embryos
234 in *nxhl* ^{ATG} MO) with comparison to control (n=128 embryos) and low survival rate at 3 dpf (45.78%,
235 n=225 embryos in *nxhl* ^{e1i1} MO and 17.68%, n=198 embryos in *nxhl* ^{ATG} MO) compared to controls

236 (n=218 embryos) (Figure 4B (k) and Figure S11B). This confirmed that knockdown of *nxhl* certainly
237 causes morphological defects in the heart and caudal fin in zebrafish.
238
239 **Loss of *nxhl* causes vascular defects and impairs formation of the caudal vein plexus (CVP) in**
240 **zebrafish**
241 Since *nxhl* silence caused pericardial edema and caudal fin defects, we speculated that there is a positive
242 connection between the vascular system and these phenotypes. We then used the *Tg(fli1a:EGFP)*^{y1}
243 zebrafish as a model to find out if the vascular development is impaired by knockdown of *nxhl.* Embryos
244 were injected with 4 ng control MO or *nxhl* ^{e1i1} MO, and the results exhibited that loss of *nxhl* caused
245 intersegmental vessel (ISV) growth defect and disruption of the honeycomb structure in the CVP at 52
246 hpf (Figure 4C (a-f)). Also, a thinner ISV growth and ectopic sprouts of dorsal aorta were observed at the
247 rear-somite with only 10% of complete ISVs (n=365 embryos) in *nxhl* ^{e1i1} morphants compared to 98% of
248 complete ISVs in controls (n=335 embryos). In control embryos, the parachordal vessels (PAV) were
249 formed normally, while the *nxhl* ^{e1i1} MO knockdown prevented the parachordal vessels (PAV) formation
250 (Figure 4C (g)). We also observed that CVPs formed honeycomb-like structures at the tail around 52 hpf
251 in control embryos, but *nxhl* ^{e1i1} MO knockdown caused specific defects in CVP formation (Figure 4D
252 (a-f)). Quantification of loop formation and the area at CVP showed a 8.2-fold and 3-fold decrease in *nxhl*
253 ^{e1i1} MO morphants (n=10 embryos) at 52 hpf, respectively (Figure 4D (g and h)). Our data indicate that
254 *nxhl* plays a critical role in controlling vascular integrity and in regulating ISV and CVP formation during
255 angiogenesis, which is an explanation strongly consistent with the phenotypes observed. What is the
256 mechanism behind?
257
258 ***Nxhl* may play role in angiogenesis via endothelial cells and its human homolog Harbi1 is a**
259 **potential anti-cancer therapeutic target**
260 Endothelial cells are the building elements of blood vessels. As *nxhl* plays a critical role in controlling
261 vascular integrity, we then speculated that it may affect angiogenesis via endothelial cells. This is
262 supported by the significant enrichment of the genes involved in blood vessel morphogenesis when *nxhl*
263 was knocked down (Table S29). Since human Harbi1 gene shares DDE_Tnp_4 domain with *nxhl* and *nxh*
264 (Figure S10), it is supposed to play similar role to *nxhl* and *nxh* in vascular epithelioid cells. Next, we
265 designed siRNAs targeting human Harbi1, transferred siRNA into human umbilical vein endothelial
266 (HUVEC) cells and investigated their cell migration, invasion, and tube formation. We observed that

267 silencing of Harbi1 not only significantly inhibits the cell migration and invasion abilities but also the
268 tube formation compared with controls in HUVECs (Figure 5A; Figure S13A). This highlights the
269 anti-angiogenesis function of Harbi1, indicating that *nxh* and *nxhl* are functionally related. Because
270 targeting angiogenesis has great potential in anti-tumor or anti-cancer therapy[32,33], we next asked whether
271 this anti-angiogenesis function of Harbi1 affects the angiogenesis of the cancer cells. We choose
272 SMMC7721 (hepatocellular carcinoma, HCC), A549 (non-small cell lung cancer, NSCLC), and HCT116
273 (colon carcinoma, Colo) cancer cell lines, transferred with si-Harbi1, then their supernatants incubated
274 with HUVECs, respectively. We observed the abilities of tube formation in all of the cancer cells are
275 significantly inhibited compared with controls (Figure 5B; Figure S13B; Figure S13C). This shows that
276 silencing of Harbi1 does suppress the angiogenesis of the cancer cells. We further silenced the Harbi1 in
277 SMMC7721 cell lines, and found that loss of Harbi1 not only significantly inhibits the proliferation,
278 migration and colony formation of HCC cells, but suppresses the expressions of EMT marker genes
279 (Figure 5C; Figure S13D). We then predicted whether Harbi1 links to the HCC progression based on the
280 data in the cancer genome atlas (TCGA). We found that expressions of Harbi1 in HCC tissues are
281 significantly higher than normal tissues (Figure S13E). We next tested this in 72 HCC patients who had
282 been carried out the surgical operation using qRT-PCR. We found that the clinical test results are
283 consistent with TCGA predictions (Figure 5D; Figure S13E). Such function of Harbi1 was never reported
284 before on cancer[34-36]. Our data indicate that *nxhl* and *nxh*, like their human homolog Harbi1, play role in
285 angiogenesis process via endothelial cells and Harbi1 is a potential anti-cancer therapeutic target.
286

287 **Loss of *nxhl* leads to the down-regulation of the landmark gene network associated with heart and**
288 **vascular development in zebrafish**

289 Next, we asked how *nxhl* affects angiogenesis. We first examined transcriptome sequencing (RNA-seq)
290 data from zebrafish after injection of 4 ng *nxhl* ^{e1i1} MO at 3 dpf. We found that loss of *nxhl* greatly
291 changes the transcriptome with 1955 down-regulated and 698 up-regulated (Figure S14; Table S21). We
292 noticed that in the KEGG pathways associated with heart and vascular development are significantly
293 enriched in the *nxhl*-silenced group (Figure S15; Table S22-26). We speculated that the transcription of
294 genes linked to heart and vascular development may also be significantly changed in the *nxhl*-silenced
295 zebrafish. We then picked and examined the expression of 18 genes previously documented to be closely
296 related to heart defects and/ or angiogenesis[37-41] using qPCR. Consistent with the RNA-seq data, we
297 found that 13 of these genes (*ptprb, tie2, nr2f1a, s1pr1, hey2, dot1L, hand2, erbb2, klf2a, mef2cb, mef2aa,*

298 *ephB2a* and *cx40.*8) were significantly decreased while two genes (*vegfaa and vegfr2*) increased sharply.
299 *S1pr2, egfl7, and nrg2a* were kept unchanged (Figure 6A-D). Notably, the arterial marker *ephB2a* and
300 venous marker *erbb2* were decreased in *nxhl* morphants compared to the wild-type (Figure 6D). Normally,
301 the increase of v*egfaa and vegfr2* is linked to the enhancement of vascular system[42,43]. However, in our
302 study, both genes increased while others decreased when *nxhl* was silenced. We speculated this was a
303 consequence of a negative feedback regulation to avoid an excessive decrease in the vascular system. We
304 next found that at the protein level, when *nxhl* is silenced, the *nxhl* protein also decreases. *ptprb*, the most
305 decreased gene, is also greatly reduced at the protein level. The s1pr1, hand2, dot1L, and hey2 proteins
306 were also downregulated compared with controls (Figure 6F). As previously reported, *ptprb, tie2, nr2f1a,*
307 *s1pr1, vegfaa* and *vegfr2* normally contribute to vascular development and deletion of each of them leads
308 to defects on the vascular system during embryo development[42-46], while loss of *dot1L, hand2, erbb2,*
309 *mef2cb, mef2aa, ephB2a* or *cx40.8* always results in cardiovascular system or heart development
310 defects[39,47-53]. *Hey2 and klf2a* have been implicated in the regulation of both angiogenesis and heart
311 development[54,55]. Based on these reports, we built a schematic diagram of the network as shown in Figure
312 6F. This network demonstrated that silencing of *nxhl* did downregulate the key genes that are essential for
313 heart and /or vascular development. To this end, our results showed that loss of *nxhl* greatly affects the
314 expression of these key genes in the network, suggesting that the heart and vascular phenotypes caused by
315 *nxhl* silencing are greatly due to the regulation of these genes. Thus, we then asked how *nxhl* mediates the
316 network.
317

318 **_Ptprb_ plays similar function to _nxhl_ in zebrafish and is downstream of _nxhl_**

319 As described above, we noticed that *ptprb* is the most downregulated gene after silence of *nxhl* and the
320 one that is closely linked to both vascular integrity and angiogenesis[56-61]. To test whether there is a
321 positive connection between *nxhl* and *ptprb,* we knocked down *ptprb* by injection of 4 ng *ptprb* MO
322 morphants designed (see Table 30). We observed that embryos injected with *ptprb* MO present slight
323 pericardial edema, shortened body axis and severe body axis bending (Figure 7A (a-i)； Figure S16).
324 Moreover, heartbeat and circulation in the caudal vein (CV) is visible in the control fish (Supplementary
325 Movie 3,4), but is abnormal in *ptprb*-MO-injected fish (Supplementary Movie 5-8). We also recorded a
326 high percentage of embryos with defects (75.48%, n=208 embryos in *ptprb* MO and 0.94%, n=212
327 embryos in control) and lower survival rate at 50 hpf compared with controls (Figure 7A (j)). The
328 efficiency of *ptprb* MO at 50 hpf demonstrated that injection of 4 ng of *ptprb* morpholino was highly

329 efficient in producing a knockdown of *ptprb* (Figure 7A(k, l)). In the vascular system, loss of *ptprb* leads
330 to an indefinite absence or deformity of DLAVs (blue arrowhead) and ISVs in the tail end (white and
331 yellow arrowhead), and a decrease of PAV (red arrowhead) formation (Figure 7B (a-i)). Knockdown of
332 *ptprb* caused significantly decrease of the mean diametre of ISVs in *ptprb*-MO injected embryos (Figure
333 7B (j, k)). We also observed that CVPs formed honeycomb-like structures at the tail around 50 hpf in
334 control embryos, but *ptprb* MO knockdown causes CVP sinus cavities defects (Figure 7C (a- c)).
335 Quantification of loop formation and area at CVP showed a 5.6-fold and 2.2-fold decrease in *ptprb* MO
336 (n=83 embryos) at 50 hpf, respectively (Figure 7 C (d, e)). These data are to some extent consistent with
337 previous reports[41] and strongly suggest that loss of *ptprb* results in heart and vascular phenotypes similar
338 to those of *nxhl.* Moreover, we also investigated the expression of the 15 genes that were examined in
339 *nxhl*-knockdown experiment. We found that most of the genes present an expression profile similar to that
340 in *nxhl*-knockdown experiment, except vegfaa and vegfr2 (Figure 7D (a, b)). We then logically concluded
341 that *nxhl* and *ptprb* should act in the same signaling pathway. However, which one is downstream of the
342 other is unclear. Therefore, we examined the gene expression of *nxhl* after silencing of *ptprb*, and we
343 found that it was kept unchanged (Figure 7A (m)), but we observed a significant decrease of *ptprb*
344 expression after silencing of *nxhl.* These observations demonstrate that *ptprb* acts at the downstream of
345 *nxhl.*
346

347 **_Nxhl-NCL-ptprb_ signaling links heart and vascular molecular network to regulate _nxhl_-induced**
348 **phenotypes**

349 As mentioned above, among the 18 genes associated with heart and vascular development, 15 genes were
350 significantly changed by the *nxhl* and *ptprb* knockdown. We suppose these genes may be part of a
351 regulatory network of their own. We then built a schematic diagram of the network according to previous
352 reports (Figure 6E). This network presents connections between most of these genes, suggesting a
353 cooperative regulation mechanism on the heart and vascular development. As we already knew that *ptprb*
354 acts downstream of *nxhl*, we next asked if *nxhl* directly interacts with *ptprb* to mediate the network. Thus,
355 we designed *nxhl* probes and conducted a ChIRP-MS experiment in zebrafish to find out those proteins
356 interacting with *nxhl*. Eleven proteins with change folds above 2 were discovered (Figure 8 A (a, b)). This
357 indicates that the *nxhl* RNA may interact with these proteins. Unexpectedly, *ptprb* was not found in these
358 interacting proteins (Figure 8A (a, b)). This suggests that proteins other than *ptprb* may be involved. We
359 next focused on the proteins of the group of eleven proteins that are associated with the vascular system,

360 and nucleolin (NCL) (Figure 8A (c, d)) aroused our interest because of its molecular conservation and
361 important functions on angiogenesis[62,63]. To figure out whether *nxhl* interacts with NCL, we performed
362 RNA Immunoprecipitation (RIP) using the NCL protein as bait protein in 293T cells (Figure S17) and
363 then detected the *nxhl* RNA using qPCR. We found that *nxhl* RNA is significantly higher than that in IgG
364 control in the RNAs pulled-down by the NCL protein (Figure 8B (b, c, d), and Figure S18, 19). The RNA
365 pulled down was amplified and the sequencing results confirmed that it is *nxhl* gene. This indicates that
366 the NCL protein reversely interacts with *nxhl* RNA. Therefore, these experiments prove that *nxhl* RNA
367 and NCL interact physically.

368 However, still no evidence was shown on the relationship between *nxhl* and *ptprb*. Could it be that
369 NCL interacts with *ptprb,* thus bridging *nxhl* and *ptprb?* Such scenario was never proposed or
370 documented before. However, a report showed that VEGF, which plays roles on vascular development as
371 *nxhl* does, does interact with NCL[64]. We then supposed that NCL might also interact with *ptprb* or its
372 human homologue VE-PTP, the key molecule in vascular development. To test this hypothesis, we
373 detected VE-PTP RNA using the same RNAs pulled-down by NCL protein, and we found that VE-PTP
374 RNA is significantly higher than that in IgG control. The RNA pulled down was amplified and the
375 sequencing results proved that it is VE-PTP gene. This indicates that the NCL protein can also interact
376 with VE-PTP RNA physically (Figure 8B (b, e, f)). We next verified this interaction in 293T cells using
377 the VE-PTP RNA pulldown experiment in the reverse way, and the result of western blotting against NCL
378 protein supports the interaction between VE-PTP and NCL (Figure 8C (a)). However, whether this
379 interaction also occurs between NCL and *ptprb* in zebrafish is unclear. We next designed a zebrafish *ptprb*
380 gene-specific probe to pull down the proteins that interact with *ptprb* in the juvenile zebrafish. We found
381 that the NCL protein band is much stronger that in controls (Figure 8C (b)). These results indicate that the
382 NCL protein not only interacts with VEP-PTP in 293T cells but also with *ptprb* in zebrafish. So far, we
383 proved that *nxhl* and NCL, NCL and *ptprb* interact physically. However, how *nxhl* regulates NCL and
384 *ptprb* is unclear. To address this issue, we micro-injected 4 ng *nxhl*-e1i1-MO in one cell stage embryo,
385 and found that resemble phenotypes were induced as that shown in Figure 4 (Figure S20). Meanwhile, we
386 found that loss of *nxhl* not only causes a significant decrease of NCL mRNA and total protein level but
387 also leads to reduction of T76 phosphorylation level and an increase of the K88 acetylation level of the
388 NCL protein (Figure 8 D (a, b)). This suggests that knockdown of *nxhl* significantly affects the expression
389 of NCL, which plays vital functions in angiogenesis[63], although the impact of phosphorylation and
390 acetylation of NCL protein on the heart and vascular development have not been deeply understood yet[62].

391 Then we investigated the expression of the downstream gene *ptprb*, and found that loss of *nxhl* also
392 decreases *ptprb* at both mRNA and protein levels (Figure 8D (c, d)). These results suggest that silence of
393 *nxhl* leads to heart and vascular defect due to the downregulation of both NCL and *ptprb* via the decrease
394 in the *nxhl*-NCL and NCL-*ptprb* interaction.

395 Taken together, we conclude that the NCL protein can not only interact with the upstream *nxhl* but also
396 the downstream *ptprb.* This uncovers the existence of regulatory gene upstream of *ptprb* gene, one of the
397 most important vascular-controllers. Certainly, this logically proves that *nxh* may act in the same way in
398 the golden pompano. Based on these data, we built a new schematic diagram of the network that shows
399 the *nxhl*-NCL-*ptprb* signaling links of the heart and vascular system molecular network (Figure 8E (a) vs.
400 Figure 6E). We also made a schematic diagram to describe the possible mechanism underlying
401 *nxhl*-induced phenotypes (Figure S21), highlighting the finding that loss of *nxhl* significantly and broadly
402 downregulates heart- and vascular- associated landmark gene network via the *nxhl* -NCL-*ptprb* pathway.
403 Given the extreme importance of the heart and vascular development, and the broad connections with
404 landmark genes described in the network, we believe the finding of this novel signaling pathway to be of
405 considerable relevance for the study of the heart and vascular development regulatory network.

406

407 **Discussion**

408 Taken together, we report the high-quality chromosome-level genome for the golden pompano, which
409 underwent the Ts3R WGD and shows biased-subgenomes retention. For the first time, we show the
410 existence of minor and major ZGA during golden pompano embryo development. We also, for the first
411 time, report that the karyotypes-retained gene *nxh* belongs to the minor ZGA gene, and its zebrafish
412 homolog *nxhl* depletion significantly delays the ZGA program in zebrafish. The *nxhl* plays key roles on
413 the heart and vascular development through regulating a heart- and vascular-associated network of
414 landmark genes. Moreover, we first highlight the interactions between *nxhl* and NCL, NCL and *ptprb*,
415 and the existence of a hitherto unreported *nxhl*-NCL-*ptprb* signaling pathway as a regulatory complex
416 mediating both heart and vascular development. Finally, we show that the novel angiogenesis and
417 anti-cancer function of Harbi1, the human homolog of *nxh*, and it might be a potential anti-cancer
418 therapeutic target. A detailed discussion can be found in the supplementary information.

419

420 **Acknowledgements**

421 We thank Biomarker Technologies Corporation for genome and transcriptome sequencing. This research

was supported by the Guangxi science and technology major project (GuiKeAA18242031, GuiKeAA18242031-2, GuiKeAA17204080, GuiKe AA17204080-3), Guangxi Key Laboratory for Aquatic Genetic Breeding and Healthy Aquaculture, Guangxi Institute of Fishery Sciences (14-045-10 (14-A-01-02),15-140-23(15-A-01-01, 15-A-01-02, 15-A-01-03),16-380-38(16-A-01-01, 16-A-01-02), 17-259-29 and 19-A-01-05) and Guangxi research institutes of basic research and public service special operations (GXIF-2016-03, GXIF-2016-09, GXIF-2016-18, GXIF-2016-19）.

Author Contributions

H.L.L and X.H.C designed the scientific objectives and oversaw the project. H.L.L., X.H.C., Y.Z.Z., J.X.P., and Y.L discussed the primary ideas of the article. Y.D.Z, J.X.P., Y.L., Y.H., Q.Y.L., P.P.H., C.L.Y., P.Y.W., X.L.C., and P.F.F. collected samples for sequencing DNA and RNA. C.M.J. and their colleagues performed genome sequencing, assembly and annotation. C.M.J. and H.Y.Y performed phylogenomic and whole genome duplication evolution analysis. C.M.J. H.Y.Y., H.L.L., and Y.D.Z performed RNA-seq analysis. H.L.L and L.Q.L performed functional assay of zebrafish *nxhl* gene and Harbi1 gene. C.M.J, H.Y.Y., H.L.L., and Y.D.Z prepared the supplemental data and method. C.M.J. and H.L.L prepared the draft manuscript with input from all other authors. H.L.L., X.H.C., Y.L., and H.K.Z. discussed and revised the manuscript.

Data availability

The authors declare that all data reported in this study are fully and freely available from the date of publication. This Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the accession WOFJ00000000. The version described in this paper is version WOFJ01000000. The draft genome data (genome assembling and annotations) and RNA-Seq data of the embryo are available under BioProject PRJNA574895. Transcriptome (Illumina) data of *nxhl* silence are available in the Sequence Read Archive (SRA) with accession number SRR10199007 and SRR10199008 under BioProject PRJNA573544 .

Declaration of interests

The authors declare no competing interests.

Supplemental information

Supplemental Information can be found with this article online version.

References

1 Li, R., Baek, K. I., Chang, C. C., Zhou, B. & Hsiai, T. K. Mechanosensitive Pathways Involved in Cardiovascular Development and Homeostasis in Zebrafish. *J Vasc Res* **56**, 273-283, doi:10.1159/000501883 (2019).

2 Garoffolo, G. & Pesce, M. Mechanotransduction in the Cardiovascular System: From Developmental Origins to Homeostasis and Pathology. *Cells* **8**, doi:10.3390/cells8121607 (2019).

3 Crosier, P. S. *et al.* Pathways in blood and vessel development revealed through zebrafish genetics. *Int J Dev Biol* **46**, 493-502 (2002).

4 Glickman, N. S. & Yelon, D. Cardiac development in zebrafish: coordination of form and function. *Semin Cell Dev Biol* **13**, 507-513 (2002).

5 Palfy, M., Joseph, S. R. & Vastenhouw, N. L. The timing of zygotic genome activation. *Current opinion in genetics & development* **43**, 53-60, doi:10.1016/j.gde.2016.12.001 (2017).

6 Blythe, S. A. & Wieschaus, E. F. Coordinating Cell Cycle Remodeling with Transcriptional Activation at the Drosophila MBT. *Curr Top Dev Biol* **113**, 113-148, doi:10.1016/bs.ctdb.2015.06.002 (2015).

7 Harrison, M. M. & Eisen, M. B. Transcriptional Activation of the Zygotic Genome in Drosophila. *Curr Top Dev Biol* **113**, 85-112, doi:10.1016/bs.ctdb.2015.07.028 (2015).

8 Sandler, J. E. & Stathopoulos, A. Quantitative Single-Embryo Profile of Drosophila Genome Activation and the Dorsal-Ventral Patterning Network. *Genetics* **202**, 1575-1584, doi:10.1534/genetics.116.186783 (2016).

9 Heyn, P. *et al.* The earliest transcribed zygotic genes are short, newly evolved, and different across species. *Cell Rep* **6**, 285-292, doi:10.1016/j.celrep.2013.12.030 (2014).

10 Abe, K. I. *et al.* Minor zygotic gene activation is essential for mouse preimplantation development. *Proceedings of the National Academy of Sciences of the United States of America* **115**, E6780-E6788, doi:10.1073/pnas.1804309115 (2018).

11 Li, X. *et al.* The Tumor Vessel Targeting Strategy: A Double-Edged Sword in Tumor Metastasis. *Cells* **8**, doi:10.3390/cells8121602 (2019).

12 Viallard, C. & Larrivee, B. Tumor angiogenesis and vascular normalization: alternative therapeutic targets. *Angiogenesis* **20**, 409-426, doi:10.1007/s10456-017-9562-9 (2017).

13 Sung, Y. C. *et al.* Delivery of nitric oxide with a nanocarrier promotes tumour vessel normalization and potentiates anti-cancer therapies. *Nat Nanotechnol* **14**, 1160-1169, doi:10.1038/s41565-019-0570-3 (2019).

14 Smith, J. J. *et al.* Sequencing of the sea lamprey (Petromyzon marinus) genome provides insights into vertebrate evolution. *Nature genetics* **45**, 415-421, 421e411-412, doi:10.1038/ng.2568 (2013).

15 Nakatani, Y., Takeda, H., Kohara, Y. & Morishita, S. Reconstruction of the vertebrate ancestral genome reveals dynamic genome reorganization in early vertebrates. *Genome research* **17**, 1254-1265, doi:10.1101/gr.6316407 (2007).

16 Meyer, A. & Schartl, M. Gene and genome duplications in vertebrates: the one-to-four (-to-eight in fish) rule and the evolution of novel gene functions. *Curr Opin Cell Biol* **11**, 699-704 (1999).

17 Arnegard, M. E., Zwickl, D. J., Lu, Y. & Zakon, H. H. Old gene duplication facilitates origin and diversification of an innovative communication system--twice. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 22172-22177, doi:10.1073/pnas.1011803107 (2010).

18 Braasch, I., Brunet, F., Volff, J. N. & Schartl, M. Pigmentation pathway evolution after whole-genome duplication in fish. *Genome Biol Evol* **1**, 479-493, doi:10.1093/gbe/evp050 (2009).

19 Cheng, F. *et al.* Epigenetic regulation of subgenome dominance following whole genome triplication in Brassica rapa.

New Phytol **211**, 288-299, doi:10.1111/nph.13884 (2016).

20 Hughes, T. E., Langdale, J. A. & Kelly, S. The impact of widespread regulatory neofunctionalization on homeolog gene evolution following whole-genome duplication in maize. *Genome research* **24**, 1348-1355, doi:10.1101/gr.172684.114 (2014).

21 Mammoto, T., Mammoto, A. & Ingber, D. E. Mechanobiology and developmental control. *Annu Rev Cell Dev Biol* **29**, 27-61, doi:10.1146/annurev-cellbio-101512-122340 (2013).

22 Nelson, J. S. Fishes of the World.Third edition [M]. *New York: John Wiley&Sons, Inc.* (1994).

23 Krzywinski, M. *et al.* Circos: an information aesthetic for comparative genomics. *Genome research* **19**, 1639-1645, doi:10.1101/gr.092759.109 (2009).

24 Crete-Lafreniere, A., Weir, L. K. & Bernatchez, L. Framing the Salmonidae family phylogenetic portrait: a more complete picture from increased taxon sampling. *PloS one* **7**, e46662, doi:10.1371/journal.pone.0046662 (2012).

25 Cheng, F. *et al.* Deciphering the diploid ancestral genome of the Mesohexaploid Brassica rapa. *The Plant cell* **25**, 1541-1554, doi:10.1105/tpc.113.110486 (2013).

26 Emery, M. *et al.* Preferential retention of genes from one parental genome after polyploidy illustrates the nature and scope of the genomic conflicts induced by hybridization. *PLoS genetics* **14**, e1007267, doi:10.1371/journal.pgen.1007267 (2018).

27 Liang, Z. & Schnable, J. C. Functional Divergence between Subgenomes and Gene Pairs after Whole Genome Duplications. *Molecular plant* **11**, 388-397, doi:10.1016/j.molp.2017.12.010 (2018).

28 Braasch, I. *et al.* The spotted gar genome illuminates vertebrate evolution and facilitates human-teleost comparisons. *Nature genetics* **48**, 427-437, doi:10.1038/ng.3526 (2016).

29 Lien, S. *et al.* The Atlantic salmon genome provides insights into rediploidization. *Nature* **533**, 200-205, doi:10.1038/nature17164 (2016).

30 Lee, M. T., Bonneau, A. R. & Giraldez, A. J. Zygotic genome activation during the maternal-to-zygotic transition. *Annu Rev Cell Dev Biol* **30**, 581-613, doi:10.1146/annurev-cellbio-100913-013027 (2014).

31 Lee, M. T. *et al.* Nanog, Pou5f1 and SoxB1 activate zygotic gene expression during the maternal-to-zygotic transition. *Nature* **503**, 360-364, doi:10.1038/nature12632 (2013).

32 Castet, F. *et al.* Uveal Melanoma, Angiogenesis and Immunotherapy, Is There Any Hope? *Cancers* **11**, doi:10.3390/cancers11060834 (2019).

33 Qin, S. *et al.* Recent advances on anti-angiogenesis receptor tyrosine kinase inhibitors in cancer therapy. *Journal of hematology & oncology* **12**, 27, doi:10.1186/s13045-019-0718-5 (2019).

34 Smith, J. J., Sumiyama, K. & Amemiya, C. T. A living fossil in the genome of a living fossil: Harbinger transposons in the coelacanth genome. *Molecular biology and evolution* **29**, 985-993, doi:10.1093/molbev/msr267 (2012).

35 Sinzelle, L. *et al.* Transposition of a reconstructed Harbinger element in human cells and functional homology with two transposon-derived cellular genes. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 4715-4720, doi:10.1073/pnas.0707746105 (2008).

36 Kapitonov, V. V. & Jurka, J. Harbinger transposons and an ancient HARBI1 gene derived from a transposase. *DNA and cell biology* **23**, 311-324, doi:10.1089/104454904323090949 (2004).

37 Pereira, C. F. *et al.* Hematopoietic Reprogramming In Vitro Informs In Vivo Identification of Hemogenic Precursors to Definitive Hematopoietic Stem Cells. *Developmental cell* **36**, 525-539, doi:10.1016/j.devcel.2016.02.011 (2016).

38 Boselli, F., Freund, J. B. & Vermot, J. Blood flow mechanics in cardiovascular development. *Cellular and molecular life sciences : CMLS* **72**, 2545-2559, doi:10.1007/s00018-015-1885-3 (2015).

39 Hinits, Y. *et al.* Zebrafish Mef2ca and Mef2cb are essential for both first and second heart field cardiomyocyte differentiation. *Developmental biology* **369**, 199-210, doi:10.1016/j.ydbio.2012.06.019 (2012).

40 Ruiz-Herguido, C. *et al.* Hematopoietic stem cell development requires transient Wnt/beta-catenin activity. *The Journal of experimental medicine* **209**, 1457-1468, doi:10.1084/jem.20120225 (2012).

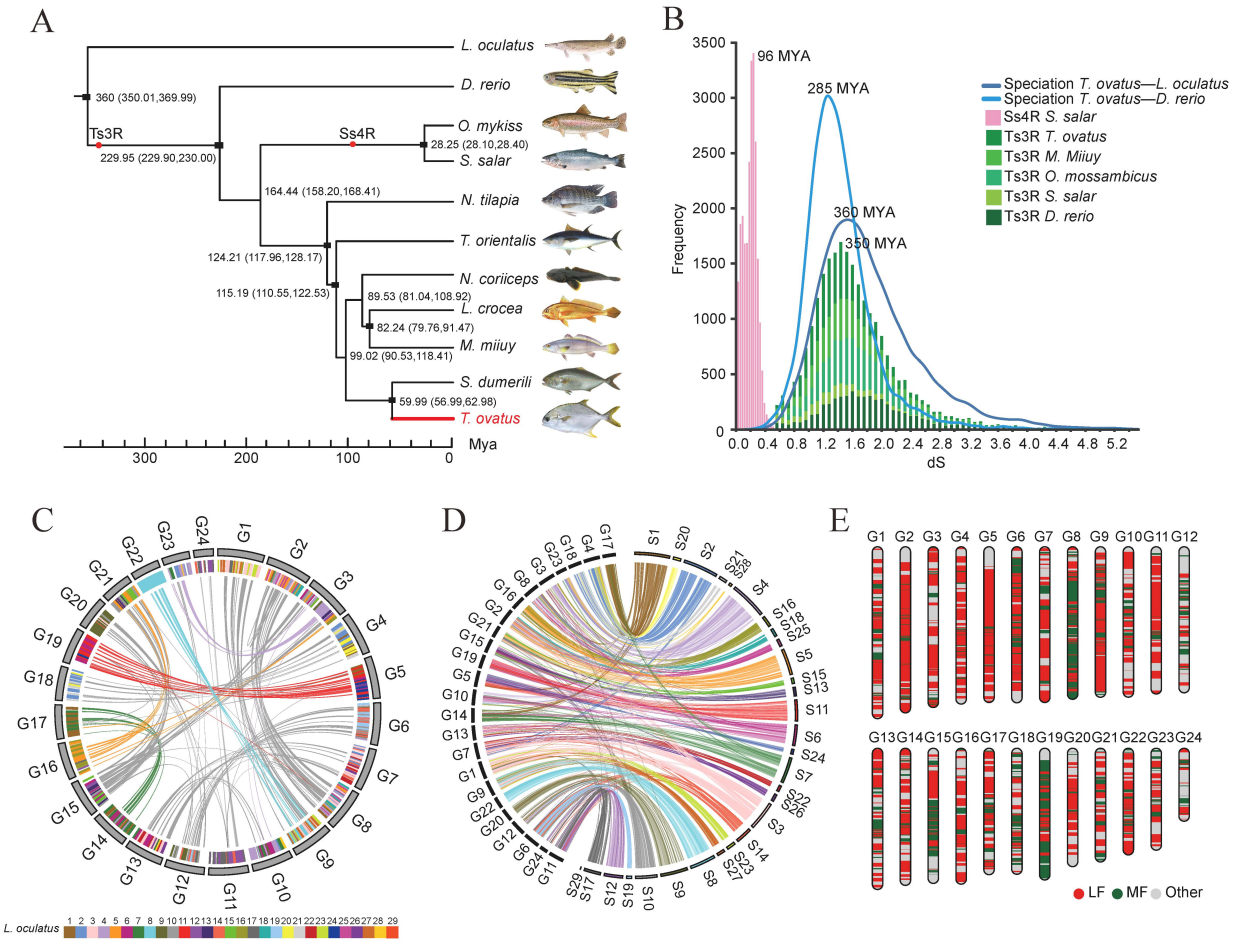
41 Carra, S. *et al.* Ve-ptp modulates vascular integrity by promoting adherens junction maturation. *PloS one* **7**, e51245,

543 doi:10.1371/journal.pone.0051245 (2012).
544 42 Toselli, C. M., Wilkinson, B. M., Paterson, J. & Kieffer, T. J. Vegfa/vegfr2 signaling is necessary for zebrafish islet
545 vessel development, but is dispensable for beta-cell and alpha-cell formation. *Scientific reports* **9**, 3594,
546 doi:10.1038/s41598-019-40136-1 (2019).
547 43 Koenig, A. L. *et al.* Vegfa signaling promotes zebrafish intestinal vasculature development through endothelial cell
548 migration from the posterior cardinal vein. *Developmental biology* **411**, 115-127, doi:10.1016/j.ydbio.2016.01.002
549 (2016).
550 44 Winderlich, M. *et al.* VE-PTP controls blood vessel development by balancing Tie-2 activity. *The Journal of cell*
551 *biology* **185**, 657-671, doi:10.1083/jcb.200811159 (2009).
552 45 Wu, B. J. *et al.* Nuclear receptor subfamily 2 group F member 1a (nr2f1a) is required for vascular development in
553 zebrafish. *PloS one* **9**, e105939, doi:10.1371/journal.pone.0105939 (2014).
554 46 Mendelson, K., Zygmunt, T., Torres-Vazquez, J., Evans, T. & Hla, T. Sphingosine 1-phosphate receptor signaling
555 regulates proper embryonic vascular patterning. *The Journal of biological chemistry* **288**, 2143-2156,
556 doi:10.1074/jbc.M112.427344 (2013).
557 47 Schindler, Y. L. *et al.* Hand2 elevates cardiomyocyte production during zebrafish heart development and regeneration.
558 *Development* **141**, 3112-3122, doi:10.1242/dev.106336 (2014).
559 48 Nguyen, A. T. *et al.* DOT1L regulates dystrophin expression and is critical for cardiac function. *Genes & development*
560 **25**, 263-274, doi:10.1101/gad.2018511 (2011).
561 49 Medrano, J. L. & Naya, F. J. The transcription factor MEF2A fine-tunes gene expression in the atrial and ventricular
562 chambers of the adult heart. *The Journal of biological chemistry* **292**, 20975-20988, doi:10.1074/jbc.M117.806422
563 (2017).
564 50 Lazic, S. & Scott, I. C. Mef2cb regulates late myocardial cell addition from a second heart field-like population of
565 progenitors in zebrafish. *Developmental biology* **354**, 123-133, doi:10.1016/j.ydbio.2011.03.028 (2011).
566 51 Liu, J. *et al.* A dual role for ErbB2 signaling in cardiac trabeculation. *Development* **137**, 3867-3875,
567 doi:10.1242/dev.053736 (2010).
568 52 Gerhart, S. V., Jefferis, R. & Iovine, M. K. Cx40.8, a Cx43-like protein, forms gap junction channels inefficiently and
569 may require Cx43 for its association at the plasma membrane. *FEBS letters* **583**, 3419-3424,
570 doi:10.1016/j.febslet.2009.09.054 (2009).
571 53 Gemel, J. *et al.* Connexin40 abnormalities and atrial fibrillation in the human heart. *Journal of molecular and cellular*
572 *cardiology* **76**, 159-168, doi:10.1016/j.yjmcc.2014.08.021 (2014).
573 54 Gibb, N. *et al.* Hey2 regulates the size of the cardiac progenitor pool during vertebrate heart development.
574 *Development* **145**, doi:10.1242/dev.167510 (2018).
575 55 Vermot, J. *et al.* Reversing blood flows act through klf2a to ensure normal valvulogenesis in the developing heart.
576 *PLoS biology* **7**, e1000246, doi:10.1371/journal.pbio.1000246 (2009).
577 56 Hayashi, M. *et al.* VE-PTP regulates VEGFR2 activity in stalk cells to establish endothelial cell polarity and lumen
578 formation. *Nature communications* **4**, 1672, doi:10.1038/ncomms2683 (2013).
579 57 Braun, L. J. *et al.* VE-PTP inhibition stabilizes endothelial junctions by activating FGD5. *EMBO reports*,
580 doi:10.15252/embr.201847046 (2019).
581 58 Juettner, V. V. *et al.* VE-PTP stabilizes VE-cadherin junctions and the endothelial barrier via a
582 phosphatase-independent mechanism. *The Journal of cell biology* **218**, 1725-1742, doi:10.1083/jcb.201807210
583 (2019).
584 59 Mellberg, S. *et al.* Transcriptional profiling reveals a critical role for tyrosine phosphatase VE-PTP in regulation of
585 VEGFR2 activity and endothelial cell morphogenesis. *FASEB journal : official publication of the Federation of*
586 *American Societies for Experimental Biology* **23**, 1490-1502, doi:10.1096/fj.08-123810 (2009).
587 60 Dejana, E., Tournier-Lasserve, E. & Weinstein, B. M. The control of vascular integrity by endothelial cell junctions:
588 molecular basis and pathological implications. *Developmental cell* **16**, 209-221, doi:10.1016/j.devcel.2009.01.004

589 (2009).
590 61 Dominguez, M. G. *et al.* Vascular endothelial tyrosine phosphatase (VE-PTP)-null mice undergo vasculogenesis but
591 die embryonically because of defects in angiogenesis. *Proceedings of the National Academy of Sciences of the United*
592 *States of America* **104**, 3243-3248, doi:10.1073/pnas.0611510104 (2007).
593 62 Jia, W., Yao, Z., Zhao, J., Guan, Q. & Gao, L. New perspectives of physiological and pathological functions of
594 nucleolin (NCL). *Life sciences* **186**, 1-10, doi:10.1016/j.lfs.2017.07.025 (2017).
595 63 Huang, Y. *et al.* The angiogenic function of nucleolin is mediated by vascular endothelial growth factor and
596 nonmuscle myosin. *Blood* **107**, 3564-3571, doi:10.1182/blood-2005-07-2961 (2006).
597 64 Watanabe, T. *et al.* Nucleolin as cell surface receptor for tumor necrosis factor-alpha inducing protein: a carcinogenic
598 factor of Helicobacter pylori. *Journal of cancer research and clinical oncology* **136**, 911-921,
599 doi:10.1007/s00432-009-0733-y (2010).
600
601
602
603
604

605 **Figures and extended data:**

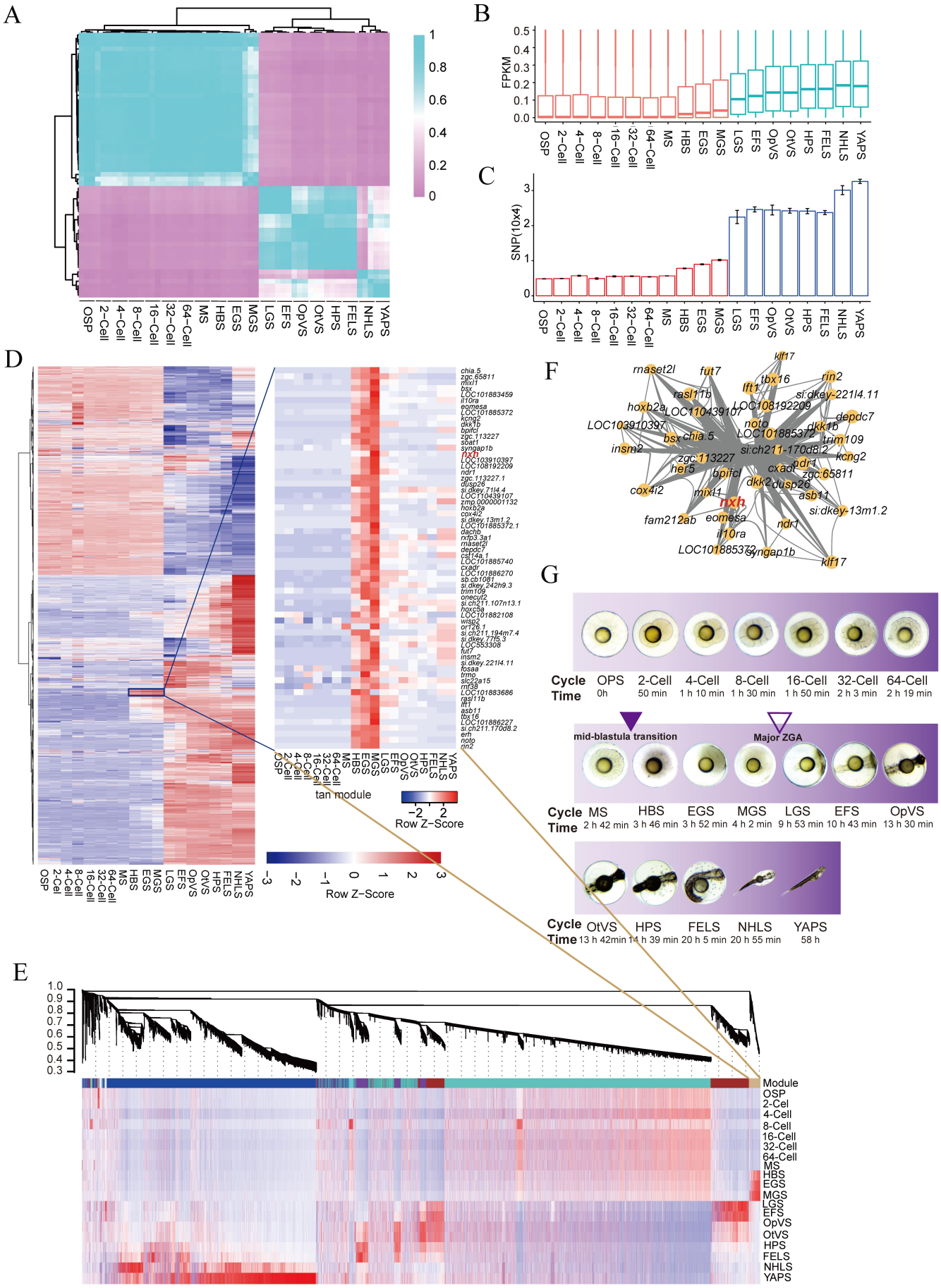
606



607

608 **Figure 1. Genome evolution of golden pompano. A** Phylogenetic relationship of Perciformes and
609 relevant teleost lineages. The position of golden pompano is highlighted in red. Red circles represent the
610 Teleost specific whole genome duplication (Ts3R), Salmonid-specific whole genome duplication (Ss4R),
611 respectively. The divergence time was estimated using the nodes with calibration times derived from the
612 Time Tree database, which were marked by a black rectangle. All estimated divergence times are shown
613 with 95% confidence intervals in brackets. **B** Inspection of whole genome duplication events based on
614 synonymous mutation rate (Ks) distribution. The x axis shows the synonymous distance until a Ks cut-off
615 of 5.2. Note that in order to represent all the data on the same frequency scale, bin sizes are different for
616 each data set. **C** Internal genome synteny of golden pompano. Double-conserved synteny between the
617 golden pompano and spot gar genomes. Each spot gar chromosome (represented as colored blocks) is
618 mostly syntenic with two different chromosomes in the golden pompano genome (syntenic golden
619 pompano regions represented by different colors according to spot gar chromosomal location), a pattern
620 typically associated with whole-genome duplication (Ts3R). Pairs of paralogous genes in spotted gar that

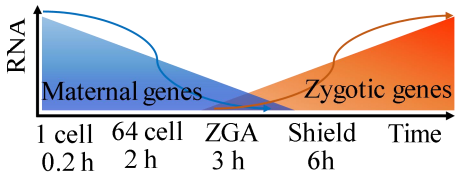
621 are inserted in a double-conserved synteny block are consistent with an origin at the Ts3R event
622 (ohnologues), while genes that are inserted in a double-conserved synteny block but have no paralogue
623 are singletons that have lost their duplicate copy since the Ts3R event. Only genes anchored to
624 chromosomes are represented. **D** Macro-synteny comparison between spotted gar and golden pompano
625 shows the overall one-to-two double-conserved synteny relationship between spotted gar to a post-Ts3R
626 teleost genome. **E** Component of less fragment (LF) and major fragment (MF) subgenomes within golden
627 pompano genome.
628


A
B
C
D
E
F
G
FPKM
SNP(10x4)
OSP
2-Cell
4-Cell
8-Cell
16-Cell
32-Cell
64-Cell
MS
HBS
EGS
MGS
LGS
EFS
OpVS
OtVS
HPS
FELS
NHLS
YAPS
tan module
Row Z-Score
Cycle Time
OPS 0h
2-Cell 50 min
4-Cell 1 h 10 min
8-Cell 1 h 30 min
16-Cell 1 h 50 min
32-Cell 2 h 3 min
64-Cell 2 h 19 min
mid-blastula transition
Major ZGA
MS 2 h 42 min
HBS 3 h 46 min
EGS 3 h 52 min
MGS 4 h 2 min
LGS 9 h 53 min
EFS 10 h 43 min
OpVS 13 h 30 min
OtVS 13 h 42min
HPS 14 h 39 min
FELS 20 h 5 min
NHLS 20 h 55 min
YAPS 58 h
Module

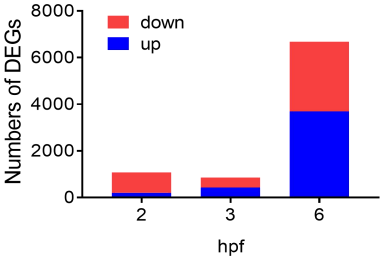
630 **Figure 2. Characterization of Nineteen Continuous Embryonic Development stages from OSP to**
631 **YAPS. A** Heatmap of correlation coefficient within transcriptomics experiments of embryonic
632 development. Expression profile of transcriptomic experiments among embryonic development process.
633 **B** Landscape of transcriptomic expression for 19 embryonic development stages. Cluster of
634 stage-specifically expressed genes were highlighted in red rectangle. **C** Accumulation of transcriptome
635 SNPs within embryonic developments. Error bar represents the standard deviation of biological replicates.
636 **D** Expanded view of expression of all genes in purple modules. The purple model is associated with
637 specific actively expressed genes in minor-ZGA process. **E** WGCNA analysis of Embryonic Development
638 Stages Revealed Gene-Network Modules Enriched in D. **F** Hub-gene network of the purple module. Size
639 of the dots represents hubness. Color of the dots represents the increasing expression level from low to
640 high in preparation stage for ZGA (MGS). Bold text highlights the genes known for *nxh* (EVM0008813).
641 **G** Timing of minor-ZGA and major-ZGA events. The photos of embryonic development process from
642 OSP to YAPS were observed using microscope. The activation of minor-ZGA is beginning at HBS stage
643 and continued to MGS stage. The major-ZGA is activated during transformation stages between MGS and
644 LGS process.
645
646
647
648
649
650
651

A

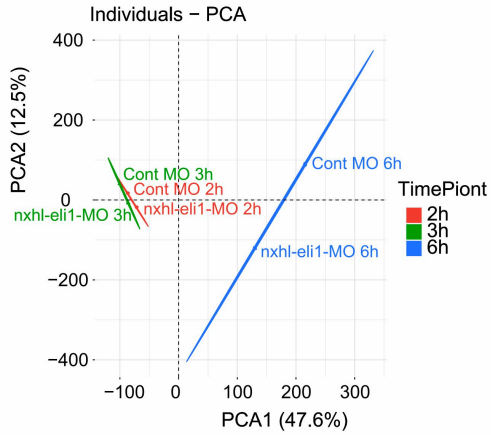
a



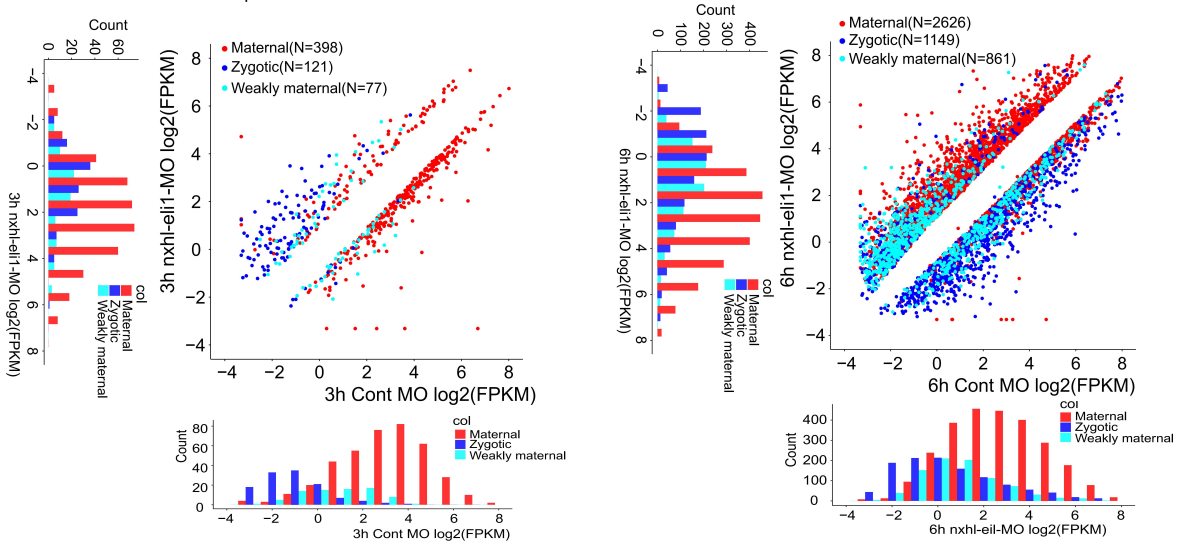
b



c

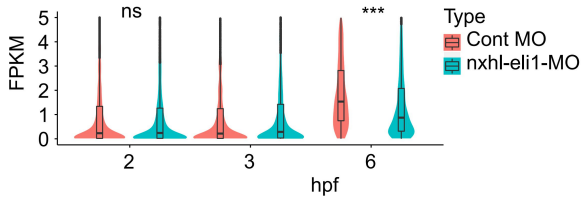


d

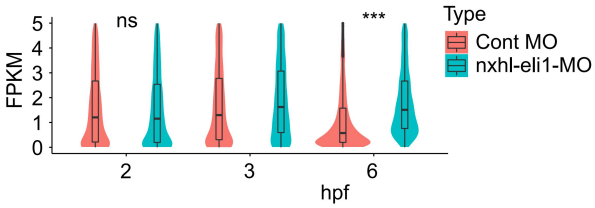


B

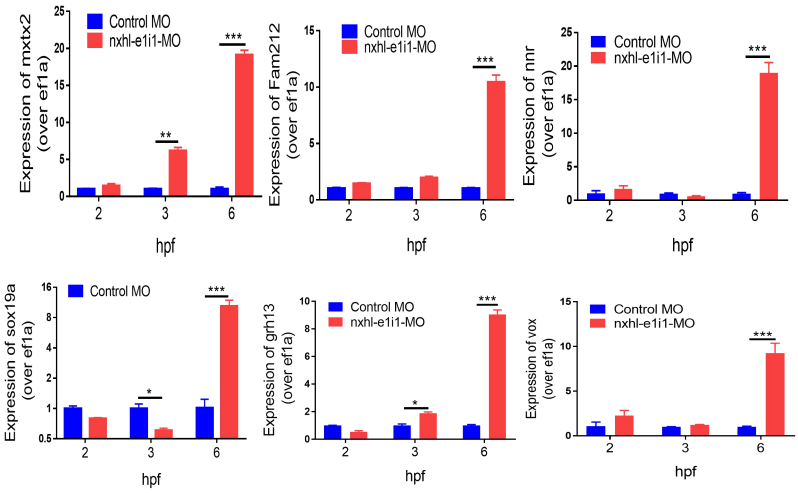
a



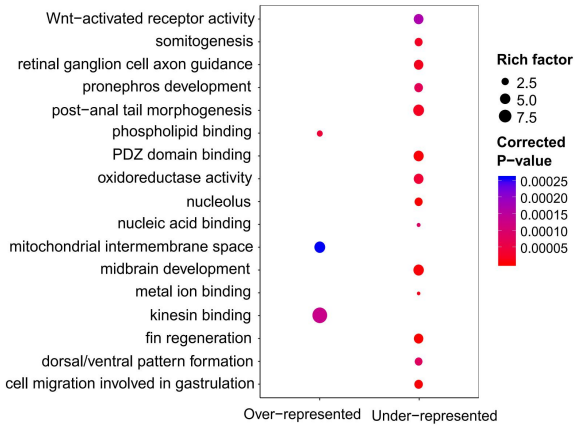
b



c



C



652

653 **Figure 3. *Nxhl* depletion delays the zygotic transcription program.** (**A**) Loss of *nxhl* delays the onset
654 of the ZGA process. **a**, As embryos reach ZGA, maternal transcripts are degraded and zygotic transcripts
655 activated. The numbers below represent time points that were sampled for RNA-seq. hpf, hours post
656 fertilization. **b**, Number of differentially expressed transcripts in *nxhl*-knockdown embryos at stages 2-6
657 hpf. **c**, PCA plot of RNA-seq triplicate samples for pools of wild type (WT) and *nxhl*-knockdown
658 conditions at time points 2-6 hpf. Samples from different conditions show clustered differences at 6 hpf. **d**,
659 Scatterplot of differentially expressed transcripts (596 at 3 hpf, 3636 at 6 hpf, FDR=0.05) between wild
660 type and *nxhl*-knockdown at 3 hpf and 6 hpf. Histograms descripts the number of increased (y axis) and
661 decreased (x axis) transcripts. (**B**) Loss of *nxhl* delays expression of zygotic genes. (**a, b**) Distribution of
662 significantly differentially expressed transcripts in *nxhl*-depleted embryos at the three time points (2, 3
663 and 6 hpf). **a**, FPKM of 3359 transcripts that were reduced in *nxhl*-depleted embryos. **b**, FPKM of 3610
664 transcripts with elevated levels in *nxhl*-depleted embryos. **c**, Quantitative PCR of ZGA marker genes.
665 Embryos were injected at the one-cell stage with 4 ng *nxhl* morpholino, and the wild type and
666 *nxhl*-depleted embryos were pooled for RNA extraction. Data were normalized to eif4g2α and are
667 represent as mean±SEM from three independent experiments. *P* values: *<0.05, **<0.01 and ***<0.001;
668 unpaired t-test. (**C**) Gene ontologies of differentially expressed transcripts in *nxhl*-depleted embryos. R
669 package clusterProfiler was used to analyze differentially expressed transcripts at 6 hpf upon *nxhl*
670 depletion.
671
672
673
674
675

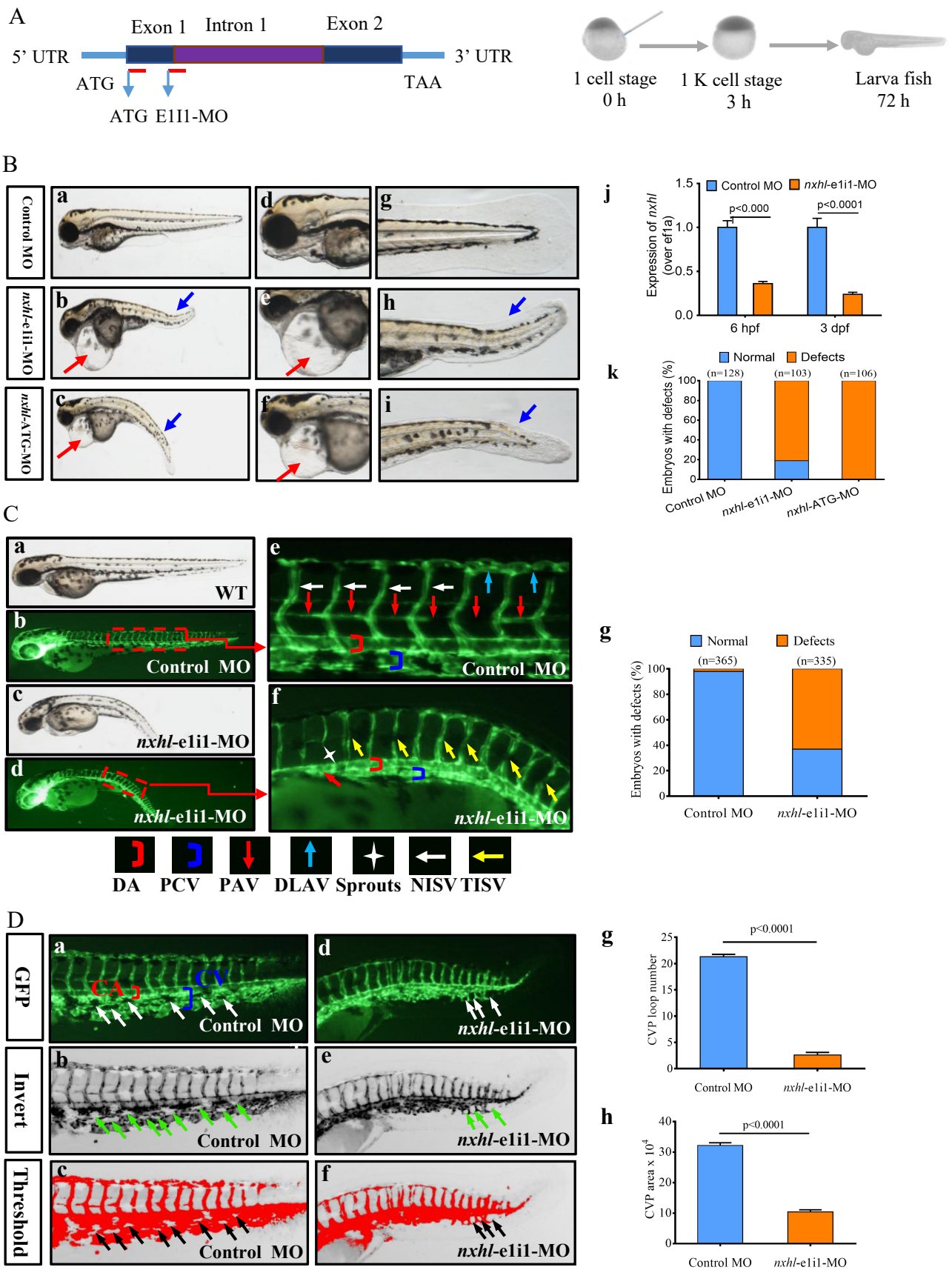


Figure 4. Loss of *nxhl* gene affects ZGA and causes morphological and vascular defects. A Schematic map of *nxhl* gene and the MO design and micro-injection strategy. **B** (**a-i**) Gross morphology at 3 dpf in wild-type AB strain. Compared with control MO, knock down *nxhl* present pericardial oedema

680 (**b**, **e, c, f**, red arrow) and caudal fin defects (**b, c**, **h, i**, blue arrow). Heart beat and circulation in caudal
681 vein (CV) is visible in the control fish, but is abnormal in *nxhl* -MO injected fish (Supplementary Movie1,
682 Supplementary Movie2). (**j**) Validation of MO against *nxhl.* Endogenous *nxhl* in control and *nxhl*
683 morphants assessed by qPCR. Samples were collected at 6 hpf after introduction of 4ng of MO at the
684 one-cell stage (N=30). (**k**) The bar graph shows the percentage of embryos with development defects after
685 knockdown of *nxhl* with *v*-e1i1-MO *and v*-ATG-MO. **C** Morpholino knockdown of *nxhl* causes vascular
686 defects. *Tg(fli1a:EGFP)*^{y1} zebrafish embryos were injected with 4ng control MO or 4 ng *v*-e1i1-MO. **(a-f)**
687 Representative bright field and fluorescent images of *Tg(fli1a:EGFP)*^{y1} embryos at 52 hpf. **(b, e)** Image of
688 trunk regions taken at 52 hpf, with the vascular structures visualized by GFP fluorescence and labelled
689 ISV (intersegmental vessel) and DLAV (dorsal longitudinal anastomotic vessel) showed regular
690 development in the embryo injected with control MO. Compared with control MO, embryos injected with
691 *nxhl*-e1i1-MO present thinner ISVs (yellow arrows) and ectopic sprouts (asterisk) of dorsal aorta **(d, f)**. In
692 control embryos, the parachordal vessels (PAV) form normally (**e**, red arrows). Compared with control,
693 MO knock down *nxhl* prevents the PAV formation, the precursor to the lymphatic system. **g**, The bar
694 graph shows the percentage of embryos with vascular defects after knockdown of *nxhl* with
695 *nxhl*-e1i1-MO. The boxed regions are shown at higher magnification in the right panels. hpf, hours post
696 fertilization.**D** *nxhl* knockdown impairs formation of the CVP in zebrafish. In control embryos, caudal
697 vein plexus (CVP) were formed honeycomb-like structures at the tail around 52 hpf (**a-c**, arrowheads). In
698 contrast, *nxhl* knockdown resulted in specific defects in CVP formation (**d-f**). **g,** Quantification of loop
699 formation at CVP shows a 8.2-fold decreased in *nxhl*-e1i-MO injected embryos at 52 hpf. **h,**
700 Quantification of area at CVP shows a 3-fold decreased in *nxhl*-e1i-MO injected embryos at 52 hpf. CA,
701 caudal artery; CV, caudal vein. NISV, normal intersegmental vessel; TISV, thinner intersegmental vessel.
702 $p<0.05$ represents statistically significant.
703
704
705

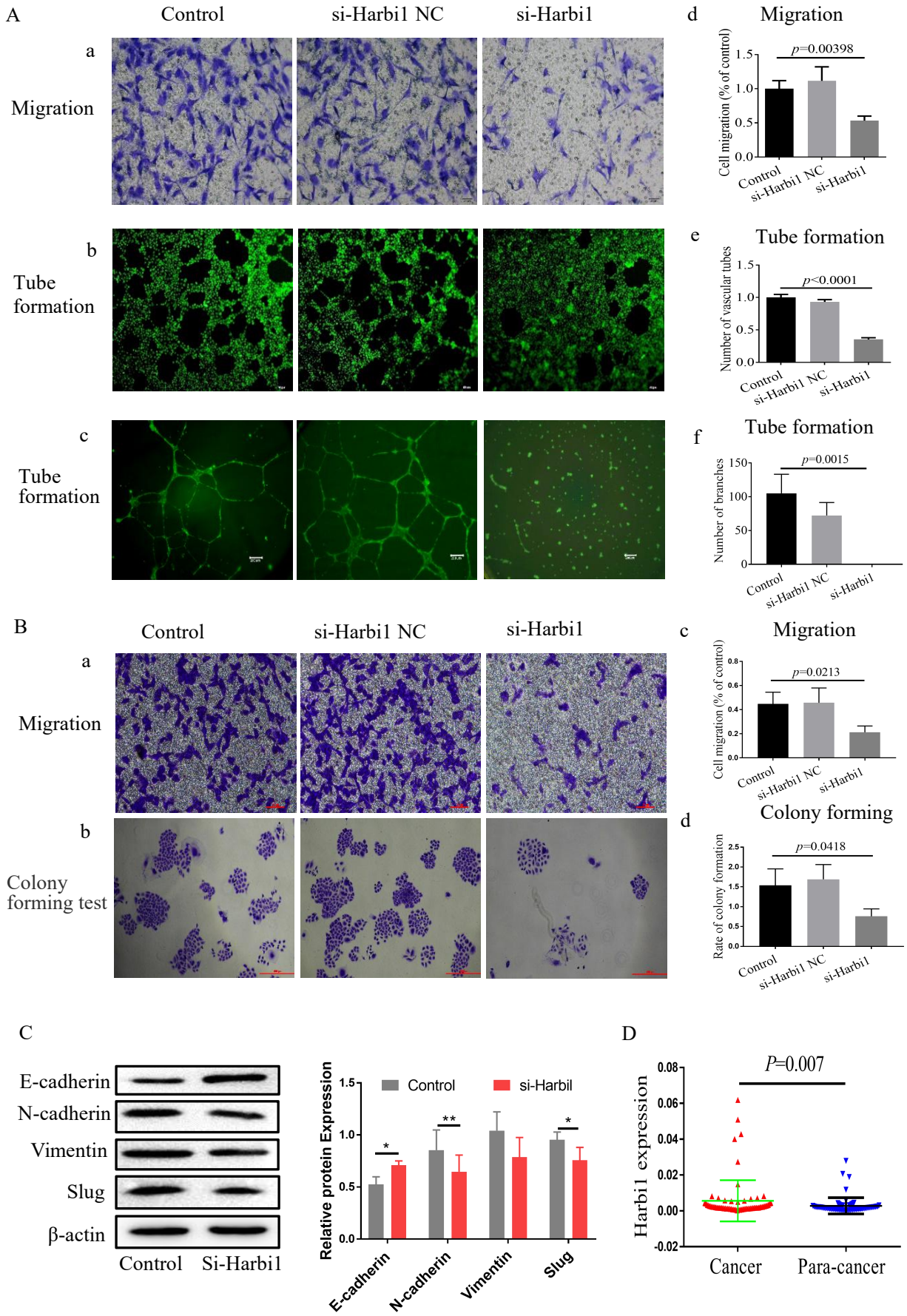


Figure 5. Silencing of Harbi1 not only decreases the angiogenesis of HUVEC cells and cancer cell lines but inhibits migration and EMT of HCC cells. A Silencing of Harbi1 gene suppresses HUVEC

709 cell migration and tube formation abilities, and inhibits angiogenesis of HCC cell. The migratory
710 potential of HUVECs treated with si-Harbi1 was determined using transwell chambers as described in the
711 "Materials and methods" section. (**a, e**) Representative images of migrated cells stained with crystal violet
712 and inhibition of migration in si-Harbi1 treated HUVEC cells. The data represent as mean±SEM from
713 three independent experiments. $p<0.05$ represents statistically significant. (**b, f**) Representative images of
714 tube formation in si-Harbi1 treated HUVEC cells and inhibition of tube formation in vitro. Images were
715 taken 6 h after addition of the supernatant. Tube networks were quantified using the Image J software and
716 expressed as branches. The data represent as mean±SEM from three independent experiments. Scale bars,
717 50 μm. (**c, g)** Representative images of tube formation in si-Harbi1 treated HCC cells. HUVEC cells were
718 plated on matrigel-coated plates at a density of 1.5 × 105 cells/well and incubated for 24 h before the cell
719 culture supernatant of si-Harbi1 treated hepatocellular carcinoma (SMMC7721). The data represent as
720 mean±SEM from three independent experiments. Scale bars, 100 μm. Images were taken 24h after
721 addition of the supernatant. Tube networks were quantified using the Image J software and expressed as
722 branches. **B** Silencing of Harbi1 inhibits migration and EMT of HCC cells. (**a, c**) Representative images
723 of migrated cells stained with crystal violet and inhibition of migration in si-Harbi1 treated SMMC7721
724 cell. (**b, d**) Representative images of colony formation cells stained with crystal violet and inhibition of
725 colony forming in si-Harbi1 treated SMMC7721 cell. **C** Silencing of Harbi1 inhibits
726 epithelial-mesenchymal transition of SMMC7721 cells. Four EMT marker proteins were detected using
727 WB method in control and si-Harbi1-treated SMMC7721 cells. The intensities of the bands are presented.
728 **D** Harbi1expression in 72 HCC tissues and para-cancer tissues. qPCR was used to determine the Harbi1
729 levels. The data represent as mean±SEM from three independent experiments. $p<0.05$ represents
730 statistically significant.
731
732
733

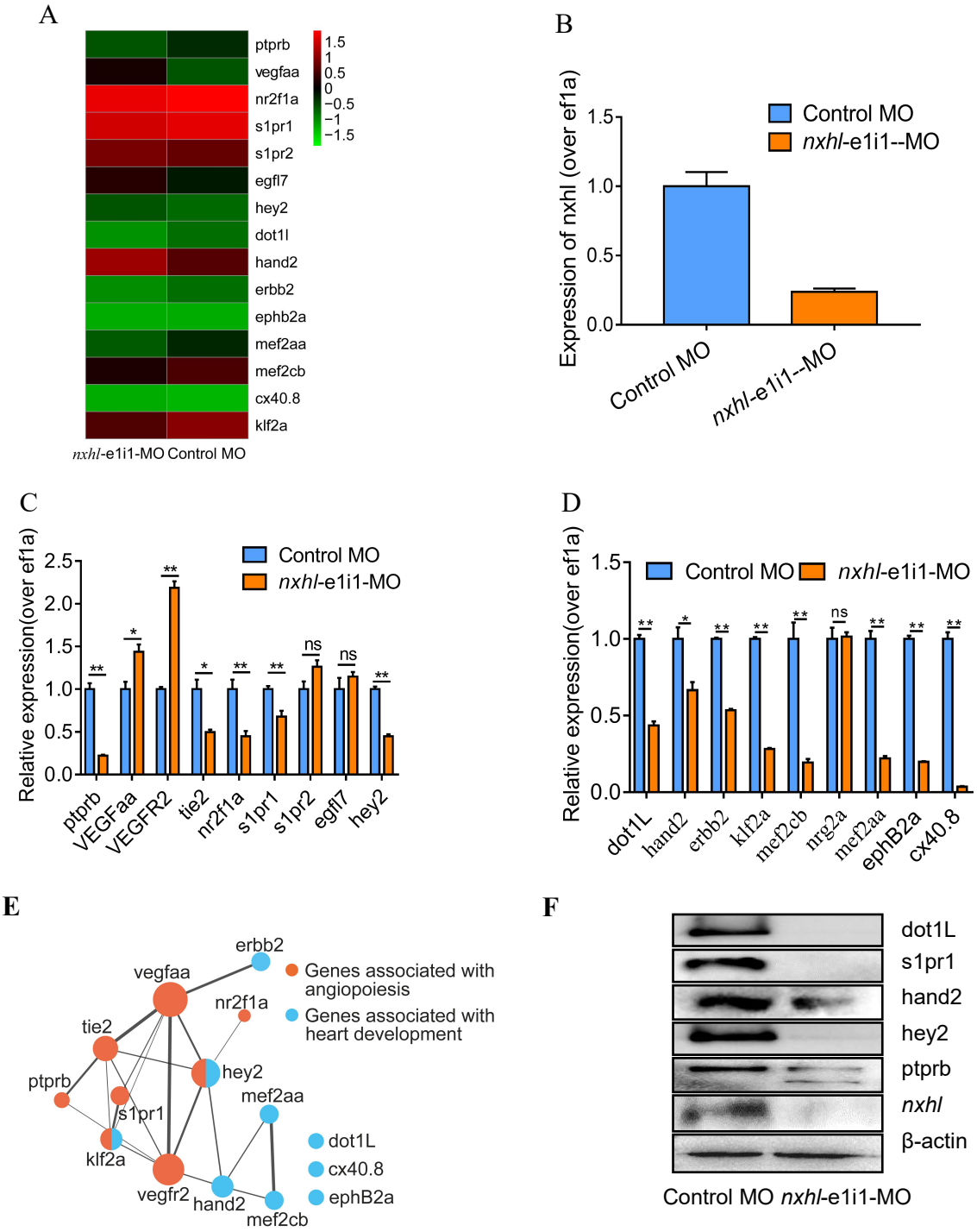
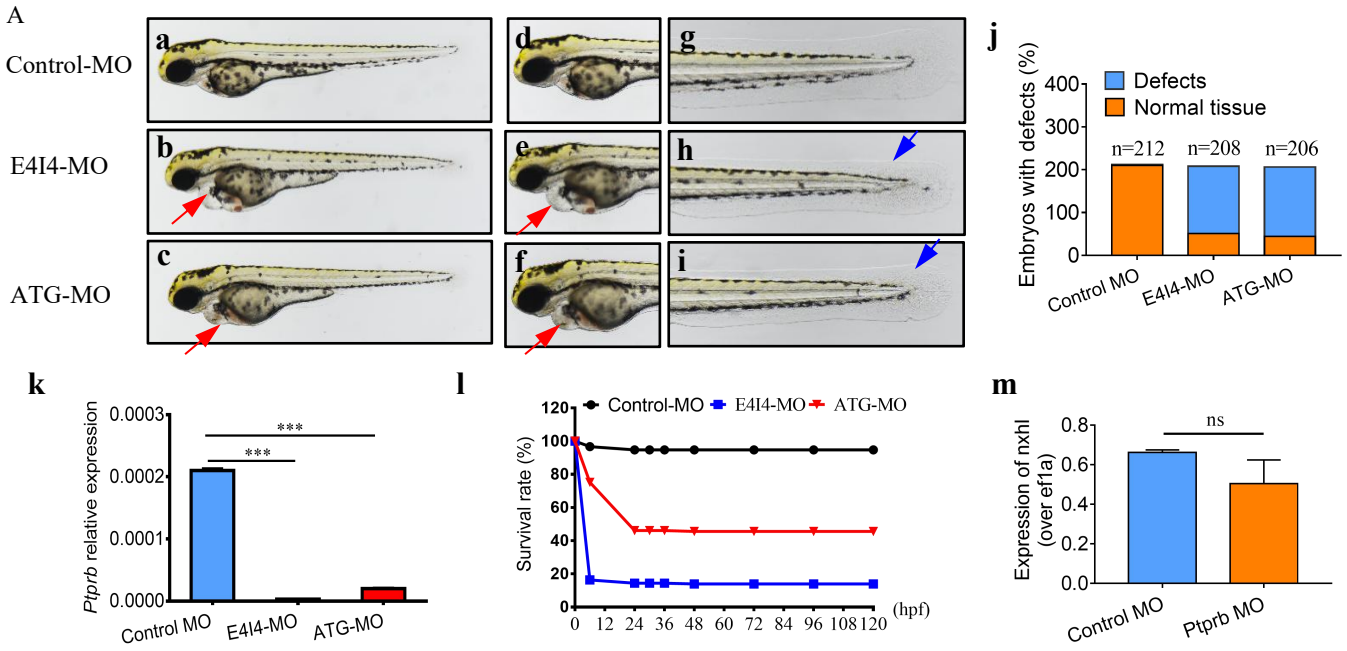
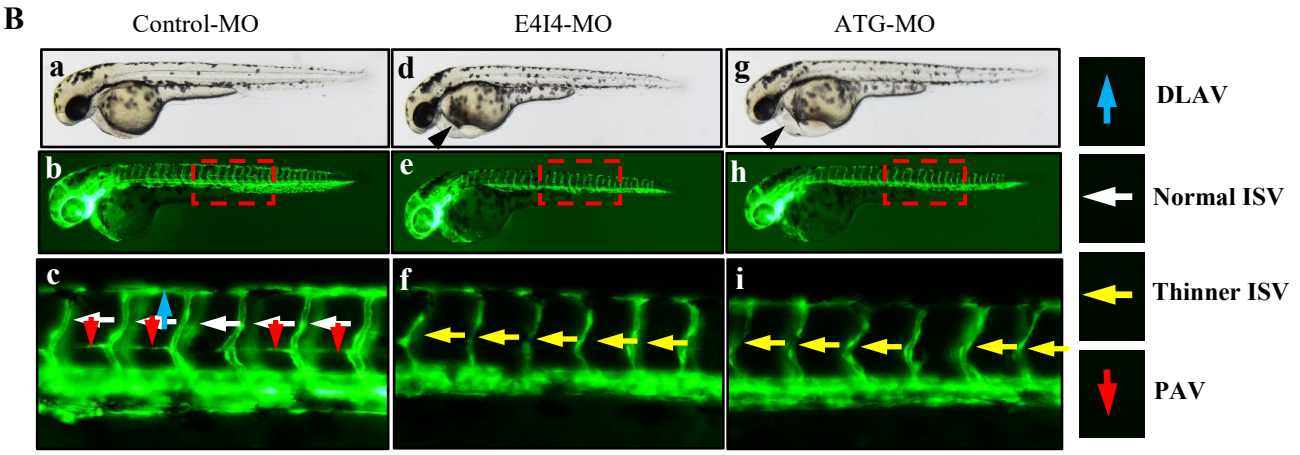
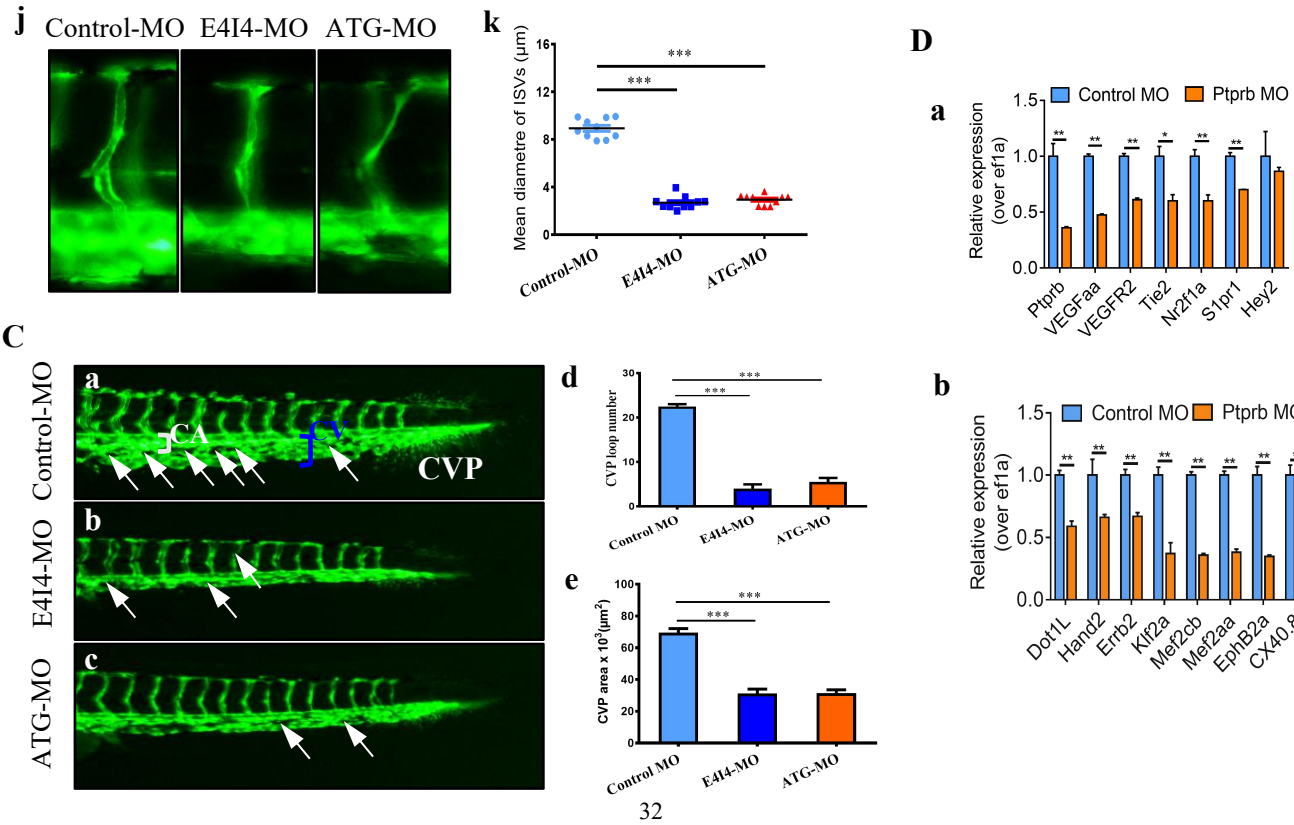


Figure 6. Loss of *nxhl* results in the down-regulation of genes associated with vascular and heart development. A Heatmap of the 15 selected genes from zebrafishes after injection of 4ng *nxhl* ^{e1i1} MO at 3 dpf examined by RNA-seq. **B** Expression of *nxhl* post injection of *nxhl* ^{e1i1} MO 3 dpf. **C** Expression of genes associated with angiopoiesis post injection of *nxhl* ^{e1i1} MO 3 dpf using QPCR. The data represent as mean±SEM from three independent experiments. **D** Expression of genes associated with heart development post injection of *nxhl* ^{e1i1} MO 3 dpf using QPCR. The data represent as mean±SEM from three independent experiments. $p<0.05$ represents statistically significant. **E** Networks of the genes previously reported to be associated with angiopoiesis and heart development. Cytoscope V3.6.1 was

745 used to build this network. **F** Protein levels of the selected genes associated with angiopoiesis and heart
746 development post injection of *nxhl* ^{e1i1} MO 3 dpf by using Western blotting. β-actin antibody was used as
747 internal control.
748
749
750
751
752
753
754
755
756


A
Control-MO
E4I4-MO
ATG-MO
a
b
c
d
e
f
g
h
i
j
Embryos with defects (%)
Defects
Normal tissue
n=212
n=208
n=206
Control MO
E4I4-MO
ATG-MO
k
Ptprb relative expression
0.0003
0.0002
0.0001
0.0000
***
***
Control MO
E4I4-MO
ATG-MO
l
Control-MO
E4I4-MO
ATG-MO
Survival rate (%)
120
100
80
60
40
20
0
0 12 24 36 48 60 72 84 96 108 120
(hpf)
m
Expression of nxhl (over ef1a)
0.8
0.6
0.4
0.2
0.0
ns
Control MO
Ptprb MO


B
Control-MO
E4I4-MO
ATG-MO
a
b
c
d
e
f
g
h
i
DLAV
Normal ISV
Thinner ISV
PAV


j
Control-MO
E4I4-MO
ATG-MO
k
Mean diametre of ISVs (μm)
16
12
8
4
0
***
***
Control-MO
E4I4-MO
ATG-MO
D
a
Control MO
Ptprb MO
Relative expression (over ef1a)
1.5
1.0
0.5
0.0
Ptprb
VEGFaa
VEGFR2
Tie2
Nr2f1a
S1pr1
Hey2
b
Control MO
Ptprb MO
Relative expression (over ef1a)
1.5
1.0
0.5
0.0
Dot1L
Hand2
Errb2
Klf2a
Mef2cb
Mef2aa
EphB2a
CX40.8
C
Control-MO
E4I4-MO
ATG-MO
a
b
c
CA
CV
CVP
d
CVP loop number
30
20
10
0
***
***
Control MO
E4I4-MO
ATG-MO
e
CVP area x 10³(μm²)
100
80
60
40
20
0
***
***
Control MO
E4I4-MO
ATG-MO

758 **Figure 7. Loss of *ptprb* gene causes morphological and vascular defects. A** Loss of *ptprb* induces
759 morphological defects. (**a-i**) Gross morphology at 3 dpf. Compared with control MO, knock down *ptprb*
760 present pericardial oedema (**e, f**, red arrow) and caudal fin defects (**h, i**, blue arrow). (**j**) The bar graph
761 shows the percentage of embryos with development defects after knockdown of *ptprb*. (**k**) Validation of
762 MO against *ptprb.* Endogenous *ptprb* in control and *ptprb* morphants assessed by qPCR. Samples were
763 collected at 3dpf after introduction of 4ng of MO at the one-cell stage (N=30). The data represent as
764 mean±SEM from three independent experiments. (**l**) A time-course plot of percent survival in control vs.
765 *ptprb* morphants for 3 days. dpf, days post fertilization. (m) Expression of *nxhl* post injection of *Ptprb*
766 MO 3 dpf. The data represent as mean±SEM from three independent experiments. **B** Morpholino
767 knockdown of *ptprb* causes vascular defects. *Tg(fli1a:EGFP)*^{y1} zebrafish embryos were injected with 4ng
768 control MO or 4ng *ptprb* MO. **(a-i)** Representative bright field and fluorescent images of
769 *Tg(fli1a:EGFP)*^{y1} embryos at 50 hpf. **(b, c)** Image of trunk regions taken at 50 hpf, with the vascular
770 structures visualized by eGFP fluorescence and labelled ISV (intersegmental vessel) and DLAV (dorsal
771 longitudinal anastomotic vessel) showed regular development in the embryo injected with control MO.
772 Compared with control MO, embryos injected with *ptprb* MO present pericardial oedema **(d, g,** black
773 arrowheads and thinner ISVs (**f, i,** yellow arrows). In control embryos, the parachordal vessels (PAV)
774 form normally (**c**, red arrows). Compared with control, MO knock down *ptprb* prevents the parachordal
775 vessels (PAV) formation, the precursor to the lymphatic system. The boxed regions are shown at higher
776 magnification in the bottom panels. **(j, k)** Quantification of the mean diametre of ISVs shows
777 significantly decrease in *ptprb*-MO injected embryos. Columns, mean; SEM (n =10; ANOVA;) $p<0.05$
778 represents statistically significant.***P < 0.0001. DLAV, dorsal longitudinal anastomotic vessels; ISV,
779 intersegmental vessel; hpf, hours post fertilization**. C** *ptprb* knockdown impairs formation of the CVP in
780 zebrafish. (**a**) In control embryos, caudal vein plexus (CVP) were formed honeycomb-like structures at
781 the tail around 50 hpf (white arrow). In contrast, *ptprb* knockdown resulted in specific defects in CVP
782 formation (**b, c**). (**d)** Quantification of loop formation at CVP shows a 5.6-fold decreased in *ptprb* MO
783 injected embryos at 50 hpf. (**e)** Quantification of area at CVP shows a 2.2-fold decreased in *ptprb* MO
784 injected embryos at 50 hpf. The data represent as mean±SEM from three independent experiments. CVP,
785 caudal vein plexus; CA, caudal artery; CV, caudal vein. **D** Expression of genes associated with
786 angiopoiesis (a) and heart development (b) post injection of *ptprb* MO 50 hpf using QPCR. $p<0.05$
787 represents statistically significant.
788

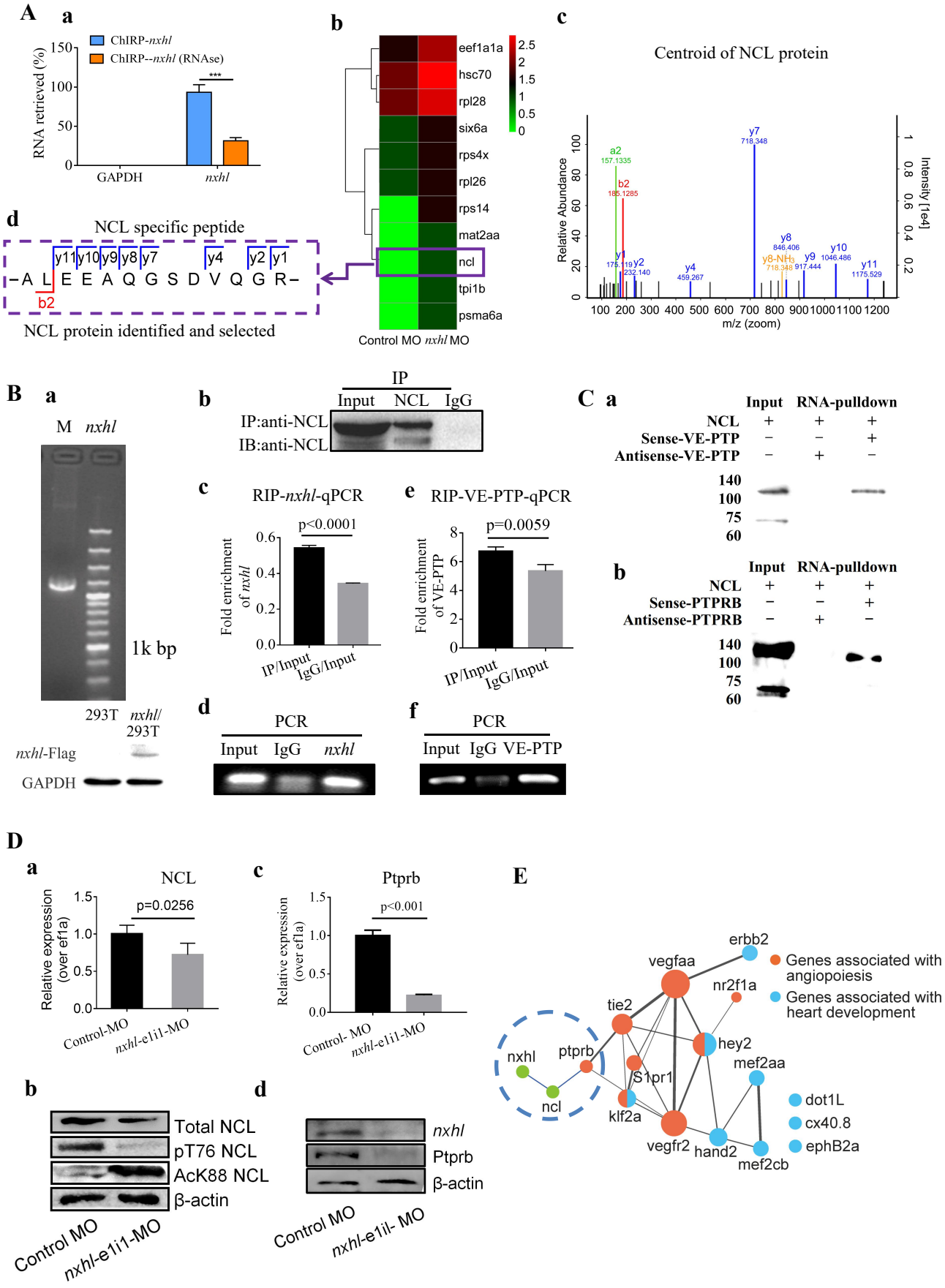


Figure 8. Mechanism of *nxhl* on angiopoiesis development of zebrafish. A ChIRP-MS identification of *nxhl* RNA binding proteins. (**a**) qPCR identification of *nxhl* RNA in the eluted RNAs. Graph shows

794 more than 90% *nxhl* RNA was retrieved, and no GAPDH was detected. (**b**) Heat map of major proteins
795 enriched and significantly (change fold >2 and p<0.05) retrieved by *nxhl* and control probes, analyzed by
796 LC/MS-MS. Among them, NCL protein (purple boxed) was selected as candidate for follow-up study
797 because of its broad association with angiopoiesis. (**c**) Centroid of NCL protein. NCL protein was pull
798 down and identified by LC/MS-MS and centroid of NCL shows its unique m/z feature. The specific
799 peptide (**d**) identifies NCL protein. **B** Physiological interaction between *nxhl,* VE-PTP mRNA and NCL
800 protein, respectively. (**a**) The mRNA expression of *nxhl* was determined by qPCR and Western blotting
801 against Flag antibody was used to identify the successful expression of pcDNA3.1- Flag-*nxhl* plasmid in
802 293T cells. (**b**) Immunoprecipitation (IP) against NCL antibody followed the Western blotting against
803 NCL antibody was performed to verify whether NCL protein was pulled down successfully. (**c, d**) The
804 interaction between *nxhl* mRNA and NCL protein. qPCR and general PCR gel detection for *nxhl* mRNA
805 expression in the NCL-pulled down RNA. The data represent as mean±SEM from three independent
806 experiments. (**e, f**) The interaction between VE-PTP mRNA and NCL protein. qPCR and general PCR gel
807 detection for VE-PTP mRNA expression in the NCL-pulled down RNA. The data represent as
808 mean±SEM from three independent experiments. **C** Physiological interaction between NCL protein and
809 VE-PTP and *ptprb* mRNA, respectively. (**a**) The interaction between VE-PTP mRNA and NCL protein.
810 Western blotting was performed to detect NCL protein in the VE-PTP-biotin probe -pulled down proteins
811 in 293T cells. (**b**) The interaction between *ptprb* mRNA and NCL protein. Western blotting was
812 performed to detect NCL protein in the ptprb-biotin probe -pulled down proteins in zebrafish tissues. **D**
813 Loss of *nxhl* affects the expression of nucleolin (NCL) at both mRNA and protein levels. (**a**) The mRNA
814 expression of NCL was determined by qPCR. The data represent as mean±SEM from three independent
815 experiments. (**b**)The total NCL protein, phosphorylated NCL and acetylated NCL were detected by
816 Western blotting using specific NCL antibodies. (**c**) The mRNA expression of *ptprb* was determined by
817 qPCR. The data represent as mean±SEM from three independent experiments. (**d**) The total *nxhl* and
818 ptprb protein were detected by Western blotting using specific *nxhl* and *ptprb* antibodies. **E** Schematic
819 model illustrating the mechanism of *nxhl* in zebrafish angiopoiesis and heart development. Interaction
820 map of the proteins associated with angiopoiesis and heart development in zebrafish in this study. The
821 interactions between *nxhl* mRNA and NCL protein, NCL protein and *ptprb* mRNA are new-found
822 interactions in this study. *nxhl* may affect the angiopoiesis and heart development via the regulation of the
823 network present.
824

825 **Supplementary figure legends**

826 **Figure S1**. **Investigation of genome features for Golden Pompano. A** K-mer analysis for genome size
827 estimation (k-mer = 21). The estimated genome size is approximate 657 Mb with a heterozygosity rate of
828 0.19%. The genome size was estimated by using the formula: Genome size = K-mer num/Peak depth, and
829 the heterozygosity rate causes a sub-peak at a position half of that of the main peak, whereas a certain
830 repeat rate can cause a similar peak at the position of multiple integers of the main peak. **B** Distribution of
831 GC contents among different sequencing depth. The x-axis represents GC content of 10 kb
832 non-overlapping window.

833 **Figure S2. Distribution of quality and length score of Pacbio subreads. Black line indicated**
834 **accumulation. A** Histogram plot of quality score of SMRT subreads. **B** Histogram plot of read length of
835 SMRT subreads.

836 **Figure S3. Heatmap of interaction frequency of Hi-C data.** The Hi-C data was generated from muscle
837 tissue.

838 **Figure S4. Overview of golden pompano.** Numbers on the circumference are at the megabase scale. **A**
839 Gene density of female *Trachinotus ovatus* (window size = 500 Kb). **B** TE content density of female *T.*
840 *ovatus* (window size = 500 Kb). **C** Genome markers (optical) density of female *T. ovatus* (window size =
841 500 Kb). **D** Hi-C depth of female *T. ovatus* (window size = 500 Kb). **E** GC content of female *T. ovatus*
842 (window size = 500 Kb). **F** Color bands in the middle of the Circos plot connect segmental duplication
843 (minimum five gene pairs) from Teleost-specific whole genome duplication (Ts3R) events.

844 **Figure S5. Gene families evolution of Golden Pompano. A** Venn diagram of gene family comparison in
845 five fish species. **B** Gene family expansion and contraction of golden pompano. Gains and losses are
846 indicated along branches and nodes.

847 **Figure S6. Expression Divergence and Selection Bias among Ancestral Subgenomes. A** Boxplot of
848 expression level of LF, MF and Other gene sets. **B** Selection bias associated with ancestral subgenomes
849 fragmentation. The Ka/Ks values were calculated by orthologous pairs between golden pompano and
850 spotted gar which is outgroup species without Ts3R genome duplication events.

851 **Figure S7. Expression profile of introns in Golden Pompano embryo development.** 19 embryo
852 development stages were collected and transcriptome-sequenced. The FPKM of introns were used to
853 draw the heatmap. A boundary lines between OSP-MGS and LGS-YAPS stages.

854 **Figure S8. Dynamic dominantly expressed genes within LF and MF subgenomes during embryonic**
855 **development process.** Genes whose expression level (FPKM) >= 0.1 are defined as expressed genes.

856 Gene pair having at least one expressed gene is considered for expression bias analysis. Trend of
857 dominant genes is inverse from LGS stage, which is the beginning time of major-ZGA event.

858 **Figure S9. QPCR results of the 8 selected genes in five embryo development stages.** 8 genes in 5
859 embryo stages (between MS and LGS) were randomly picked to verify the gene expression profiles
860 perfected by sequencing. Total RNA was extracted from embryos and reverse transcribed using the the
861 PrimeScript RT reagent Kit. Quantification of gene expression was performed in triplicates using Bio-rad
862 iQ SYBR Green Supermix (Bio-rad) with detection on the Realplex system (Eppendorf). Relative gene
863 expression quantification was based on the comparative threshold cycle method ($2^{-\Delta\Delta Ct}$) using 18s
864 RNA as endogenous control gene.

865 **Figure S10. Characterization of *nxh* gene (EVM0008813). A** Micro-synteny analysis of *nxh* locus
866 among spotted gar, zebrafish, gold pompano and stickleback. Two inversions and one insertion occurred
867 in *nxh* locus region of golden pompano genomes. **B** Validation of expression level for *nxh* by QPCR
868 technology. 18s RNA was considered as internal marker. Gene structure of *nxh* was showed at upper
869 region. **C** Domains of *nxh* and other homologous protein. The domains were identified in SMART
870 database (http://smart.embl.de/).

871 **Figure S11. Confirmation of *nxhl* gene knockdown.** 4 ng morphants per fish embryo was injected. **A**
872 The gel electrophoresis of PCR products of *nxhl* gene in control and *nxhl*-MO groups. Compared with
873 control MO, *nxhl* MO shows quite weak expression of *nxhl* gene. **B** A time-course plot of percent survival
874 in control vs. *nxhl* morphants for 3 days. hpf, hours post fertilization; dpf, days post fertilization.

875 **Figure S12. Phenotypes of *nxhl* zebrafish morphants.** 4 ng morphants per fish embryo was injected. **A**
876 Gross morphology at 3 dpf. Compared with control MO, knock down *nxhl* present pericardial oedema
877 (red arrow) and caudal fin defects (blue arrow). **B** Representative bright field and fluorescent images of
878 *Tg(fli1a:EGFP)*^{y1} embryos at 52 hpf after injection of *nxhl* morphants.

879 **Figure S13 Silencing of Harbi1 not only decreases the angiogenesis of HUVEC cells and cancer cell**
880 **lines but inhibits migration and EMT of HCC cells. A** Silencing of Harbi1 gene suppresses HUVEC
881 cell invasion abilities. The invasive potential of HUVECs treated with si-Harbi1 was determined using
882 transwell chambers as described in the "Materials and methods" section. (**a, b**) Representative images of
883 invasive cells stained with crystal violet in si-Harbi1 treated HUVEC cells. The data represent as
884 mean±SEM from three independent experiments. $p<0.05$ represents statistically significant. **B** Silencing
885 of Harbi1 gene inhibits angiogenesis of two cancer cell lines in vitro. HUVEC cells were plated on
886 matrigel-coated plates at a density of 1.5 × 105 cells/well and incubated for 24 h before the cell culture

887 supernatant of si-Harbi1 treated non-small-cell lung carcinoma (A549) and colon carcinoma (HCT116)
888 cell lines was added, respectively. (**a, c**) Representative images of tube formation in si-Harbi1 treated
889 A549 cells. (**b, d**) Representative images of tube formation in si-Harbi1 treated HCT116 cells. The data
890 represent as mean±SEM from three independent experiments. Scale bars, 100 μm. Images were taken 24h
891 after addition of the supernatant. Tube networks were quantified using the ImageJ software and expressed
892 as branches. **C** Silencing of Harbi1 inhibits migration and EMT of HCC cells. (**a, c**) Representative
893 images of migrated cells stained with crystal violet and inhibition of migration in si-Harbi1 treated
894 SMMC7721 cell. (**b, d**) Representative images of colony formation cells stained with crystal violet and
895 inhibition of colony forming in si-Harbi1 treated SMMC7721 cell. **D** Silencing of Harbi1 inhibits
896 proliferation of HCC cells. MTT method was used to test the cell proliferation. **E** Harbi1 gene expression
897 in HCC samples based on the TCGA data. The data represent as mean±SEM from three independent
898 experiments. $p<0.05$ represents statistically significant.

899 **Figure S14. Gene expression profile post knockdown of *nxhl* comparing with control.** (A). Heatmap
900 of the gene clustering after knockdown of *nxhl* gene comparing with control. (B). Scatter of the
901 differential expressed genes after injection of *nxhl* morphants.

902 **Figure S15.** KEGG enrichment of genes differentially expressed after knockdown of *nxhl* gene
903 comparing with control.

904 **Figure S16. Phenotypes of *ptprb* zebrafish morphants.** 4 ng morphants per fish embryo was injected. **A**
905 Gross morphology at 48-hpf. Compared with control MO, knock down *ptprb* present pericardial oedema
906 (red arrow) and caudal fin defects (blue arrow). **B** Representative bright field and fluorescent images of
907 *Tg(fli1a:EGFP)*^{y1} embryos at 48-hpf after injection of *ptprb* morphants.

908 **Figure S17. Identification of pcDNA3.1-*nxhl* expression plasmid. A** PCR amplification of *nxhl* gene
909 (1267 bp, including enzyme cutting sites NheI and KpnI, and protective bases). **B** Four monoclones were
910 selected by PCR and used for sequencing identification. **C** Sequencing result of *nxhl* gene. The green
911 underline is the fragment inserted into the vector, the vector sequence is on both sides. The red underline
912 is the sequencing result of enzyme cutting site, which shows that the target sequence has been correctly
913 constructed into the vector.

914 **Figure S18. Identification of pcDNA3.1-*nxhl*-Flag expression plasmid. A** PCR amplification of
915 *nxhl*-Flag gene（1306 bp, including enzyme cutting sites NheI and KpnI, and protective bases）based on
916 pcDNA3.1-*nxhl* constructs. **B** Sequencing result of *nxhl*-Flag fragment. The green underline is the
917 fragment inserted into the vector, which shows that the target sequence has been correctly constructed

918 into the vector.

919 **Figure S19. SDS-PAGE electrophoresis of *ptprb* and *nxhl* protein. A** *ptprb* and *nxhl* protein was
920 expressed in pet-b2m Rosetta expression system with kanamycin resistance, respectively. *ptprb* (33 kDa)
921 and *nxhl* (62 kDa) was successfully expressed and the purity of the recombinant protein was 85%. The
922 purified protein was used to immunize rabbits to produce antibody, respectively. M, protein marker. 1,
923 whole bacterial protein. 2, pet-b2m-*nxhl* protein induced by IPTG. 3 and 4 pet-b2m-ptprb protein induced
924 by IPTG. **B, C** Antibody titters of *nxhl* and *ptprb* after three times immunizations of *nxhl* and *ptprb*
925 antigens, respectively. 2ug/ml *nxhl* and *ptprb* protein (100ul/well) was used to coat the plates, respectively.
926 The final antibody concentration of *nxhl* and *ptprb* is 10mg/ml with titter of 1:256000, respectively.

927 **Figure S20. Silence of *nxhl* gene leads to identical hear and vascular phenotypes to that in Figure 4.**
928 **A** Loss of *nxhl* induces morphological defects. Compared with control MO, knock down *nxhl* present
929 pericardial oedema (red arrow) and caudal fin defects (blue arrow). Samples were collected at 3 dpf after
930 introduction of 4 ng of MO at the one-cell stage (N=20). **B** Morpholino knockdown of *nxhl* causes
931 vascular defects. *Tg(fli1a:EGFP)*^{y1} zebrafish embryos were injected with 4ng control MO or 4ng *nxhl*
932 MO. **(a, b)** representative bright field images and (c, d) representative fluorescent images of
933 *Tg(fli1a:EGFP)*^{y1} embryos at 3 dpf. **(e, f)** Image of trunk regions taken at 3 dpf, with the vascular
934 structures visualized by GFP fluorescence and labelled ISV (intersegmental vessel) and DLAV (dorsal
935 longitudinal anastomotic vessel) showed regular development in the embryo injected with control MO.
936 Compared with control MO, embryos injected with *nxhl* MO present thinner ISVs (yellow arrows). **(g, h)**
937 *nxhl* knockdown impairs formation of the CVP in zebrafish. In control embryos, caudal vein plexus (CVP)
938 were formed honeycomb-like structures at the tail around 3 dpf (red arrow). In contrast, *nxhl* knockdown
939 resulted in specific defects in CVP formation. CA, caudal artery; CV, caudal vein. These phenotypes were
940 identical to that in Figure 4.

941 **Figure S21.** Possible mechanism of *nxhl* in zebrafish angiopoiesis and heart development. Knockdown of
942 *nxhl* may downregulate the *nxhl*-NCL-*ptprb* complex, subsequently regulate the proteins associated with
943 angiopoiesis and heart development, and finally result in heart pericardial oedema, vascular patterning
944 and integrity (lumen formation) defects.

945

946 **Supplementary Tables:**

947 **Table S1a.** Summary of genome sequencing strategy for golden pompano.

948 **Table S1b.** Statistics of PacBio sub-reads length distribution.

949 **Table S1c.** Summary of BioNano data collection and assembly statistics.

950 **Table S1d.** Summary of Hi-C data for golden pompanos.

951 **Table S2a.** Summary of assembly results for golden pompano.

952 **Table S2b.** Statistics of chromosomes for golden pompano based on Hi-C data.

953 **Table S3a.** PacBio sub-reads validation for golden pompano genome assembly.

954 **Table S3b.** Genome completeness assement by BUSCO.

955 **Table S3c.** Summary of pooled transcriptome data assisted for genome annotation.

956 **Table S3d.** Genome completeness evaluated by ESTs/unigenes.

957 **Table S4.** Repeat content of female and male golden pompano.

958 **Table S5.** Summary of nc-RNA of golden pompano.

959 **Table S6a.** Comparison of gene features among golden pompano, zebrafish, large yellow croaker,
960 antarctic bullhead notothen and Pacific bluefin tuna.

961 **Table S6b.** Summary of gene models annotated by different non-redundant databases.

962 **Table S7.** Lineage Specific (LS) genes annotation for Golden pompano.

963 **Table S8.** KEEG analysis of positive selection genes for Golden pompano.

964 **Table S9.** Positive selection genes of Pacific bluefin tuna.

965 **Table S10.** GO term analysis of positive selection genes for Pacific bluefin tuna.

966 **Table S11.** KEEG analysis of positive selection genes for Pacific bluefin tuna.

967 **Table S12.** Positive selection genes of Golden pompano.

968 **Table S13.** GO term analysis of positive selection genes for Golden pompano.

969 **Table S14.** KEEG analysis of positive selection genes for Golden pompano.

970 **Table S15.** Common positive selection genes of Golden pompano and Pafici bluefin tuna.

971 **Table S16.** Pfam annotation of lineage specific genes for gold pompano.

972 **Table S17.** Summary of transcript factor (TF) among five genome sequenced fished in Perciformes and
973 Zebrafish.

974 **Table S18.** Summary of syntenic gene pairs between Spotted gar and Golden pompano.

975 **Table S19.** The enriched molecular function of G17, G19, G20 and G22.

976 **Table S20.** The enriched KEGG pathway of G17, G19, G20 and G22.

977 **Table S21.** Annotation of the differentially expressed genes.

978 **Table S22.** Differentially expression analysis between *nxhl* MO and Control MO.

979 **Table S23.** GO analysis of the down-regulated genes between *nxhl* MO and Control MO.

980 **Table S24.** GO analysis of the up-regulated genes between *nxhl* MO and Control MO.

981 **Table S25.** KEGG pathway analysis.

982 **Table S26.** All genes involved in KEGG pathway.

983 **Table S27** All genes differentially expressed after nxhl depletion.

984 **Table S28** Go terms of the signifficantly differentially expressed genes.

985 **Table S29.** GO enrichment of genes involved in blood vessel morphogenesis and their differential
986 expression value.

987 **Table S30.** Primers and probes used in this study.

988

989