

1 **Protamine lacking piscine spermatozoa are transcriptionally**  
2 **active**

3

4 **Júlia Castro-Arnau<sup>1</sup>, François Chauvigné<sup>1</sup>, Jessica Gómez-Garrido<sup>2</sup>, Anna Esteve-**  
5 **Codina<sup>2</sup>, Marc Dabad<sup>2</sup>, Tyler Alioto<sup>2,3</sup>, Roderick Nigel Finn<sup>1,4</sup> and Joan Cerdà<sup>1\*</sup>**

6

7 <sup>1</sup>Institute of Agrifood Research and Technology-Institute of Biotechnology and  
8 Biomedicine, Universitat Autònoma de Barcelona, 08193 Barcelona, Spain; <sup>2</sup>Centre  
9 Nacional d'Anàlisi Genòmica-Centre for Genomic Regulation, Barcelona Institute of  
10 Science and Technology, 08028 Barcelona, Spain; <sup>3</sup>Universitat Pompeu Fabra, 08003  
11 Barcelona, Spain; <sup>4</sup>Department of Biological Sciences, Bergen High Technology Centre,  
12 University of Bergen, 5020 Bergen, Norway

13

14 **Key words:** sperm maturation, extratesticular ducts, transcription, GnRH, PDGF, sperm  
15 motility, male fertility

16

17

18

19 \*To whom correspondence may be addressed. Email: [joan.cerda@irta.cat](mailto:joan.cerda@irta.cat)

20

21 **Abstract**

22

23 Transcriptional quiescence of post-meiotic spermatozoa associated with protamine-  
24 mediated chromatin condensation is widely recognized in animals. How sperm acquire the  
25 extratesticular maturational competence to move and fertilize the egg is therefore thought to  
26 occur via non-transcriptional mechanisms. Here, using transcriptional profiling during  
27 spermatozoon differentiation in a fish that does not condense chromatin with protamines,  
28 we uncover spatially distinct roles of the GnRH receptor and PDGF signaling pathways  
29 between the somatic epithelia of the extratesticular ducts and the maturing spermatozoa. In  
30 vitro induction and inhibition experiments demonstrate that the endocrine signaling  
31 pathways are conserved in different lineages of fish and activate de novo transcription of  
32 spermatozoon genes required for the acquisition of full motility. These experiments further  
33 confirmed that mitochondrial translation is important for sperm maturation in anamniotes  
34 as in amniotes, but that transcriptional quiescence of post-meiotic spermatozoa is not a pan  
35 vertebrate phenomenon. On the contrary, the data show that the identified signal  
36 transduction pathways between the soma and the sperm upregulate effector genes essential  
37 for maturational competence and male fertility.

38

39

## 40 Introduction

41

42 Vertebrate spermatogenesis proceeds in a multistage process from mitotic expansion of  
43 spermatogonial stem cells to form primary spermatocytes (spermatocytogenesis) through  
44 two meiotic divisions to form spermatids (spermatidogenesis) and a tertiary phase of  
45 differentiation (spermiogenesis) to form the highly polarized spermatozoa that retain a  
46 recombined haploid genome (*de Kretser et al., 1998; Schulz et al., 2010; Nishimura and*  
47 *L'Hernault, 2017*). This is regardless of whether the germ cells develop in cysts in  
48 anamniotes (fishes and amphibians) or in the acystic epithelial lining of the seminiferous  
49 tubules in amniotes (reptiles, birds and mammals) (*Yoshida, 2016*). At the culmination of  
50 testicular spermatogenesis however, the fully differentiated spermatozoa are typically not  
51 capable of fertilizing the egg (*Nixon et al., 2020; Pérez, 2020*). They require a maturational  
52 phase, which confers the physiological ability to move, recognize and penetrate the egg  
53 (*Nixon et al., 2020; Pérez, 2020*). In most vertebrates, this process proceeds during sperm  
54 storage and transit through the extratesticular excurrent ducts (ETDs) and tubular systems  
55 that emanate from the testis (*Sullivan and Mieusset, 2016*). Such ETDs are thought to have  
56 evolved in the common ancestor of jawed vertebrates becoming ever more convoluted to  
57 form the epididymis in amniotes (*Jones, 2002*). Since humans are members of this latter  
58 group, considerable research has been invested to understand the epididymal regulation of  
59 sperm maturation and the aetiology of asthenozoospermia (*Sullivan and Mieusset, 2016*).  
60 By contrast, almost nothing is known of the molecular signaling pathways that regulate  
61 sperm maturation in anamniotes.

62 Both transcriptomic and proteomic studies in mammals have established that gene  
63 expression is highly segmented along the length of the epididymis (*Sullivan and Mieusset,*  
64 *2016; Belleannée et al., 2012; Zhao et al., 2019*). Conversely, despite presenting hundreds  
65 of proteins and carrying thousands of RNAs of different types, the transcriptional and  
66 translational activities of the maturing spermatozoa are virtually silent (*Fisher et al., 2012;*  
67 *Grunewald et al., 2005; Ren et al., 2017; Freitas et al., 2020*). Such quiescence is in stark  
68 contrast to the stellar transcriptional activity of the spermatogenic cells, which exceed all  
69 other cell types by expressing >80% of the protein coding genes in the genome (*Soumillon*  
70 *et al., 2013; Xia et al., 2020*). The onset of transcriptional and translational quiescence

71 occurs during the spermiogenic differentiation phase when large numbers of rRNAs are  
72 degraded, the cytoplasm and nucleoplasm are discarded, and the histones of the DNA-  
73 packing nucleosomes are gradually replaced by protamines (*Ren et al., 2017; Rathke et al.,*  
74 *2014*). The many types of RNAs still present in the sperm are thus thought to be the  
75 remnants of the high transcriptional activity of the spermatocytic and spermatidogenic  
76 phases (*Ren et al., 2017*). Alternatively, many proteins and some types of RNAs may be  
77 delivered via epididymal exosomes-epididysomes (*James et al., 2020*), which partially  
78 solves the problem of the lack of cytoplasmic ribosomes for protein translation. Other  
79 studies suggest that mitochondrial ribosomal pathways, rather than the canonical  
80 cytoplasmic mechanisms, remain active and yield paternal factors that are important for  
81 sperm maturation, capacitation in the female reproductive tract, fertility and early zygotic  
82 development (*Gur and Breitbart, 2006; Zhao et al., 2009; Rajamanickam et al., 2017;*  
83 *Zhu et al., 2019*). In all cases, however, *de novo* transcription in the maturing mammalian  
84 sperm is not considered to be a major source of RNAs or proteins.

85 Interestingly, not all vertebrate sperm retain protamines in their nuclei. Despite  
86 protamines first being discovered in fish, the Rhine salmon (*Miescher, 1874*), it has  
87 become evident that the spermatozoa of several lineages of anamniotes lack such highly  
88 arginine-enriched forms of the sperm nuclear basic proteins (SNBP) (*Ausió, 1999; Eirin-*  
89 *López et al., 2006*). Even when present, it has been shown that protamines may not be  
90 involved in spermatogenic nuclear condensation (*Shimizu et al., 2000*). It is thus not known  
91 whether transcriptional quiescence is a general feature of post-meiotic sperm maturation in  
92 vertebrates, and if not what role such late-stage transcription might play. To address these  
93 questions, we conducted transcriptome profiling of haploid germ cells and ejaculated  
94 spermatozoa from a species of fish, the gilthead seabream (*Sparus aurata*), which produces  
95 profuse amounts of sperm without protamines (*Kurtz et al., 2009*). Gene set enrichment  
96 analysis revealed the regulation of a high number of transcripts potentially involved in  
97 transcription, translation and chromatin organization in spermatozoa, as well as of several  
98 signaling pathways, of which the gonadotropin-releasing hormone receptor (GnRHR) and  
99 platelet-derived growth factor (PDGF) were amongst the most dominant. Experimental  
100 investigation of the origin of these signaling pathways uncovered their expression in  
101 sequential segments of the ETDs where their paracrine signaling induces the *de novo*

102 transcription of genes in the post-meiotic spermatozoa. The importance of these  
103 mechanisms for sperm maturation in seabream, as well as in zebrafish (*Danio rerio*), as a  
104 model from a more ancestral teleost lineage that produces small volumes of sperm lacking  
105 protamines (*Wu et al., 2011*), was confirmed through motility tests in the presence of  
106 transcription and translation inhibitors. The present data thus uncover soma to germ cell  
107 signaling pathways during sperm maturation in vertebrates and reveal that post-meiotic  
108 spermatozoa may not remain transcriptionally silent. On the contrary, such late-stage  
109 transcriptional activation induced by ETD epithelial endocrine signaling upregulates  
110 gametic cell effector genes that are required for the acquisition of full sperm motility.

111

## 112 **Results**

113

### 114 **Transcriptome profiling of haploid germ cells and mature spermatozoa**

115 The changes in gene expression during the differentiation and maturation of seabream  
116 spermatozoa were investigated by whole-transcriptome RNA-seq of haploid germ cells  
117 (HGCs) and ejaculated (mature) spermatozoa (SPZ<sub>EJ</sub>). The HGCs were isolated by  
118 fluorescence-activated cell sorting (FACS), whereas SPZ<sub>EJ</sub> were collected by manual  
119 stripping of naturally spermiating males. Flow cytometry of the extract from the seabream  
120 whole mature testis showed that the percentage of diploid and haploid cells reached 16%  
121 and 84% of the total cells, respectively (*Figure 1A*). The percentage of diploid cells was  
122 lower than expected because the centrifugation steps of the extract before cell sorting  
123 partially depleted this population. Flow cytometry identified two subpopulations of haploid  
124 cells based on their relative size and SYBR Green I fluorescence intensity: a subpopulation  
125 formed by spermatocytes II and spermatids (SPC II and SPD, respectively), which we refer  
126 here as HGCs, and another subpopulation formed by intratesticular spermatozoa (SPZ<sub>I</sub>)  
127 (*Figure 1A and B*). The percentage of HGCs and SPZ<sub>I</sub> in the testicular extracts was of 34 ±  
128 4% and 66 ± 4% ( $n = 15$ ), respectively (*Figure 1B*).

129 Microscopic examination of the HGC-enriched population after FACS confirmed the  
130 presence of SPC II, and round and elongating SPD in this fraction (*Figure 1C*). Whole-  
131 mount immunostaining revealed strong expression of Lys<sup>9</sup> acetylated histone 3 (H3K9ac)  
132 and meiotic recombination protein Spo11 in SPC II, which progressively decreased in

133 round and elongating SPD, and completely vanished in SPZ<sub>EJ</sub> (**Figure 1C**).  
134 Immunolocalization of  $\alpha$ -tubulin (Tuba) showed that the protein was spread in the  
135 cytoplasm in SPC II and round SPD, became also detectable in the nascent flagellar region  
136 of elongating SPD, and was finally distributed along the flagellum of differentiated SPZ<sub>EJ</sub>  
137 (**Figure 1C**). These observations indicate a high occurrence of meiotic recombination in  
138 SPC II and a progressive DNA condensation during the differentiation of SPC II into SPD  
139 and spermatozoa, which are conserved features during vertebrate germ cell development  
140 (**Kurtz et al., 2009; Hazzouri et al., 2000**). Therefore, these data confirmed that the sorted  
141 population of cells from the mature seabream testis correspond to HGCs before  
142 differentiation into spermatozoa.

143 Four unstranded RNA libraries (replicates) for low-input RNA were subsequently  
144 constructed for each of the two HGC and SPZ<sub>EJ</sub> cell types; each replicate being a pool of  
145 cells collected from three different males. Library sequencing rendered 30-62 million reads  
146 per library comprising a yield of 5-10 Gb. From these data, we produced a new integrative  
147 *S. aurata* genome annotation before the RNA-seq analysis. This new annotation was carried  
148 out by re-annotating the available *S. aurata* reference genome (**Pauletto et al., 2018**), and  
149 by adding 202 *de novo* assembled transcripts that were not present in the genome assembly.  
150 In total, 31,501 protein-coding genes were annotated, which produced 57,396 transcripts  
151 (1.82 transcripts per gene) and encoded for 51,365 unique protein products. Functional  
152 labels were assigned to 62% of the annotated proteins. In addition, 165,898 non-coding  
153 transcripts were annotated, of which 159,925 are long non-coding RNA (lncRNA) genes  
154 and 5,973 correspond to short non-coding RNAs.

155 Principal component analysis (PCA) of the expression data showed that FACS-  
156 purified HGC and SPZ<sub>EJ</sub> formed two relatively well-separated clusters, suggesting that the  
157 developmental stage has a large effect on the pattern of gene expression (**Figure 1D**).  
158 However, while the four HGC biological replicates clustered very close, those of SPZ<sub>EJ</sub>  
159 were more distant, indicating a higher variability in the transcriptome of the SPZ<sub>EJ</sub>  
160 replicates. Nevertheless, the RNA-seq analysis revealed a total of 7,287 differentially  
161 expressed genes (DEGs) (adjusted *p*-value < 0.01) between both cell types, of which nearly  
162 half (3,447) were upregulated in SPZ<sub>EJ</sub> when compared to HGCs (**Figure 1E-G**). In  
163 addition, 239 transcripts were detected only in SPZ<sub>EJ</sub> (**Figure 1G**). Finally, as previously

164 reported in the human spermatozoon (*Corral-Vázquez et al., 2021*), we also found a high  
165 number of differentially expressed lncRNAs (9,059 sequences) of which 5,114 were  
166 upregulated and 3,446 unique in SPZ<sub>EJ</sub> (*Figure 1G*).

167 The quality of the RNA-seq data and the reliability of the DEGs identified were  
168 validated on randomly selected 45 DEGs by real-time quantitative reverse transcription  
169 PCR (qRT-PCR) in three biological replicates. Fold changes from qRT-PCR were  
170 compared with the RNA-seq expression profiles (*Figure 1H*). The dynamic expression  
171 patterns of all genes were consistent with the RNA-seq analysis, showing a high correlation  
172 (Pearson's correlation coefficient of 0.892) between RNA-seq and qRT-PCR data. These  
173 results therefore indicated the reliability of the RNA-seq for mRNA differential expression  
174 analysis.

175

## 176 **Functional enrichment analysis of DEGs during spermatozoa differentiation and** 177 **maturation**

178 Gene ontology (GO) term-enriched analysis of DEGs in SPZ<sub>EJ</sub> with significant differences  
179 revealed that a large number of biological processes were represented. The five top-ranked  
180 GO terms were regulation of biological, cellular and metabolic processes, and organic  
181 substance and metabolic processes (*Figure 2-Supplement 1A*). Further analysis of GO term  
182 distribution indicated that the most represented biological process was the regulation of  
183 gene expression, followed by positive regulation of macromolecule and cellular  
184 metabolism, regulation of signal transduction, and regulation of cellular biosynthesis  
185 (*Figure 2-Supplement 1B*). Interestingly, genes with GO terms such as cellular response to  
186 stimulus, cell communication, signal transduction, response to external or chemical  
187 stimulus, cell adhesion, and cell surface receptor signaling pathway, were only upregulated  
188 in SPZ<sub>EJ</sub> (*Figure 2-Supplement 1A*). For the GO molecular function, the top enriched  
189 terms were binding to ribonucleotides and purine nucleotides, whereas the terms Ca<sup>2+</sup>,  
190 phosphatidylinositol and actin binding, ion channel activity, and transmembrane transport  
191 of inorganic cations and organic anions appeared to be only upregulated in SPZ<sub>EJ</sub> (*Figure*  
192 *2-Supplement 1C*). Taken together these findings indicate the enrichment of gene  
193 expression, metabolic and signaling processes in SPZ<sub>EJ</sub>.

194 To gather more information on genes with a potential impact on spermatozoa  
195 function, the DEGs in SPZ<sub>EJ</sub> were manually classified into five functional categories by  
196 using GO analysis and the Uniprot database. These categories included transcription,  
197 translation and chromatin organization (1,056 genes), receptors (433 genes), metabolism of  
198 proteins, lipids and carbohydrates (492 genes), cytoskeleton and cell movement (520  
199 genes), and channels, exchangers and transporters (308 genes) (**Figure 2A**). The genes  
200 upregulated in SPZ<sub>EJ</sub> related to transcription, translation and chromatin organization (443  
201 genes) mainly correspond to transcription factors (42.5%) and regulators of transcription  
202 (20.1%), followed by ribosome structure (12%), regulators of translation (5.4%), chromatin  
203 and RNA binding (4.1 and 4.7%, respectively), and histones and histone modification  
204 (6.8%) (**Figure 2B**). Most of the receptor-encoding upregulated genes (303 genes) were G  
205 protein-coupled receptors (36.6%), tyrosine phosphatase and kinase receptors (11.5%),  
206 cytokine receptors (7.3%), as well as other receptors mainly including Fc receptors and  
207 novel immune-type receptors (**Figure 2C**). For metabolic processes (253 genes), the most  
208 enriched genes in SPZ<sub>EJ</sub> were those related to glycolysis and gluconeogenesis (8.7%), the  
209 metabolism of glycogen and other polysaccharides (11.8%), fatty acids (26.1%) and amino  
210 acids (11.5%), and proteases (17.4%) (**Figure 2D**). Finally, genes encoding for proteins  
211 involved in cytoskeletal organization (32.1%), actin binding (21.4%) and motor proteins  
212 (15.7%) were the most abundant upregulated genes involved in the cytoskeleton and cell  
213 movement (149 genes) (**Figure 2E**), whereas in the group including upregulated genes  
214 encoding for channels, exchangers and transporters (139 genes) the K<sup>+</sup> and metal specific  
215 channels (19.9%), cation channels (13.2%) and peptide and amino acid transporters  
216 (16.9%) were the most enriched in SPZ<sub>EJ</sub> (**Figure 2F**).

217 In an effort to identify specific transcription/translation and metabolic processes  
218 enriched in SPZ<sub>EJ</sub>, we built the protein interactome network of DEGs classified into these  
219 two categories by using the STRING protein-protein interaction (PPI) database for known  
220 PPIs (*Szklarczyk et al., 2019*) with very stringent inclusion criteria. As a result, a connected  
221 network comprising 766 proteins and 3,588 connections was mapped for the proteins  
222 encoded by genes involved in transcription and translation (**Figure 3A**). These proteins  
223 could be divided into five major subclusters based on their known biological functions  
224 established through GO analysis, including mitochondrial translation, tRNA



225 aminoacylation, translation initiation, cytosolic ribosomes and mRNA splicing (**Figure**  
226 **3A**). All of the DEGs grouped into the cytosolic ribosome subunit subcluster, and half of  
227 the DEGs belonging to the tRNA aminoacylation, mitochondrial translation, and translation  
228 initiation subclusters, were upregulated (**Figure 3A**). These findings, together with the  
229 previous observation of the high abundance of upregulated genes encoding for transcription  
230 factors and transcription regulators in SPZ<sub>EJ</sub>, suggest that both mitochondrial and  
231 cytoplasmic translation activity occurs during the differentiation and maturation of  
232 spermatozoa.

233 The metabolism interactome showed 379 proteins and 821 connections divided into  
234 fifteen subclusters, from which those corresponding to glycolysis/gluconeogenesis, pentose  
235 phosphate pathway (PPP) and sphingolipid, galactose, glycogen and glutathione  
236 metabolism, were the most upregulated in SPZ<sub>EJ</sub> (**Figure 3B**). Further mapping of the 76  
237 DEGs coding for enzymes involved in respiratory pathways indicated that most of the  
238 genes of the tricarboxylic acid (TCA) cycle, as well as three genes coding for specific  
239 enzymes of gluconeogenesis, such as phosphoenol-pyruvate carboxykinase (*pck2*), fructose  
240 1,6-bisphosphatase (*fbp1*) and glucose 6-phosphatase (*g6pc*), were downregulated or not  
241 differentially expressed in SPZ<sub>EJ</sub> (**Figure 3-Supplement 1**). In contrast, most of the  
242 glycolytic enzyme-encoding genes, including the two key enzymes hexokinase-1 (*hk1*) and  
243 pyruvate kinase (*pkm*), as well as many of the genes coding for enzymes catalyzing  
244 oxidative phosphorylation (OXPHOS), were upregulated (**Figure 3-Supplement 1**),  
245 suggesting that both glycolysis and OXPHOS are possibly important pathways for ATP  
246 generation in seabream spermatozoa.

247

### 248 **The PDGF and GnRHR signaling pathways are upregulated in SPZ<sub>EJ</sub>**

249 In order to identify signaling pathways enriched in SPZ<sub>EJ</sub>, pathway analysis was done for  
250 the 7,287 DEGs using the PANTHER classification system (*Mi et al., 2019*). The analysis  
251 identified a total of 960 transcripts belonging to 37 different signaling pathways, including  
252 13 receptor pathways (**Figure 4A**). Highly enriched and significant pathways were integrin,  
253 epidermal growth factor receptor (EGFR), fibroblast growth factor (FGF), cholecystokinin  
254 receptor (CCKR) and inflammation mediated by chemokine and cytokine. These results  
255 possibly reflect the activation of mechanisms for actin remodeling (*Breitbart et al., 2011*),

256 the acquisition and regulation of motility and chemotaxis (*Caballero-Campo et al., 2014;*  
257 *Tan and Thomas, 2015; Zhou et al., 2015; Saucedo et al., 2018*), and the formation of an  
258 active network of proteins prior to fertilization crucial for the sperm-egg fusion (*Chen et*  
259 *al., 1999; Frolikova et al., 2016*), during the differentiation and maturation of spermatozoa.

260 However, amongst the most dominant pathways in terms of number of genes  
261 identified and lowest *p*-values were the GnRHR and PDGF signaling pathways.  
262 Hierarchical clustering heatmaps showed that most of the genes related to these two  
263 pathways encoding for receptors, kinases, transcription factors or calcium binding proteins  
264 were upregulated, such as the Pdgf receptor b (*pdgfrb*), phosphatidylinositol 4,5-  
265 bisphosphate 3-kinase (*pik3*), nuclear factor kappaB subunit p65 (*rela*), NF-kappa-B  
266 inhibitor alpha (*nfkbia*), signal transducer and activator of transcription 1 (*stat1*), GTPase  
267 Kras (*kras*), RAF proto-oncogene Ser/Thr-protein kinase (*raf1*), mitogen-activated protein  
268 kinase 1 (*mapk1*), dual specificity mitogen-activated protein kinase kinase 2 (*map2k2*) or  
269 transcription factor AP-1 (*jun*) in the PDGF pathway (**Figure 4B**), and adenylate cyclase  
270 type 7 (*adcy7*), cAMP-dependent protein kinase catalytic subunit PRKX (*prkx*), c-AMP-  
271 dependent transcription factor ATF-4 (*atf4*), guanine nucleotide-binding protein G(q)  
272 subunit alpha (*gnaq*), inositol 1,4,5-trisphosphate receptor type 1 and 2 (*itpr1* and *itpr2*),  
273 calmodulin-1 (*calm1*), protein kinase beta, delta and epsilon (*prkcb*, *prkcd* and *prkce*) or  
274 early growth response protein 1 (*egr1*) (**Figure 4C**), in the GnRHR pathway. These data  
275 were validated by qRT-PCR for a number of genes, including three *Gnrhrs* identified in our  
276 transcriptome (*gnrhr1*, *gnrhr2* and *gnrhr3*) for which the RNA-seq did not detect  
277 significantly different expression levels (**Figure 4D**). The qRT-PCR analysis showed  
278 however that *gnrhr2* and *gnrhr3* were in fact upregulated in SPZ<sub>EJ</sub> (**Figure 4D**) Altogether,  
279 these data suggest the activation of the GnRHR and PDGF signaling pathways during  
280 seabream spermiogenesis.

281

## 282 **Seabream *Gnrh* and *Pdgf* paralogs are sequentially expressed in the ETD epithelia**

283 Since seabream SPZ<sub>EJ</sub> show elevated expression of components of the GnRHR and PDGF  
284 signaling pathways, we speculated that these pathways might be involved in the maturation  
285 of the spermatozoon in the ETDs, the efferent (ED) and sperm (SD) ducts (Figure 5A and  
286 *B*), prior to ejaculation. To investigate this hypothesis, we first evaluated whether two

287 molecular forms of GnRh present in gilthead seabream, the seabream GnRh (*sbgnrh*) and  
288 salmon GnRh (*sgnrh*) (**Powell et al., 1994**), as well as different Pdgf paralogs identified in  
289 our transcriptome and in the seabream genome (*pdgfaa*, *-ab*, *-ba*, *-bb*, *-c* and *-d*), are  
290 expressed in the testis, ED and SD. *In situ* hybridization using DIG-labeled, paralog-  
291 specific riboprobes showed strong *sbgnrh* expression in SPC in the testis, whereas the  
292 expression was also prominent in nonciliated cells lining the lumen of the ED (**Figure 5C**).  
293 In contrast, *sbgnrh* transcripts were almost undetectable in the ciliated epithelial cells of the  
294 proximal and distal regions of the SD (**Figure 5C**). The *sgnrh* mRNA was also detected  
295 exclusively in testicular SPC, while a faint signal was observed in the epithelial cells from  
296 ED but not from the SD (**Figure 5-Supplement 1**). The expression of the *sbgnrh* paralog  
297 correlated with the immunostaining of GnRh peptides using an anti-GnRH antibody,  
298 confirming that the sbGnRh was produced in testicular SPC and nonciliated epithelial cells  
299 of the ED, the expression of the neuropeptide being progressively decreased along the SD  
300 (**Figure 5D**).

301 The cell localization of *pdgf* expression in testis and ETDs by *in situ* hybridization  
302 revealed distinct expression patterns of the different *pdgf* paralogs. In the testis, expression  
303 of *pdgfaa* was specific of the somatic Sertoli cells, but only when they showed embedded  
304 developing spermatogonia (SPG) (**Figure 6-Supplement 1**), while no positive signals were  
305 detected for *pdgfab* (**Figure 6A**). The *pdgfba* mRNA was detected in SPG, with much  
306 weaker signals in SPC and SPD (**Figure 6B**), whereas the transcripts of the duplicated  
307 *pdgfb* paralog, as well as those of the *pdgfd*, were also localized in SPC but they were  
308 much less abundant in SPD (**Figure 6C and D**). The expression of *pdgfc* was only detected  
309 in Leydig cells (**Figure 6-Supplement 1**). In the ETDs, *pdgfab* expression was strong in the  
310 epithelial cells of the ED, being weak in the SD (**Figure 6A**), whereas the expression of  
311 *pdgfba* and *-bb* was more intense in the luminal surface of the SD proximal and distal  
312 regions, respectively (**Figure 6B and C**). In contrast, *pdgfd* expression was low in the  
313 epithelium throughout the ETDs, but somewhat more intense in the proximal region of the  
314 SD (**Figure 6D**), while the expression of *pdgfaa* and *-c* was almost or completely  
315 undetectable (**Figure 6-Supplement 1**).

316 Taken together, these findings demonstrate the local production of GnRh peptides by  
317 the epithelial cells of the ED, as well as the sequential expression of different *pdgf* paralogs

318 along the ETDs, which would be consistent with a physiological role of these hormones  
319 during the maturation of spermatozoa.

320

### 321 **GnRH and PDGF regulate transcription and motility of immature seabream** 322 **spermatozoa**

323 To investigate the physiological state of the spermatozoa from the ED (SPZ<sub>ED</sub>), we  
324 compared their function with respect to that of SPZ<sub>EJ</sub>. Time-course monitoring of sperm  
325 motion kinetics upon seawater activation using computer-assisted sperm analysis (CASA)  
326 revealed that SPZ<sub>ED</sub> showed a reduced percentage of motility and progressivity, and an  
327 impaired curvilinear velocity (VCL), with respect the SPZ<sub>EJ</sub> (**Figure 7A**). These data  
328 therefore indicate that SPZ<sub>ED</sub> can be classified as immature gametes, which acquire full  
329 motility potential during their journey throughout the ETDs.

330 Further RT-PCR analysis showed that both SPZ<sub>ED</sub> and SPZ<sub>EJ</sub> express *gnrhr1*, *gnrhr2*  
331 and *gnrhr3* transcripts (**Figure 7B**), while expression of *pdgfra* is specific of SPZ<sub>ED</sub>, and  
332 that of *pdgfrb* is prevalent in SPZ<sub>ED</sub> and SPZ<sub>EJ</sub> (**Figure 7C**). Therefore, we tested the  
333 hypothesis that the activation of these receptors by their cognate ligands in SPZ<sub>ED</sub> may play  
334 a role in the acquisition of full motility. For this, SPZ<sub>ED</sub> were incubated with sbGnRH,  
335 sGnRH or mouse recombinant PDGF-BB (rPDGF-BB), and subsequently activated in  
336 seawater to determine changes in motility. Exposure to the three hormones significantly  
337 increased the motility, progressivity and VCL of the SPZ<sub>ED</sub>, although the positive effect of  
338 rPDGF-BB on the velocity appeared to be more persistent over time than that of sbGnRH  
339 (**Figure 7D and Figure 7-Supplement 1**). However, the stimulation of SPZ<sub>ED</sub> motility by  
340 both sbGnRH and rPDGF-BB was completely abolished by preincubation of spermatozoa  
341 with the transcription inhibitor actinomycin D or the mitochondrial translation inhibitor  
342 chloramphenicol (**Figure 7E and Figure 7-Supplement 1**), suggesting that the sbGnRH-  
343 and rPDGF-BB-mediated regulation of motility is dependent on transcription and  
344 mitochondrial translation in spermatozoa.

345 To investigate the transcription-dependent actions of sbGnRH and PDGF on the  
346 motility of SPZ<sub>ED</sub>, we employed a targeted approach by evaluating the hormone-induced  
347 changes in the expression levels of selected genes. This included genes that encode  
348 aquaporins and ion channels, receptors, components of the motile apparatus, and enzymes

349 involved in respiratory pathways, most of them regulated in SPZ<sub>EJ</sub> as indicated by RNA-seq  
350 profiling (**Figure 2 and 3**), and which control or can potentially modulate sperm motility  
351 (**Boj et al., 2015; Chauvigné et al., 2015; Alavi et al., 2019; Chauvigné et al., 2021**). The  
352 data indicated that both sbGnRH and rPDGF-BB upregulated the expression of many of  
353 these genes in SPZ<sub>ED</sub>, but not all the same genes were affected by the two hormones  
354 (**Figure 7F**). Thus, sbGnRH stimulated the expression of *aqp1ab2*, *-3a* and *-8bb*, whereas  
355 rPDGF-BB increased the amount of the same transcripts as well as those of *aqp1ab1* and *-*  
356 *11a*. In contrast, sbGnRH induced higher expression levels of different ion channels (*trpv1*,  
357 *cnga3*, *-4*, *cngb3*, *cng*, *trpm4*, *piezo1*, *kcnc4*, *kcng4*, and *kcnh1*) compared to rPDGF-BB,  
358 whereas the growth factor upregulated some of the same genes (*trpv1*, *cnga3*, *-4*, *cngb3*,  
359 and *cng*), as well as that of *lrrc8d* and *vdac1*, which were not regulated by the  
360 neuropeptide. Interestingly, all the receptors analyzed (*gnrhr1*, *gnrhr2*, *gnrhr3*, *pdgfrb*, *ccr6*  
361 and *ccr6-L*) and most of the genes related to sperm flagellar motility (*cfap43*, *-44*, *-65*,  
362 *dnah1*, *lrrc6*, *spag16*, *spef2*, and *ttll1*) were upregulated only by sbGnRH, while the  
363 rPDGF-BB exclusively increased the expression of *ttll12*. Finally, the data showed that  
364 sbGnRH stimulated the expression of several glycolytic enzymes, such as *gpi*, *pkm*, *aldoa*  
365 and *pgam1*, whereas rPDGF-BB only upregulated *pgam1* and *g6pd*, the latter enzyme  
366 catalyzing the rate-limiting step of the PPP. Many of the genes upregulated by sbGnRH or  
367 rPDGF-BB in SPZ<sub>ED</sub> *in vitro* also appeared to be upregulated in SPZ<sub>EJ</sub> *in vivo* in the RNA-  
368 seq (**Figure 7F**). However, most of the sbGnRH regulated genes coding for motility factors  
369 were downregulated in SPZ<sub>EJ</sub> (**Figure 7F**), which may reflect an early and transitory  
370 activation of these genes during the maturation of spermatozoa in the ED *in vivo*. In any  
371 case, our findings suggest that both sbGnRH and PDGF play a role in the maturation of  
372 SPZ<sub>ED</sub> through transcriptional activation of genes involved in the acquisition and  
373 maintenance of flagellar motility.

374

### 375 **Transcription-dependent regulation of sperm maturation by GnRH and PDGF is** 376 **conserved in zebrafish.**

377 To examine whether the GnRHR and PDGF signaling pathways regulating seabream sperm  
378 maturation could be conserved in teleosts from more ancestral lineages, we first localized  
379 both GnRH and PDGF expressing cells in the testis and ETDs of the zebrafish (**Figure 8A**).

380 Immunostaining and *in situ* hybridization experiments showed the expression of GnRH  
381 peptides, as well as of *pdgfaa* and *-bb* transcripts, in the epithelial cells lining the ETD  
382 (**Figure 8B-D**), thus suggesting the existence of the epithelial GnRH and PDGF signaling  
383 pathways in the ETD of zebrafish as observed in seabream.

384 To further assess whether sbGnRH and rPDGF-BB can induce sperm maturation in  
385 zebrafish, we first confirmed that spermatozoa from the ETD (SPZ<sub>ETD</sub>) show lower motility  
386 than SPZ<sub>EJ</sub> upon activation in freshwater (**Figure 8-Supplement 2**), and that they express  
387 the four GnRH (*gnrhr1*, *-r2*, *-r3* and *-r4*) and two PDGF (*pdgfra* and *-b*) receptors formerly  
388 identified in zebrafish (**Tello et al., 2008; Eberhart et al., 2008**) (**Figure 8E**). This allowed  
389 us to classify the zebrafish SPZ<sub>ETD</sub> as immature sperm cells. *In vitro* incubation of SPZ<sub>ETD</sub>  
390 with sbGnRH or rPDGF-BB before activation significantly increased the motility,  
391 progressivity and VCL of the spermatozoa, each of which was completely blocked by the  
392 addition of actinomycin D (**Figure 8F**). These data demonstrate that the transcription-  
393 dependent maturation of SPZ<sub>ETD</sub> induced by GnRH and PDGF is a conserved mechanism in  
394 teleosts.

395

## 396 Discussion

397

398 The present results reveal that the widely accepted view of virtual transcriptional silence  
399 during post-meiotic spermatozoon maturation (**Fisher et al., 2012; Grunewald et al., 2005;**  
400 **Ren et al., 2017; Freitas et al., 2020; Puga Molina et al., 2018**) is not a conserved  
401 phenomenon in vertebrates. By selecting a species of fish that does not incorporate  
402 protamines during the chromatin condensation phase of spermiogenesis and conducting  
403 transcriptome profiling and gene set enrichment analysis during spermatozoon  
404 differentiation, we uncovered novel endocrine signaling pathways in the ETD required for  
405 the acquisition of sperm motility competence. This mechanism does not rely on the delivery  
406 of specific mRNAs via epididysome-like vesicles as in mammals (**James et al., 2020**), but  
407 on *de novo* transcription and translation mechanisms occurring in maturing spermatozoa.

408 In most teleosts, sperm maturation, the phase during which non-functional gametes  
409 develop into mature spermatozoa, fully capable of vigorous motility and fertilization, is  
410 believed to occur in the ETDs (**Schulz et al., 2010**). Previous studies have shown that

411 administration of some hormones, such as progestins, androgens and gonadotropins, can  
412 increase the seminal plasma pH in the ETD, which results in the elevation of intra-sperm  
413 cAMP levels, increase hydration, or induce the secretion of sperm-immobilizing ions by the  
414 ETD epithelium (*Schulz et al., 2010; Marshall et al., 1993*). However, the cellular sources  
415 of these hormones in the ETD and their potential signal transducing effects in the maturing  
416 spermatozoa are completely unknown. In the present study, the transcriptomic analysis of  
417 the enriched signaling pathways in seabream SPZ<sub>EJ</sub> revealed the presence of conserved  
418 GnRH and PDGF endocrine pathways in the teleost ETD. These hormones are produced in  
419 a spatially distinct expression sequence by the epithelial cells lining the ED and SD in a  
420 manner that resembles the regional expression of genes in the mammalian epididymis  
421 (*Sullivan and Mieusset, 2016; Belleannée et al., 2012; Zhao et al., 2019*). Such  
422 spatiotemporal expression is thus likely an ancient signaling mechanism that evolved early  
423 in the development of ETDs in jawed vertebrates. In the present context, therefore, different  
424 paralogs of the GnRH and PDGF hormones provide a programmed sequence of paracrine  
425 signals to activate their cognate receptors in SPZ<sub>ED</sub> to transduce *de novo* transcription and  
426 translation.

427 To validate this model, we performed *in vitro* experiments to maturationally induce  
428 SPZ<sub>ED</sub> with sbGnRH or rPDGF-BB, and conducted qRT-PCR of 45 different genes, which  
429 regulate or can potentially modulate sperm motility. We further validated the maturational  
430 status of the endocrine-induced and non-induced SPZ<sub>ED</sub> in the presence and absence of the  
431 transcription inhibitor actinomycin D following seawater activation and CASA analysis.  
432 These data show that for both seabream and zebrafish, the motility of the SPZ<sub>ED</sub> only  
433 increases in the presence of hormones, and only in the absence of the transcriptional  
434 inhibitor. For seabream these experiments also demonstrate the importance of  
435 mitochondrial translation for the increase in motility, as observed in mammals (*Gur and*  
436 *Breitbart, 2006; Zhao et al., 2009; Rajamanickam et al., 2017; Zhu et al., 2019*), while  
437 the qRT-PCR data confirm the effect of the hormones on the upregulation of a suite of  
438 downstream effector genes. In this latter respect, sbGnRH and rPDGF-BB show similar  
439 regulatory induction of aquaporins and ion channels, but sbGnRH has a more potent effect  
440 on the *de novo* transcription of receptors, motility factors and some key enzymes in glucose  
441 metabolism. Since we show that the sbGnRH neuropeptide is primarily expressed in the ED

442 epithelium, the data suggest that the upregulation of sperm receptors, motility factors and  
443 genes associated with glucose metabolism is induced early in the maturational process.  
444 Conversely, the sequential expression of paralogous seabream *pdgf* receptors in separate  
445 regions of the ED and SD, indicate that regulation of aquaporins and ion channels occurs  
446 throughout the ETD.

447 The early upregulation of *pdgfrb* and *gnrhr1*, -2 and -3 by sbGnRH in SPZ<sub>ED</sub> is a  
448 clear indication of the acquisition of developmental competence, since expression of these  
449 receptors in the maturing spermatozoa assembles the signal transduction pathways capable  
450 of responding to the cognate hormones that are secreted from the somatic ETD. Thus,  
451 although region-specific gene expression is known in the mammalian epididymis (*Sullivan*  
452 *and Miousset, 2016; Belleannée et al., 2012; Zhao et al., 2019*), and GnRH receptors are  
453 known to be expressed in primate spermatozoa (*de Villiers et al., 2021*), to the best of our  
454 knowledge, the epithelial ETD endocrine transduction of receptor-mediated gene  
455 transcription of the maturing spermatozoa has not previously been reported for vertebrates.

456 Interestingly, several of the genes that are hormonally upregulated in the maturing  
457 seabream spermatozoa have been shown to play important roles during the activation and  
458 maintenance of sperm motility. This includes aquaporins that facilitate the efflux of water  
459 for motility activation (*Boj et al., 2015; Chauvigné et al., 2013*), or the mitochondrial  
460 efflux of hydrogen peroxide for the maintenance of ATP production and flagellar  
461 contractions (*Chauvigné et al., 2015, 2021*). Other upregulated genes encode ion channels  
462 involved in sperm motility, such as *Trpv1* (*Majhi et al., 2013; Chen et al., 2020*) and  
463 different cyclic nucleotide-gated channels (*Fechner et al., 2015*), or which are potentially  
464 implicated in cell volume regulation, such as *Vdac1* (*Triphan et al., 2008*), *Lrrc8d*  
465 (*Jentsch, 2016*) and *Trpv4* (*Benfenati et al., 2011*). Each of these two processes is  
466 considered important for the activation and maintenance of sperm motility in marine  
467 teleosts (*Boj et al., 2015; Alavi et al., 2019*). The motility factors upregulated by sbGnRH  
468 were flagellar proteins involved in sperm flagellum axoneme organization and function  
469 (*Cfap*, *Spag16*, *Spef2*, *Ttll1* and *Ttll12*) and dynein motor proteins (*Dnah1*, *Lrcc6*), which  
470 are likely necessary for flagellar function in piscine spermatozoa as in mammals (*Vogel et*  
471 *al., 2010, Sironen et al., 2011; Dzyuba and Cosson, 2014; Feng et al., 2020; Wu et al.,*  
472 *2021*). The sbGnRH also activated the expression of *Ropn1* early in the maturation process,



473 which is an axonemal protein that plays a role in PKA-dependent signaling cascades  
474 required for spermatozoon capacitation (*Fiedler et al., 2013*). These findings therefore  
475 reinforce the notion that GnRH and PDGF signaling from the ETD epithelium plays a  
476 paracrine role to specifically induce the maturational expression of genes required for the  
477 activation and prolongation of sperm motility in the external aquatic environment.

478 In mammals, the vast majority of the paternal genome is packaged in protamines with  
479 transcriptional silence being coupled to heterochromatin condensation (*Sassone-Corsi,*  
480 *2002*). In anamniotes, however, protamines may not be involved in nuclear chromatin  
481 condensation, or are completely lacking from the spermatozoon nucleus as in the species  
482 selected in the present study (*Shimizu et al., 2000; Kurtz et al., 2009; Wu et al., 2011;*  
483 *Wike et al., 2021*). In seabream and zebrafish, spermiogenic nuclear chromatin  
484 condensation occurs without the replacement of the somatic-like and H1-family linker  
485 histones, so that they retain the nucleosome organisation with their nuclei remaining less  
486 condensed than those of species that incorporate protamines (*Kurtz et al., 2009; Wike et al.,*  
487 *2021; Saperas et al., 1993*). This is due to the absence of a second phase of spermiogenic  
488 chromatin condensation, which occurs when histones are displaced by SNBPs or  
489 protamines (*Kurtz et al., 2009; Saperas et al., 1993*). In such cases, and indeed in a highly  
490 diverse range of species, including invertebrates, the first chromatin condensation transition  
491 is also characterized by low level acetylation that is not related to histone replacement  
492 (*Kurtz et al., 2007, 2009*). It seems plausible that the *de novo* transcription observed for the  
493 maturing spermatozoa of seabream in the present study, may therefore occur during this  
494 phase. In any event, the regional signaling of the ETD appears to be conserved in the  
495 epididymis of amniotes, but not the spermatozoon transcription. Future studies should  
496 investigate the chromatin architecture reorganization and epigenetic marks in teleost SPZ<sub>ED</sub>  
497 that allow transcription and translation at this stage.

498 In summary, using a combination of transcriptional profiling, immunolocalization, *in*  
499 *situ* hybridization, and *in vitro* induction and inhibition experiments of sperm motility, we  
500 uncover novel endocrine signaling pathways in the ETD epithelium that transduce the *de*  
501 *novo* transcription of gametic effector genes required for fish sperm maturation. The  
502 experiments confirmed that the requirement of mitochondrial translation for the acquisition  
503 of full sperm motility is conserved between amniotes and anamniotes, but that

504 transcriptional silence of post-meiotic spermatozoa is not a pan vertebrate phenomenon. In  
505 fishes, *de novo* transcriptional activation induced by soma to gamete signal transduction  
506 pathways is necessary for the acquisition of fertility competence.

507

## 508 **Materials and Methods**

509

### 510 **Animals and sample collection**

511 Adult gilthead seabream males were raised in captivity at Institut de Recerca i Tecnologia  
512 Agroalimentàries (IRTA) aquaculture facilities in San Carlos de la Rápita (Tarragona,  
513 Spain) and maintained in the laboratory as previously described (*Chauvigné et al., 2013*).  
514 Samples of testis and SPZ<sub>EJ</sub> were obtained from males during the natural reproductive  
515 season (November-February) as previously described (*Chauvigné et al., 2013*), whereas the  
516 SPZ<sub>ED</sub> was extracted with a micropipette after an incision in the dorsal region of the  
517 dissected testis close to the SD. Zebrafish were obtained from the PRBB Animal Facility  
518 (Barcelona, Spain) and kept at 28°C with 14-hour light and 10-hour dark cycle and fed  
519 daily with dry small granular pellets (Zebrafish Management Ltd) and newly hatched brine  
520 shrimp *Artemia franciscana*. To collect SPZ<sub>EJ</sub>, males were anaesthetised with 100 ppm 2-  
521 phenoxyethanol and euthanized, and the testis mixed with 30 µl of non-activating SS300  
522 solution (in mg/ml: 8.15 NaCl, 0.67 KCl, 0.11 CaCl<sub>2</sub>, 0.12 MgSO<sub>4</sub>, 0.18 glucose, 2.42  
523 Tris-Cl pH 8.0; 300 mOsm) (*Chauvigné et al., 2021*). Subsequently, the testis was mixed  
524 with 30 µl of fresh SS300 solution and slightly crushed using a micropipette to isolate  
525 SPZ<sub>ETD</sub>.

526 Procedures relating to the care and use of animals and sample collection were  
527 approved by the Ethics Committee (EC) of Institut de Recerca i Tecnologia  
528 Agroalimentàries (IRTA) and Universitat Autònoma de Barcelona (UAB), following the  
529 International Guiding Principles for Research Involving Animals (EU 2010/63).

530

### 531 **Cell cytometry and FACS**

532 Testis samples (~30 mg) employed for FACS were collected from seabream males  
533 showing >80% of motile and progressive spermatozoa, and more than 2 min of motility  
534 duration. Biopsies were cut into small pieces of ~1 g and treated with 0.2% collagenase

535 (Merck type 1A) for 1 h under agitation in non-activating medium (NAM; in mg/ml: 3.5  
536 NaCl, 0.11 KCl, 1.23 MgCl<sub>2</sub>, 0.39 CaCl<sub>2</sub>, 1.68 NaHCO<sub>3</sub>, 0.08 glucose, 1 bovine serum  
537 albumine [BSA], pH 7.7; 280 mOsm) (5I) supplemented with 200 µg/mL  
538 penicillin/streptomycin (Life Technologies Corp.). Samples were centrifuged at 200 × g for  
539 1 min to remove cell aggregates, and the supernatant centrifuged again at 400 × g for 1 min  
540 to enrich in haploid cells. The cells were centrifuged at 400 × g for 5 min and the pellet  
541 resuspended in 1 ml NAM. The concentration of cells was determined by light microscopy  
542 and the ISASv1 software (Proiser), and this was adjusted to 150 × 10<sup>6</sup> cells/ml. Cells were  
543 then stained with 200 nM of a solution of SYBR Green I (SGI) fluorescent nucleic acid  
544 stain (Molecular Probes, Life Technologies Corp.) for 45-60 min in the dark at room  
545 temperature, just prior to flow cytometry.

546 FACS was performed with a MoFlo XDP cell sorter (Beckman Coulter) equipped  
547 with three lasers (blue solid state of 488nm, red diode of 635nm, and argon ion UV laser of  
548 351nm). Sterilized PBS served as the sheath fluid. The sorter was set in 4-way purify sort  
549 mode and with a flow sorting rate of ~1500 events/s. The sorted population of HGC was  
550 collected in 4 ml of NAM in 15 ml tubes and centrifuged at 200 × g for 15 min. The  
551 resulting pellet was resuspended in 100 µl of NAM to obtain aliquots of 3 to 5 × 10<sup>6</sup> cells,  
552 which were centrifuged again at 200 × g and frozen in liquid nitrogen and stored at -80°C.

553

#### 554 **RNA extraction, library preparation, and sequencing**

555 Total RNA from HGC (3 × 10<sup>7</sup> cells) and SPZ<sub>EJ</sub> (3-30 × 10<sup>7</sup> cells) was extracted with the  
556 RNeasy Plus Mini Kit (Qiagen), and the purity and concentration of the extracted RNA was  
557 evaluated with the Agilent 2100 Bioanalyzer (Agilent Technologies). Four unstranded  
558 RNA libraries (replicates) for low-input RNA were constructed for each of the HGC and  
559 SPZ<sub>EJ</sub> groups; each replicate being a pool of cells collected from three different males. The  
560 libraries from the total RNA were prepared following the SMARTseq2 protocol for low-  
561 input RNA (*Picelli et al., 2014*) with some modifications. Briefly, reverse transcription  
562 with 2 ng RNA was performed using SuperScript II (Invitrogen) in the presence of oligo-  
563 dT30VN (1 µM; 5'-AAGCAGTGGTATCAACGCAGAGTACT<sub>30</sub>VN-3'), template-  
564 switching oligonucleotides (1 µM) and betaine (1 M). The cDNA was amplified using the  
565 KAPA Hifi Hotstart ReadyMix (Merck), 100 nM ISPCR primer (5'-

566 AAGCAGTGGTATCAACGCAGAGT-3') and 12 cycles of amplification. Following  
567 purification with Agencourt Ampure XP beads (1:1 ratio; Beckmann Coulter), product size  
568 distribution and quantity were assessed on a Bioanalyzer High Sensitivity DNA Kit  
569 (Agilent). The amplified cDNA (200 ng) was fragmented for 10 min at 55 °C using  
570 Nextera® XT (Illumina) and amplified for 12 cycles with indexed Nextera® PCR primers.  
571 The library was purified twice with Agencourt Ampure XP beads (0.8:1 ratio) and  
572 quantified on a Bioanalyzer using a High Sensitivity DNA Kit.

573 The libraries were sequenced on HiSeq2500 (Illumina) in paired-end mode with a  
574 read length of 2 x 76bp using TruSeq SBS Kit v4. We generated more than 30 million  
575 paired-end reads for each sample in a fraction of a sequencing v4 flow cell lane, following  
576 the manufacturer's protocol. Image analysis, base calling and quality scoring of the run  
577 were processed using the manufacturer's software Real Time Analysis (RTA 1.18.66.3) and  
578 followed by generation of FASTQ sequence files by CASAVA 1.8.

579

## 580 **Genome annotation**

581 To improve the gilthead seabream reference genome (*Pauletto et al., 2018*) for the  
582 differential expression analysis, the genome was reannotated, and a *de novo* transcriptome  
583 assembly was generated from which those transcripts not present in the genome assembly  
584 were added to the analysis.

585 **Genome reannotation.** Repeats present in the seabream genome assembly were  
586 annotated with RepeatMasker v4-0-7 (<http://www.repeatmasker.org>) using the zebrafish  
587 repeat library included in RepeatMasker. The gene annotation was obtained by combining  
588 transcript alignments, protein alignments and *ab initio* gene predictions. First, the RNA-seq  
589 reads were aligned to the genome with STAR v-2.5.3a (*Dobin et al., 2013*). Subsequently,  
590 transcript models were generated using Stringtie v1.0.4 (*Pertea et al., 2015*) and PASA  
591 assemblies were produced with PASA v2.0.2 (*Haas et al., 2008*) by adding also the  
592 114,155 *S. aurata* ESTs present in NCBI (October 2017). Secondly, the complete  
593 Actinopterygii proteomes were downloaded from Uniprot in October 2017 and aligned to  
594 the genome using Spaln v2.4.7 (*Iwata and Gotoh, 2012*). *Ab initio* gene predictions were  
595 performed on the repeat masked assembly with three different programs: GeneID v1.4  
596 (*Parra et al., 2000*), Augustus v3.2.3 (*Stanke et al., 2006*) and Genemark-ES v2.3e

597 (*Lomsadze et al., 2014*) with and without incorporating evidence from the RNA-seq data.  
598 The gene predictors were run with trained parameters for human except Genemark that runs  
599 on a self-trained manner. Finally, all the data was combined into consensus CDS models  
600 using EvidenceModeler-1.1.1 (*Haas et al., 2008*). Additionally, UTRs and alternative  
601 splicing forms were annotated through two rounds of PASA annotation updates. Functional  
602 annotation was performed on the annotated proteins with Blast2go (*Conesa et al., 2005*),  
603 using Blastp (*Altschul et al., 1990*) search against the nr database (March 2018) and  
604 Interproscan (*Jones et al., 2014*) to detect protein domains on the annotated proteins.

605 The annotation of ncRNAs was carried out by the following steps. First, the program  
606 cmsearch v1.1 (*Cui et al., 2016*) included in the Infernal software (*Nawrocki et al., 2015*)  
607 was run against the RFAM v12.0 database of RNA families (*Nawrocki et al., 2015*). The  
608 tRNAscan-SE v1.23 (*Chan and Lowe, 2019*) was also run in order to detect the transfer  
609 RNA genes present in the genome assembly. To detect the lncRNAs we selected those  
610 Pasa-assemblies that had not been included into the annotation of protein-coding genes in  
611 order to get all those expressed genes that were not translated into a protein. Finally, those  
612 PASA-assemblies without protein-coding gene annotation that were longer than 200 bp and  
613 whose length was not covered at least in an 80% by a small ncRNA were incorporated into  
614 the ncRNA annotation as lncRNAs. The resulting transcripts were clustered into genes  
615 using shared splice sites or significant sequence overlap as criteria for designation as the  
616 same gene.

617 ***Complementing the annotation with de novo assembled transcripts.*** The RNA-seq  
618 reads were assembled with Trinity v2.2.0 (*Grabherr et al., 2011*) allowing for trimming  
619 and normalization of the reads. Next, Rapclust v0.1 (*Trapnell et al., 2013*) was run, in  
620 which the process of pseudoalignment was first performed with Sailfish v0.10.0 (*Li et al.,*  
621 ***2010***), and then Rapclust was used to cluster the assembled sequences into contained  
622 isoforms in order to reduce redundancy and to cluster together all the isoforms that are  
623 likely to belong to the same gene. For evaluation of the resulting transcriptomes we  
624 estimated their completeness with BUSCO v3.0.2 (*Simao et al., 2015*) using an  
625 Actinopterygii specific dataset of 4584 genes. After obtaining the reference transcriptome,  
626 open reading frames (ORFs) were annotated in the assembled transcripts with Transdecoder  
627 (*Haas et al., 2013*) and functional annotation was performed on the annotated proteins with

628 Blast2GO, as described above. Finally, the assembled transcripts were mapped against the  
629 seabream reference genome assembly with GMAP (*Wu et al., 2005*). Those transcripts for  
630 which less than 50% of their length aligned to the genome, and with a complete ORF and  
631 functional annotation, were added to the reference genome as separate annotated contigs.

632

### 633 **Differential expression analysis**

634 RNA-seq reads were mapped against the improved version of the seabream reference  
635 genome with STAR v2.5.3a using ENCODE parameters for long RNAs. Genes were  
636 quantified with RSEM v1.3.0 (*Li and Dewey, 2011*) using the improved annotation. Sample  
637 similarities were inspected with a PCA using the top 500 most variable genes and the 'rlog'  
638 transformation of the counts. Differential expression analysis was performed with DESeq2  
639 v1.18 (*Love et al., 2014*) with default options, and genes with a false discovery rate (FDR)  
640 < 1% were considered significant. Heatmaps with the 'rlog' transformed counts of the  
641 DEGs were carried out with the 'pheatmap' R package. Venn diagrams and volcano plots  
642 were performed with the 'VenDiagramm' R package and 'ggplot2' R package,  
643 respectively.

644

### 645 **Gene classification, Ontology, and Pathway Analysis of DEGs**

646 The GO enrichment of DEGs and signaling pathway analyses were performed using the  
647 PANTHER v14.1 Classification System and analysis tools (<http://www.pantherdb.org/>).  
648 GO terms and pathways with FDR < 0.05% were considered significant. Scattered plot of  
649 pathway analysis was carried out with 'ggplot2' R package. Functional categories  
650 classification were also done manually using the Uniprot database  
651 (<https://www.uniprot.org/>) and QuickGO browser (<http://www.ebi.ac.uk/QuickGO>).  
652 Interactome analyses were conducted using the STRING database v11.0b (32) with a high-  
653 confidence interaction score (0.9), and plots were performed using Cytoscape v3.8.2  
654 (<https://cytoscape.org/>).

655

### 656 **In situ hybridization**

657 Samples of seabream and zebrafish testis and ETDs were fixed in 4% paraformaldehyde  
658 (PFA) prepared in phosphate buffer saline (PBS: 137 mM NaCl, 2.7 mM KCl, 100 mM

659 Na<sub>2</sub>HPO<sub>4</sub>, 2 mM KH<sub>2</sub>PO<sub>4</sub>, pH 7.4) overnight at 4°C. Samples were washed in PBS,  
660 dehydrated with increasing concentration of ethanol (50%, 70%, 95%, 100%) and xylene  
661 (100 %), and embedded in Paraplast<sup>®</sup> (Merck). In situ hybridization was performed on 7-  
662 µm thick sections using digoxigenin-incorporated cRNA probes synthesized with SP6 and  
663 T7 RNA polymerases using the DIG RNA Labeling Mix (Merck 11277073910). Probes  
664 were specific for each target mRNA and did not share more than 35% identity between  
665 related transcripts (*Supplementary file 1*). Hybridization was performed at 45°C overnight  
666 with probe concentration at 2.5 µg/ml (*sbgnrh* and *sapdgfaa*, *-ba* and *-bb*) or 5 µg/ml  
667 (*sgnrh*, *sapdgfab*, *-c* and *-d*, and *drpdgfaa*, *-ab* and *-bb*). The post-hybridization washing  
668 included a first wash in 50% formamide in 2 x SSC at 45°C for 30 min, followed by two  
669 washes in 2 x SSC for 10 min at 45°C, and a final wash in 0.2 x SSC at 50°C. After  
670 blocking in TBST with 0.5% BSA, hybridized riboprobes were detected with alkaline  
671 phosphatase coupled rabbit anti-digoxigenin antibody (1:500; Merck 11093274910) for two  
672 hours at room temperature, and subsequent chromogenic revelation (NBT/BCIP Stock  
673 solution, Merck 11681451001). The reaction was stopped in distilled water and slides were  
674 mounted with Fluoromount<sup>™</sup> aqueous mounting medium (Merck F4680).

675

#### 676 **Immunofluorescence microscopy**

677 Sorted germ cells and SPZ<sub>EJ</sub> were processed as described previously (51, 71) and attached  
678 to UltraStick/UltraFrost Adhesion slides (Electron Microscopy Sciences). Samples were  
679 fixed in 4% PFA in PBS for 15 min before antigen retrieval in three consecutive 5-min  
680 incubations with boiling citrate (10 mM at pH 6), followed by triton X-100 (0.2% in PBS)  
681 for 15 min. After blocking for one hour in PBST with 5% normal goat serum (Merck  
682 G9023) and 0.1% BSA, antibodies were applied overnight at 4°C in a humidified chamber.  
683 The primary antibodies were α-tubulin (Merck T9026; 1:1,000), H3K9ac (Abcam ab4441;  
684 1:1000), and Spo11 (Santa Cruz Biotechnology sc-33146; 1:1000). Anti-mouse or anti-  
685 rabbit IgG coupled with Alexa-555 (A-21422, Invitrogen, and AP510C, Merck,  
686 respectively) were applied for one hour at room temperature and cells were counterstained  
687 with 4',6-diamidino-2-phenylindole dihydrochloride (DAPI; Merck G8294; 1:3000) before  
688 mounting with Fluoromount<sup>™</sup>.

689 The biopsies of testis and ETDs were fixed and processed as previously described  
690 (*Chauvigné et al., 2013*). Sections were permeabilized with 0.2% Triton X100 in PBS for  
691 15 min, blocked with 5% normal goat serum, and subsequently incubated with affinity  
692 purified rabbit anti-GnRH (Merck, G8294, 1:400) in PBS+0.1% BSA overnight at 4°C.  
693 After washing, sections were incubated with a sheep anti-rabbit IgG antibody, Cy3  
694 conjugate (Merck AP510C) for 2 h, the nuclei counterstained with DAPI (1:3000) for 3  
695 min, and finally mounted with Fluoromount™.

696

### 697 **Sperm motility assays and in vitro incubation of SPZ<sub>ED</sub>**

698 Freshly collected seabream SPZ<sub>EJ</sub> and SPZ<sub>ED</sub> were diluted 1:100 in NAM, whereas  
699 zebrafish SPZ<sub>EJ</sub> and SPZ<sub>ETD</sub> were no further diluted. Spermatozoa concentration and kinetic  
700 parameters were determined by computer-assisted sperm analysis (CASA) using the  
701 Integrated Semen Analysis System (ISASv1, Proiser) software as previously described  
702 (*Chauvigné et al., 2013, 2021*). The sperm kinetics analyses were run in triplicate  
703 (technical replicates) for each ejaculate. For seabream, the analyses were carried out on 4-8  
704 different males (one ejaculate per male), whereas for zebrafish the analyses were done on  
705 three to four pools of three males each.

706 The SPZ<sub>ED</sub> ( $10^7$  or  $10^9$  cells/ml for zebrafish and seabream, respectively) were  
707 incubated in vitro in NAM (seabream) or modified SS300 medium for zebrafish (65 mM  
708 KCl, 62.5 mM NaCl, 2.35 mM CaCl<sub>2</sub>, 1 mM MgSO<sub>4</sub>, 6.5 mM MgCl<sub>2</sub>, 10 mM NaHCO<sub>3</sub>, 7  
709 mM glucose, 30 mM Hepes-KOH pH 7.9, 0.015 mM BSA, pH 7.7; 330 Osm) in the  
710 presence of 100 nM of sbGnRH or sGnRH (Bachem, 4030832 and 4013835, respectively),  
711 40 nM or rPDGF-BB (ThermoFisher Scientific PMG0044), or hormone vehicles (0.5% of  
712 water, or 0.8 mM acetic acid solution; controls). The incubations were carried out for 16-20  
713 h at 16°C in a temperature-controlled incubator. After the incubation period, the sperm  
714 kinetic parameters were determined by CASA as above, and a subsample of SPZ<sub>ED</sub> was  
715 frozen in liquid nitrogen and stored at -80°C until further RNA extraction. The effect of  
716 actinomycin D (Merck A9415) and chloramphenicol (Merck C1919) on motility was tested  
717 by preincubation of SPZ<sub>ED</sub> with 100 µg/ml of the drugs for 1 h before addition of the  
718 hormones.

719



## 720 **Gene Expression Analyses**

721 RT-PCR and qRT-PCR were carried out as previously described (*Chauvigné et al., 2013,*  
722 **2014**), except that in this case the cDNA was synthesized from 1 µg (testis) or 13-20 ng  
723 (spermatozoa) of RNA using the AccuScript High-Fidelity 1st Strand cDNA Synthesis Kit  
724 (Agilent 200820) following the manufacturer's instructions. For qRT-PCR, relative gene  
725 expression levels with respect to HGC or vehicle-treated SPZ<sub>ED</sub> were determined by the 2<sup>-</sup>  
726  $\Delta\Delta C_t$  method, using glutathione-specific gamma-glutamylcyclotransferase 1 (*chac1*) or beta-  
727 actin (*bactin*) as reference genes. The analyses were done on three cDNAs synthesized  
728 from three different pools of three animals each, or on three to five cDNAs from different  
729 animals, using technical duplicates. Primer sequences are listed in *Supplementary file 2*.

730

## 731 **Statistical analysis**

732 Comparisons between two independent groups were made by the two-tailed unpaired  
733 Student's *t*-test. The statistical significance among multiple groups was analyzed by one-  
734 way ANOVA, followed by the Tukey's multiple comparison test, or by the non-parametric  
735 Kruskal-Wallis test and further Dunn's test for nonparametric post hoc comparisons, as  
736 appropriate. Percentages were square root transformed previous analyses. Statistical  
737 analyses were carried out using the SigmaPlot software v12.0 (Systat Software Inc.) and  
738 GraphPad Prism v9.1.2 (226) (GraphPad Software). In all cases, statistical significance was  
739 defined as  $P < 0.05$  (\*),  $P < 0.01$  (\*\*), or  $P < 0.001$  (\*\*\*)).

740

## 741 **Data availability**

742 The RNA-seq datasets generated in this study have been submitted to Gene Expression  
743 Omnibus (GEO) database at the National Center for Biotechnology Information (NCBI)  
744 under accession no. GSE173088. Reannotation data are available at  
745 <https://denovo.cnag.cat/Saurata>. All other data generated or analysed during this study are  
746 included in the manuscript and supporting files.

747

## 748 **Acknowledgments**

749 This work was supported by the Spanish Ministry of Science and Innovation (MICIN)  
750 Grant AGL2016-76802-R (to J.C.). F.C. and J.C.A. were supported, respectively, by the

751 “Ramon y Cajal” programe (RYC-2015-17103) and a predoctoral (BES-2017-080778)  
752 contract from Spanish MICIN. A.E.C. was funded by ISCIII/ MICIN (PT17/0009/0019)  
753 and co-funded by FEDER, whereas R.N.F. was supported by the University of Bergen,  
754 Norway. We also acknowledge support of the Spanish MICIN through the Instituto de  
755 Salud Carlos III and to the EMBL partnership, the Centro de Excelencia Severo Ochoa, the  
756 Generalitat de Catalunya through the CERCA Programme, Departament de Salut and  
757 Departament d’Empresa i Coneixement, and funds from the European Regional  
758 Development Fund (Programa Operatiu FEDER de Catalunya 2014-2020) cofinanced by  
759 the Spanish MICIN (Programa Operativo FEDER Plurirregional de España (POPE) 2014-  
760 2020).

761

## 762 **Competing interests**

763 The authors declare that no competing interests exist.

764

## 765 **References**

766

767 Alavi SMH, Cosson J, Bondarenko O, Linhart O. 2019. Sperm motility in fishes: (III) diversity of  
768 regulatory signals from membrane to the axoneme. *Theriogenology* **136**:143-165. DOI:

769 <https://doi.org/10.1016/j.theriogenology.2019.06.038>

770 Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool.

771 *Journal of Molecular Biology* **215**: 403-410. DOI: <https://doi.org/10.1016/s0022->

772 [2836\(05\)80360-2](https://doi.org/10.1016/s0022-2836(05)80360-2)

773 Ausió J. 1999. Histone H1 and evolution of sperm nuclear basic proteins. *Journal of Biological*

774 *Chemistry* **274**: 31115-31118. DOI: <https://doi.org/10.1074/jbc.274.44.31115>

775 Belleannée C, Thimon V, Sullivan R. 2012. Region-specific gene expression in the epididymis. *Cell*

776 *Tissue Research* **349**: 717-731. DOI: <https://doi.org/10.1007/s00441-012-1381-0>

777 Benfenati V, Caprini M, Dovizio M, Mylonakou MN, Ferroni S, Ottersen OP, Amiry-Moghaddam

778 M. 2011. An aquaporin-4/transient receptor potential vanilloid 4 (AQP4/TRPV4) complex is

779 essential for cell-volume control in astrocytes. *PNAS* **108**: 2563-2568. DOI:

780 <https://doi.org/10.1073/pnas.1012867108>

781 Boj M, Chauvigné F, Cerdà J. 2015. Coordinated action of aquaporins regulates sperm motility in a

782 marine teleost. *Biology of Reproduction* **93(2)**: 40. DOI:

783 <https://doi.org/10.1095/biolreprod.115.131524>

- 784 Breitbart H, Etkovitz N. 2011. Role and regulation of EGFR in actin remodeling in sperm  
785 capacitation and the acrosome reaction. *Asian Journal of Andrology* **13**: 106-110. DOI:  
786 <https://doi.org/10.1038/aja.2010.78>
- 787 Caballero-Campo P, Buffone MG, Benencia F, Conejo-García JR, Rinaudo PF, Gerton GL. 2014. A  
788 role for the chemokine receptor CCR6 in mammalian sperm motility and chemotaxis. *Journal of*  
789 *Cell Physiology* **229**: 68-78. DOI: <https://doi.org/10.1002/jcp.24418>
- 790 Chan PP, Lowe TM. 2019. tRNAscan-SE: searching for tRNA genes in genomic sequences.  
791 *Methods in Molecular Biology* **1962**: 1-14. DOI: [https://doi.org/10.1007/978-1-4939-9173-](https://doi.org/10.1007/978-1-4939-9173-0_1)  
792 [0\\_1](https://doi.org/10.1007/978-1-4939-9173-0_1)
- 793 Chauvigné F, Boj M, Vilella S, Finn RN, Cerdà J. 2013. Subcellular localization of selectively  
794 permeable aquaporins in the male germ line of a marine teleost reveals spatial redistribution in  
795 activated spermatozoa. *Biology of Reproduction* **89(2)**: 37. DOI:  
796 <https://doi.org/10.1095/biolreprod.113.110783>
- 797 Chauvigné F, Zapater C, Gasol JM, Cerdà J. 2014. Germ-line activation of the luteinizing hormone  
798 receptor directly drives spermiogenesis in a nonmammalian vertebrate. *PNAS* **111**: 1427-1432.  
799 DOI: <https://doi.org/10.1073/pnas.1317838111>
- 800 Chauvigné F, Boj M, Finn RN, Cerdà J. 2015. Mitochondrial aquaporin-8-mediated hydrogen  
801 peroxide transport is essential for teleost spermatozoon motility. *Scientific Reports* **5**: 7789.  
802 DOI: <https://doi.org/10.1038/srep07789>
- 803 Chauvigné F, Ducat C, Ferré A, Hansen T, Carrascal M, Abián J., Finn RN, Cerdà J. 2021. A  
804 multiplier peroxiporin signal transduction pathway powers piscine spermatozoa. *PNAS* **18(10)**:  
805 e2019346118. DOI: <https://doi.org/10.1073/pnas.2019346118>
- 806 Chen MS, Tung KS, Coonrod SA, Takahashi Y, Bigler D, Chang A, Yamashita Y, Kincade PW,  
807 Herr JC, White JM. 1999. Role of the integrin-associated protein CD9 in binding between sperm  
808 ADAM 2 and the egg integrin alpha6beta1: implications for murine fertilization. *PNAS* **96**:  
809 11830-11835. DOI: <https://doi.org/10.1073/pnas.96.21.11830>
- 810 Chen Y, Wang H, Wang F, Chen C, Zhang P, Song D, Luo T, Xu H, Zeng X. 2020. Sperm motility  
811 modulated by Trpv1 regulates zebrafish fertilization. *Theriogenology* **151**:41-51. DOI:  
812 <https://doi.org/10.1016/j.theriogenology.2020.03.032>
- 813 Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M. 2005. Blast2GO: a universal  
814 tool for annotation, visualization and analysis in functional genomics research.  
815 *Bioinformatics* **21**: 3674-3676. DOI: <https://doi.org/10.1093/bioinformatics/bti610>
- 816 Corral-Vazquez C, Blanco J, Aiese Cigliano R, Sarrate Z, Rivera-Egea R, Vidal F, Garrido N, Daub  
817 C, Anton E. 2021. The RNA content of human sperm reflects prior events in spermatogenesis

- 818 and potential post-fertilization effects. *Molecular Human Reproduction* **27**: gaab035. DOI:  
819 <https://doi.org/10.1093/molehr/gaab035>
- 820 Cui X, Lu Z, Wang S, Jing-Yan Wang J, Gao X. 2016. CMsearch: simultaneous exploration of  
821 protein sequence space and structure space improves not only protein homology detection but  
822 also protein structure prediction. *Bioinformatics* **32**: i332-i340. DOI:  
823 <https://doi.org/10.1093/bioinformatics/btw271>
- 824 de Kretzer DM, Loveland KL, Meinhardt A, Somorangkir D, Wreford N. 1998. Spermatogenesis.  
825 *Human Reproduction* **13**: 1-8. DOI: [https://doi.org/10.1093/humrep/13.suppl\\_1.1](https://doi.org/10.1093/humrep/13.suppl_1.1)
- 826 de Villiers C, van der Horst G, Chauke C, Megwebu Z. 2021. The expression of type I and II  
827 gonadotropin-releasing hormone receptors transcripts in Vervet monkey (*Chlorocebus aethiops*)  
828 spermatozoa. *General and Comparative Endocrinology* **310**: 113819. DOI:  
829 <https://doi.org/10.1016/j.ygcen.2021.113819>
- 830 Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras  
831 TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**: 15-21. DOI:  
832 <https://doi.org/10.1093/bioinformatics/bts635>
- 833 Dzyuba V, Cosson J. 2014. Motility of fish spermatozoa: from external signaling to flagella  
834 response. *Reproductive Biology* **14**: 165-175. DOI: <https://doi.org/10.1016/j.repbio.2013.12.005>
- 835 Eberhart JK, He X, Swartz ME, Yan YL, Song H, Boling TC, Kunerth AK, Walker MB, Kimmel  
836 CB, Postlethwait JH. 2008. MicroRNA Mirn140 modulates Pdgf signaling during palatogenesis.  
837 *Nature Genetics* **4**: 290-298. DOI: <https://doi.org/10.1038/ng.82>
- 838 Eirín-López JM, Frehlick LJ, Ausió J. 2006. Protamines, in the footsteps of linker histone evolution.  
839 *Journal of Biological Chemistry* **281**: 1-4. DOI: <https://doi.org/10.1074/jbc.R500018200>
- 840 Fechner S, Alvarez L, Bönigk W, Müller A, Berger TK, Pascal R, Trötschel C, Poetsch A, Stölting  
841 G, Siegfried KR, Kremmer E, Seifert R, Kaupp UB. 2015. A K<sup>(+)</sup>-selective CNG channel  
842 orchestrates Ca<sup>(2+)</sup> signalling in zebrafish sperm. *Elife* **4**: e07624. DOI:  
843 <https://doi.org/10.7554/eLife.07624>
- 844 Fiedler SE, Dudiki T, Vijayaraghavan S, Carr DW. 2013. Loss of R2D2 proteins ROPN1 and  
845 ROPN1L causes defects in murine sperm motility, phosphorylation, and fibrous sheath integrity.  
846 *Biology of Reproduction* **88**: 41. DOI: <https://doi.org/10.1095/biolreprod.112.105262>
- 847 Fischer BE, Wasbrough E, Meadows LA, Randlet O, Dorus S, Karr TL, Russel S. 2012. Conserved  
848 properties of *Drosophila* and human spermatozoal mRNA repertoires. *Proceedings Royal*  
849 *Society B* **279**: 2636-2644. DOI: <https://doi.org/10.1098/rspb.2012.0153>
- 850 Freitas MJ, Patrício D, Fardilha M. 2020. Sperm signaling specificity: from sperm maturation to  
851 oocyte recognition - Tissue-Specific Cell Signaling in: Silva JV, Freitas MJ, Fardilha. Springer.

- 852 DOI: [https://doi.org/10.1007/978-3-030-44436-5\\_9](https://doi.org/10.1007/978-3-030-44436-5_9)
- 853 Frolikova M, Sebkova N, Ded L, Dvorakova-Hortova K. 2016. Characterization of CD46 and  $\beta$ 1  
854 integrin dynamics during sperm acrosome reaction. *Scientific Reports* **6**: 33714. DOI:  
855 <https://doi.org/10.1038/srep33714>
- 856 Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L,  
857 Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F,  
858 Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A. 2011. Full-length  
859 transcriptome assembly from RNA-Seq data without a reference genome. *Nature*  
860 *Biotechnology* **29**: 644-652. DOI: <https://doi.org/10.1038/nbt.1883>
- 861 Grunewald S, Paasch U, Glander HJ, Anderegg U. 2005. Mature human spermatozoa do not  
862 transcribe novel RNA. *Andrologia* **37**: 69-71. DOI: [https://doi.org/10.1111/j.1439-](https://doi.org/10.1111/j.1439-0272.2005.00656.x)  
863 [0272.2005.00656.x](https://doi.org/10.1111/j.1439-0272.2005.00656.x)
- 864 Gur Y, Breitbart H. 2006. Mammalian sperm translate nuclear-encoded proteins by mitochondrial-  
865 type ribosomes. *Genes & Development* **20**: 411-416. DOI: <https://doi.org/10.1101/gad.367606>
- 866 Haas BJ, Salzberg SL, Zhu W, Pertea M, Allen JE, Orvis J, White O, Buell CR, Wortman JR. 2008.  
867 Automated eukaryotic gene structure annotation using EVIDENCEModeler and the Program to  
868 Assemble Spliced Alignments. *Genome Biology* **9**: R7. DOI: [https://doi.org/10.1186/gb-](https://doi.org/10.1186/gb-2008-9-1-r7)  
869 [2008-9-1-r7](https://doi.org/10.1186/gb-2008-9-1-r7)
- 870 Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li  
871 B, Lieber M, MacManes MD, Ott M, Orvis J, Pochet N, Strozzi F, Weeks N, Westerman R,  
872 William T, Dewey CN, Henschel R, LeDuc RD, Friedman N, Regev A. 2013. De novo  
873 transcript sequence reconstruction from RNA-seq using the Trinity platform for reference  
874 generation and analysis. *Nature Protocols* **8**: 1494-1512. DOI:  
875 <https://doi.org/10.1038/nprot.2013.084>
- 876 Hazzouri M, Pivot-Pajot C, Faure AK, Usson Y, Pelletier R, Sèle B, Khochbin S, Rousseaux S.  
877 2000. Regulated hyperacetylation of core histones during mouse spermatogenesis: involvement  
878 of histone deacetylases. *European Journal of Cell Biology* **79**: 950-960. DOI:  
879 <https://doi.org/10.1078/0171-9335-00123>
- 880 Iwata H, Gotoh O. 2012. Benchmarking spliced alignment programs including Spaln2, an extended  
881 version of Spaln that incorporates additional species-specific features. *Nucleic Acids*  
882 *Research* **40**: e161. DOI: <https://doi.org/10.1093/nar/gks708>
- 883 James ER, Carrell DT, Aston KI, Jenkins TG, Yeste M, Salas-Huetos A. 2020. The role of the  
884 epididymis and the contribution of epididymosomes to mammalian reproduction. *International*  
885 *Journal of Molecular Sciences* **21**: 5377. DOI: <https://doi.org/10.3390/ijms21155377>

- 886 Jentsch TJ. 2016. VRACs and other ion channels and transporters in the regulation of cell volume  
887 and beyond. *Nature Reviews in Molecular Cell Biology* **17**: 293-307. DOI:  
888 <https://doi.org/10.1038/nrm.2016.29>
- 889 Jones P, Binns D, Chang HY, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A,  
890 Nuka G, Pesseat S, Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong SY, Lopez R,  
891 Hunter S. 2014. InterProScan 5: genome-scale protein function classification. *Bioinformatics*  
892 **30**: 1236-1240. DOI: <https://doi.org/10.1093/bioinformatics/btu031>
- 893 Jones RC. 2002. Evolution of the vertebrate epididymis - The Epididymis: From Molecules to  
894 Clinical Practice in: Robaire B, Hinton BT. Springer. DOI: [https://doi.org/10.1007/978-1-4615-](https://doi.org/10.1007/978-1-4615-0679-9_2)  
895 [0679-9\\_2](https://doi.org/10.1007/978-1-4615-0679-9_2) Nishimura H, L'Hernault SW. 2017. Spermatogenesis. *Current Biology* **27**: R979-  
896 R1001. DOI: <https://doi.org/10.1016/j.cub.2017.07.067>
- 897 Kurtz K, Martínez-soler F, Ausió J, Chiva M. 2007. Acetylation of histone H4 in complex structural  
898 transitions of spermiogenic chromatin. *Journal of Cellular Biochemistry* **102**: 1432-1441. DOI:  
899 <https://doi.org/10.1002/jcb.21365>
- 900 Kurtz K, Saperas N, Ausió J, Chiva M. 2009. Spermiogenic nuclear protein transitions and  
901 chromatin condensation. Proposal for an ancestral model of nuclear spermiogenesis. *Journal of*  
902 *Experimental Zoology B Molecular Development and Evolution* **312B**: 149-163. DOI:  
903 <https://doi.org/10.1002/jez.b.21271>
- 904 Li B, Ruotti V, Stewart RM, Thomson JA, Dewey CN. 2010. RNA-Seq gene expression estimation  
905 with read mapping uncertainty. *Bioinformatics* **26**: 493-500. DOI:  
906 <https://doi.org/10.1093/bioinformatics/btp692>
- 907 Li B, Dewey CN. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or  
908 without a reference genome. *BMC Bioinformatics* **12**: 323. DOI: [https://doi.org/10.1186/1471-](https://doi.org/10.1186/1471-2105-12-323)  
909 [2105-12-323](https://doi.org/10.1186/1471-2105-12-323)
- 910 Li L, Feng F, Wang Y, Guo J, Yue W. 2020. Mutational effect of human CFAP43 splice-site  
911 variant causing multiple morphological abnormalities of the sperm flagella. *Andrologia* **52**:  
912 e13575. DOI: <https://doi.org/10.1111/and.13575>
- 913 Lomsadze A, Burns PD, Borodovsky M. 2014. Integration of mapped RNA-Seq reads into  
914 automatic training of eukaryotic gene finding algorithm. *Nucleic Acids Research* **42**: e119.  
915 DOI: <https://doi.org/10.1093/nar/gku557>
- 916 Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-  
917 seq data with DESeq2. *Genome Biology* **15**: 550. DOI: [https://doi.org/10.1186/s13059-014-](https://doi.org/10.1186/s13059-014-0550-8)  
918 [0550-8](https://doi.org/10.1186/s13059-014-0550-8)
- 919 Majhi RK, Kumar A, Yadav M, Swain N, Kumari S, Saha A, Pradhan A, Goswami L, Saha S,

- 920 Samanta L, Maity A, Nayak TK, Chattopadhyay S, Rajakuberan C, Kumar A, Goswami C.  
921 2013. Thermosensitive ion channel TRPV1 is endogenously expressed in the sperm of a fresh  
922 water teleost fish (*Labeo rohita*) and regulates sperm motility. *Channels (Austin)* **7**: 483-492.  
923 DOI: <https://doi.org/10.4161/chan.25793>
- 924 Marshall WS, Bryson SE, Idler DR. 1993. Gonadotropin action on brook trout sperm duct  
925 epithelium: ion transport stimulation mediated by cAMP and Ca<sup>2+</sup>. *General and Comparative*  
926 *Endocrinology* **90**: 232-242. DOI: <https://doi.org/10.1006/gcen.1993.1078>
- 927 Mi H, Muruganujan A, Ebert D, Huang X, Thomas PD. 2019. PANTHER version 14: more  
928 genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic*  
929 *Acids Research* **47**: D419-D426. DOI: <https://doi.org/10.1093/nar/gky1038>
- 930 Miescher T. 1874. Das Protamin, eine neue organische Basis aus den Samen den des Rheinlachs.  
931 *Zeitschrift für Anal Chemie* **13**: 325-326. DOI: <https://doi.org/10.1007/bf01302341>
- 932 Nawrocki EP, Burge SW, Bateman A, Daub J, Eberhardt RY, Eddy SR, Floden EW, Gardner PP,  
933 Jones TA, Tate J, Finn RD. 2015. Rfam 12.0: updates to the RNA families database. *Nucleic*  
934 *Acids Research* **43**: D130-D137. DOI: <https://doi.org/10.1093/nar/gku1063>
- 935 Nixon B, Cafe SL, Eamens AL, De Iuliis GN, Bromfield EG, Martin JH, Skerrett-Byrne DA, Dun  
936 MD. 2020. Molecular insights into the divergence and diversity of post-testicular maturation  
937 strategies. *Molecular and Cellular Endocrinology* **517**: 110955. DOI:  
938 <https://doi.org/10.1016/j.mce.2020.110955>
- 939 Parra G, Blanco E, Guigó R. 2000. GeneID in *Drosophila*. *Genome Research* **10**: 511-515. DOI:  
940 <https://doi.org/10.1101/gr.10.4.511>
- 941 Pauletto M, Manousaki T, Ferraresso S, Babbucci M, Tsakogiannis A, Louro B, Vitulo N, Quoc  
942 VH, Carraro R, Bertotto D, Franch R, Maroso F, Aslam ML, Sonesson AK, Simionati B,  
943 Malacrida G, Cestaro A, Caberlotto S, Sarropoulou E, Mylonas CC, Power DM, Patarnello T,  
944 Canario AVM, Tsigenopoulos C, Bargelloni L. 2018. Genomic analysis of *Sparus aurata* reveals  
945 the evolutionary dynamics of sex-biased genes in a sequential hermaphrodite fish.  
946 *Communications Biology* **1**: 119. DOI: <https://doi.org/10.1038/s42003-018-0122-7>
- 947 Pérez LM. 2020. Fish sperm maturation, capacitation, and motility activation - Reproduction in  
948 Aquatic Animals. From Basic Biology to Aquaculture Technology in: Yoshida M, Asturiano JF.  
949 Springer. DOI: [https://doi.org/10.1007/978-981-15-2290-1\\_5](https://doi.org/10.1007/978-981-15-2290-1_5)
- 950 Perteua M, Perteua GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL. 2015. StringTie  
951 enables improved reconstruction of a transcriptome from RNA-seq reads. *Nature*  
952 *Biotechnology* **33**: 290-295. DOI: <https://doi.org/10.1038/nbt.3122>
- 953 Picelli S, Faridani OR, Björklund ÅK, Winberg G, Sagasser S, Sandberg R. 2014. Full-length

- 954 RNA-seq from single cells using Smart-seq2. *Nature Protocols* **9**: 171-181. DOI:  
955 <https://doi.org/10.1038/nprot.2014.006>
- 956 Powell JF, Zohar Y, Elizur A, Park M, Fischer WH, Craig AG, Rivier JE, Lovejoy DA, Sherwood  
957 NM. 1994. Three forms of gonadotropin-releasing hormone characterized from brains of one  
958 species. *PNAS* **91**: 12081-12085. DOI: <https://doi.org/10.1073/pnas.91.25.12081>
- 959 Puga Molina LC, Luque GM, Balestrini PA, Marin-Brigiler CI, Romarowski A, Buffone MG. 2018.  
960 Molecular basis of human sperm capacitation. *Frontiers in Cell and Developmental Biology* **6**:  
961 72. DOI: <https://doi.org/10.3389/fcell.2018.00072>
- 962 Rajamanickam GD, Kastelic JP, Thundathil JC. 2017. Content of testis-specific isoform of Na/K-  
963 ATPase (ATP1A4) is increased during bovine sperm capacitation through translation in  
964 mitochondrial ribosomes. *Cell Tissue Research* **368**: 187-200. DOI:  
965 <https://doi.org/10.1007/s00441-016-2514-7>
- 966 Rathke C, Baarends WM, Awe S, Renkawitz-Pohl R. 2014. Chromatin dynamics during  
967 spermiogenesis. *Biochimica et Biophysica Acta* **1839**: 155-168. DOI:  
968 <https://doi.org/10.1016/j.bbagr.2013.08.004>
- 969 Ren X, Chen X, Wang Z, Wang D. 2017. Is transcription in sperm stationary or dynamic? *Journal*  
970 *of Reproduction and Development* **63**: 439-443. DOI: <https://doi.org/10.1262/jrd.2016-093>
- 971 Saperas N, Ribes E, Buesa C, García-Hegart F, Chiva M. 1993. Differences in chromatin  
972 condensation during spermiogenesis in two species of fish with distinct protamines. *Journal of*  
973 *Experimental Zoology* **265**: 185-194. DOI: <https://doi.org/10.1002/jez.1402650210>
- 974 Sassone-Corsi P. 2002. Unique chromatin remodeling and transcriptional regulation in  
975 spermatogenesis. *Science* **296**: 2176-2178. DOI: <https://doi.org/10.1126/science.1070963>
- 976 Saucedo L, Sobarzo C, Brukman NG, Guidobaldi HA, Lustig L, Giojalas LC, Buffone MG,  
977 Vazquez-Levin MH, Marín-Briggiler C. 2018. Involvement of fibroblast growth factor 2 (FGF2)  
978 and its receptors in the regulation of mouse sperm physiology. *Reproduction* **156**: 163-172. DOI:  
979 <https://doi.org/10.1530/REP-18-0133>
- 980 Schulz RW, de França LR, Lareyre JJ, Le Gac F, Chiarini-Garcia H, Nobrega RH, Miura T. 2010.  
981 Spermatogenesis in fish. *General and Comparative Endocrinology* **165**: 390-411. DOI:  
982 <https://doi.org/10.1016/j.ygcen.2009.02.013>
- 983 Shimizu Y, Mita K, Tamura M, Onitake K, Yamashita M. 2000. Requirement of protamine for  
984 maintaining nuclear condensation of medaka (*Oryzias latipes*) spermatozoa shed into water but  
985 not for promoting nuclear condensation during spermatogenesis. *International Journal of*  
986 *Developmental Biology* **44**: 195-199.
- 987 Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing



- 988 genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*  
989 **31**: 3210-3212. DOI: <https://doi.org/10.1093/bioinformatics/btv351>
- 990 Sironen A, Kotaja N, Mulhern H, Wyatt TA, Sisson JH, Pavlik JA, Miiluniemi M, Fleming MD,  
991 Lee L. 2011. Loss of SPEF2 function in mice results in spermatogenesis defects and primary  
992 ciliary dyskinesia. *Biology of Reproduction* **85**: 690-701. DOI:  
993 <https://doi.org/10.1095/biolreprod.111.091132>
- 994 Soumillon M, Necsulea A, Weier M, Brawand D, Zhang X, Gu H, Barthès P, Kokkinaki M, Nef S,  
995 Gnirke A, Dym M, de Massy B, Mikkelsen TS, Kaessman H. 2013. Cellular source and  
996 mechanisms of high transcriptome complexity in the mammalian testis. *Cell Reports* **3**: 2179-  
997 2190. DOI: <https://doi.org/10.1016/j.celrep.2013.05.031>
- 998 Stanke M, Schöffmann O, Morgenstern B, Waack S. 2006. Gene prediction in eukaryotes with a  
999 generalized hidden Markov model that uses hints from external sources. *BMC Bioinformatics*  
1000 **7**: 62. DOI: <https://doi.org/10.1186/1471-2105-7-62>
- 1001 Sullivan R, Mieusset R. 2016. The human epididymis: its function in sperm maturation. *Human*  
1002 *Reproduction Update* **22**: 574-587. DOI: <https://doi.org/10.1093/humupd/dmw015>
- 1003 Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva  
1004 NT, Morris JH, Bork P, Jensen LJ, Mering CV. 2019. STRING v11: protein-protein association  
1005 networks with increased coverage, supporting functional discovery in genome-wide  
1006 experimental datasets. *Nucleic Acids Research* **47**: D607-D613. DOI:  
1007 <https://doi.org/10.1093/nar/gky1131>
- 1008 Tan W, Thomas P. 2015. Involvement of epidermal growth factor receptors and mitogen-activated  
1009 protein kinase in progestin-induction of sperm hypermotility in Atlantic croaker through  
1010 membrane progestin receptor-alpha. *Molecular and Cellular Endocrinology* **414**: 194-201. DOI:  
1011 <https://doi.org/10.1016/j.mce.2015.06.023>
- 1012 Tello JA, Wu S, Rivier JE, Sherwood NM. 2008. Four functional GnRH receptors in zebrafish:  
1013 analysis of structure, signaling, synteny and phylogeny. *Integrative Comparative Biology* **48**:  
1014 570-587. DOI: <https://doi.org/10.1093/icb/icn070>
- 1015 Trapnell C, Hendrickson DG, Sauvageau M, Goff L, Rinn JL, Pachter L. 2013. Differential analysis  
1016 of gene regulation at transcript resolution with RNA-seq. *Nature Biotechnology* **31**: 46-53.  
1017 DOI: <https://doi.org/10.1038/nbt.2450>
- 1018 Triphan X, Menzel VA, Petrunikina AM, Cassará MC, Wemheuer W, Hinsch KD, Hinsch E. 2008.  
1019 Localisation and function of voltage-dependent anion channels (VDAC) in bovine spermatozoa.  
1020 *Pflugers Archives* **455**: 677-686. DOI: <https://doi.org/10.1007/s00424-007-0316-1>
- 1021 Vogel P, Hansen G, Fontenot G, Read R. 2010. Tubulin tyrosine ligase-like 1 deficiency results in

- 1022 chronic rhinosinusitis and abnormal development of spermatid flagella in mice. *Veterinary*  
1023 *Pathology* **47**: 703-712. DOI: <https://doi.org/10.1177/0300985810363485>
- 1024 Wike CL, Guo Y, Tan M, Nakamura R, Shaw DK, Díaz N, Whittaker-Tademy AF, Durand NC,  
1025 Aiden EL, Vaquerizas JM, Grunwald D, Takeda H, Cairns BR. 2021. Chromatin architecture  
1026 transitions from zebrafish sperm through early embryogenesis. *Genome Research* **31**: 1-14. DOI:  
1027 <https://doi.org/10.1101/gr.269860.120>
- 1028 Wu S, Li H, Wang L, Mak N, Wu X, Ge R, Sun F, Cheng CY. 2021. Motor proteins and  
1029 spermatogenesis. *Advances in Experimental Medicine and Biology* **1288**: 131-159. DOI:  
1030 [https://doi.org/10.1007/978-3-030-77779-1\\_7](https://doi.org/10.1007/978-3-030-77779-1_7)
- 1031 Wu SF, Zhang H, Cairns BR. 2011. Genes for embryo development are packaged in blocks of  
1032 multivalent chromatin in zebrafish sperm. *Genome Research* **21**: 578-589. DOI:  
1033 <https://doi.org/10.1101/gr.113167.110>
- 1034 Wu TD, Watanabe CK. 2005. GMAP: a genomic mapping and alignment program for mRNA and  
1035 EST sequences. *Bioinformatics* **21**: 1859-1875. DOI:  
1036 <https://doi.org/10.1093/bioinformatics/bti310>
- 1037 Xia B, Yan Y, Baron M, Wagner F, Barkley D, Chiodon M, Kim SY, Keefe DL, Alukal JP, Boeke  
1038 JD, Yanai I. 2020. Widespread transcriptional scanning in the testis modulates gene evolution  
1039 rates. *Cell* **180**: 248-262. DOI: <https://doi.org/10.1016/j.cell.2019.12.015>
- 1040 Yoshida S. 2016. From cyst to tubule: innovations in vertebrate spermatogenesis. *Wiley*  
1041 *Interdisciplinary Reviews in Developmental Biology* **5**: 119-131. DOI:  
1042 <https://doi.org/10.1002/wdev.204>
- 1043 Zhao C, Guo XJ, Shi HZ, Wang FQ, Huang XY, Huo R, Wang XR, Liu JY, Zhou ZM, Sha JH.  
1044 2009. Role of translation by mitochondrial-type ribosomes during sperm capacitation: An  
1045 analysis based on a proteomic approach. *Proteomics* **9**: 1385-1399. DOI:  
1046 <https://doi.org/10.1002/pmic.200800353>
- 1047 Zhao W, Quansah E, Yuan M, Gou Q, Mengal K, Li P, Wu S, Xu C, Yi C, Cai X. 2019. Region-  
1048 specific gene expression in the epididymis of Yak. *Theriogenology* **139**: 132-146. DOI:  
1049 <https://doi.org/10.1016/j.theriogenology.2019.08.006>
- 1050 Zhou Y, Ru Y, Shi H, Wang Y, Wu B, Upur H, Zhang Y. 2015. Cholecystokinin receptors regulate  
1051 sperm protein tyrosine phosphorylation via uptake of HCO<sub>3</sub><sup>-</sup>. *Reproduction* **150**: 257-268. DOI:  
1052 <https://doi.org/10.1530/REP-15-0138>
- 1053 Zhu Z, Umehara T, Okazaki T, Goto M, Fujita Y, Hoque SAM, Kawai T, Zeng W, Shimada M.  
1054 2019. Gene expression and protein synthesis in mitochondria enhance the duration of high-speed  
1055 linear motility in boar sperm. *Frontiers in Physiology* **10**: 252. DOI:

1056 <https://doi.org/10.3389/fphys.2019.00252>

1057

1058 **Figure Legends and Supplementary Files**

1059

1060 **Figure 1. Transcriptome profiling of seabream haploid germ cells and ejaculated**

1061 **spermatozoa.** (A and B) Representative flow cytometry plots of the seabream cell populations in  
1062 the whole testis. In A, the populations of testicular haploid and diploid cells are encircled. In B, the  
1063 different subpopulations of haploid germinal cells (HGC), corresponding to a mix of type II  
1064 spermatocytes (SPC II) and spermatids (SPD), and intratesticular spermatozoa (SPZ<sub>I</sub>) are shown.  
1065 (C) Representative immunostaining of Lys<sup>9</sup> acetylated histone 3 (H3K9ac), meiotic recombination  
1066 protein Spo11 and  $\alpha$ -tubulin (Tuba) in sorted HGC and ejaculated spermatozoa (SPZ<sub>EJ</sub>). For each  
1067 cell type the brightfield (left panels) and epifluorescence (right panels) images are shown. SPD<sub>R</sub>,  
1068 round spermatids; SPD<sub>E</sub>, elongating spermatids. Scale bars, 2 and 5  $\mu$ m. (D) Principal component  
1069 analysis (PCA) using the top 500 most variable genes between HGC and SPZ<sub>EJ</sub> ( $n = 4$  pools) and  
1070 the 'rlog' transformation of the counts. (E) Heatmap generated by unsupervised hierarchical  
1071 clustering of RNAseq expression z-scores computed for the 7,287 differentially expressed genes  
1072 (DEGs) ( $p$ -adj < 0.01; Log<sub>2</sub> fold change > 1) between HGC and SPZ<sub>EJ</sub>. (F) Volcano plot  
1073 representation of DEGs in the SPZ<sub>EJ</sub> versus HGC comparison. The x-axis shows Log<sub>2</sub> fold changes  
1074 in expression and the y-axis the negative logarithm of their  $p$ -value to base 10. Red and green points  
1075 mark the genes with significantly increased or decreased expression respectively in SPZ<sub>EJ</sub> compared  
1076 to HGC (FDR < 0.01). (G) Venn diagrams showing the number of common mRNAs and lncRNAs  
1077 (in intersect region) which are differentially expressed between HGCs and SPZ<sub>EJ</sub>. (H) Validation of  
1078 the RNAseq data by qRT-PCR. The plot represents the Pearson's correlation analysis of DEGs in  
1079 HGC and SPZ<sub>EJ</sub> determined by RNAseq and qRT-PCR. The Pearson's correlation coefficient (PCC)  
1080 of the Log<sub>2</sub> fold change analyzed by RNAseq (x-axis) and using qRT-PCR (y-axis), the  $p$ -value,  
1081 and the number of DEGs analyzed are indicated.

1082 **Figure 1- source data 1**

1083 Data for PCA shown in D.

1084 **Figure 1- source data 2**

1085 Data for the heat map shown in E.

1086 **Figure 1- source data 3**

1087 Data for Volcano plot shown in F.

1088 **Figure 1- source data 4**

1089 Data on the validation of the RNAseq data by qRT-PCR.

1090

1091 **Figure 2. Functional classification of DEGs during sperm differentiation and maturation.** (A)

1092 Transcriptional regulation of a subpopulation of DEGs classified into five functional categories:  
1093 transcription and translation and chromatin organization, receptors, metabolism, cytoskeleton and  
1094 cell movement, and channels, exchangers and transporters. **(B-F)** Pie charts showing the GO term  
1095 distribution of upregulated DEGs in SPZ<sub>EJ</sub> included in each of the five functional groups. The  
1096 numbers are the percentage of genes in each category.

1097 **Figure 2-source data 1**

1098 Data for the classification of DEGs.

1099 **Figure 2-Supplement 1**

1100 Gene ontology (GO) enrichment analysis of the DEGs during sperm differentiation and maturation.  
1101 GO annotation of DEGs corresponding to biological process level 2 (A) and 5 (B), and molecular  
1102 function level 5 (C). The horizontal axis displays the number of significant genes corresponding to  
1103 each functional type, whereas the vertical axis displays the second level of GO annotation.

1104 **Figure 2-Supplement 1-source data 1**

1105 Data from GO analysis of the DEGs during sperm differentiation and maturation.

1106

1107 **Figure 3. Protein-protein interaction (PPI) networks of DEGs.** The PPI information of DEGs  
1108 potentially involved in transcription and translation and chromatin organization (A), and  
1109 metabolism (B), was obtained through a database search using STRING database v11 with a high  
1110 confidence score (0.9), and imported into Cytoscape v3.8.2 for network construction. Proteins and  
1111 their interactions are shown as nodes (spheres) and edges (lines), respectively. Nodes in red or green  
1112 color indicate upregulated and downregulated DEGs, respectively. Proteins are grouped based on  
1113 their known biological functions. Abbreviations: OXPHOS, oxidative phosphorylation; PPP,  
1114 pentose phosphate pathway; TCA, tricarboxylic acid.

1115 **Figure 3-Supplement 1**

1116 Mapping of DEGs coding for enzymes involved in respiratory pathways. Schematic diagram of the  
1117 biochemical pathways of glycolysis/gluconeogenesis, pentose phosphate (PP) pathway,  
1118 tricarboxylic acid (TCA) cycle and oxidative phosphorylation (OXPHOS). Enzyme-coding DEGs in  
1119 green and red color denotes downregulation and upregulation, respectively, whereas black color  
1120 indicates no change in the expression levels.

1121

1122 **Figure 4. Pathway enrichment analysis during spermatozoa differentiation and maturation.**

1123 (A) Pathway analysis of DEGs using the PANTHER Classification System showing the 37 most  
1124 highly enriched signaling pathways (FDR < 0.05) in SPZ<sub>EJ</sub>. (B and C) Hierarchical clustering  
1125 heatmaps of DEGs related to the PDGF (B) and GnRHR (C) signaling pathways. (D) qRT-PCR

1126 validation of the changes in expression of several genes classified into the PDGF or GnRHR  
1127 pathways. Data from qRT-PCR are the mean  $\pm$  SEM ( $n = 3$  pools of 3 different fish each).

1128 **Figure 4-source data 1**

1129 Data for heatmap shown in B.

1130 **Figure 4-source data 2**

1131 Data for heatmap shown in C.

1132 **Figure 4-source data 3**

1133 Data on the qRT-PCR validation of the changes in expression of several genes classified into the  
1134 PDGF or GnRHR pathways.

1135

1136 **Figure 5. Cellular localization of GnRH expression in seabream extratesticular ducts. (A)**

1137 Anatomy of the seabream testis and extratesticular ducts, efferent duct (ED) and sperm duct (SD).

1138 **(B)** Paraffin histological sections of the different structures of the testis and testicular ducts stained  
1139 with hematoxylin and eosin. **(C)** Localization of *sbgnrh* transcripts in the testis, ED and SD by *in*

1140 *situ hybridization* on paraffin sections hybridized with antisense DIG-labeled riboprobes specific

1141 for *sbgnrh* (upper panels) or sense probes (lower panels, negative controls). **(D)** Immunostaining of

1142 GnRH peptides (red, lower panels) in the same testicular structures as in C. Corresponding

1143 brightfield (BF) images are also shown (upper panels). The reactions were visualized with Cy3-

1144 conjugated sheep anti-rabbit IgG and the nuclei were counterstained with 4',6-diamidino-2-

1145 phenylindole (DAPI; blue). Control sections incubated with the secondary antibody only did not

1146 show any staining (**Figure 5-Supplement 2**). Scale bars, 50  $\mu$ m (B and C), 10  $\mu$ m (D).

1147 Abbreviations: SPC, spermatocytes; SPZ<sub>I</sub>, intratesticular spermatozoa. The arrowheads in B-D

1148 indicate epithelial cells of the ED and SD.

1149 **Figure 5-Supplement 1**

1150 Localization of *sgnrh* transcripts in the seabream testis and efferent and spermatic ducts. Paraffin

1151 sections from the testis, efferent duct (ED) and two regions of the spermatic duct (SD) were

1152 hybridized with antisense DIG-labeled riboprobes specific for *sgnrh* or with specific sense probes

1153 (lower panels, negative controls). Scale bars, 50  $\mu$ m. Abbreviations: SPG, spermatogonia; SPC,

1154 spermatocytes; SPD, spermatids; SPZ<sub>I</sub>, intratesticular spermatozoa; SC, Sertoli cells. The

1155 arrowheads indicate epithelial cells of the efferent and spermatic ducts.

1156 **Figure 5-Supplement 2**

1157 Control sections from the seabream testis, efferent duct (ED) and sperm duct (SD) incubated with

1158 the secondary antibody only. The upper panels show the brightfield (BF) images, whereas the lower

1159 panels show the epifluorescence images. Scale bars, 10  $\mu$ m. Abbreviations: SPC, spermatocytes;

1160 SPZ<sub>I</sub>, intratesticular spermatozoa. The arrowheads indicate epithelial cells of the efferent and  
1161 spermatic ducts.

1162

1163 **Figure 6. Localization of *pdgf* transcripts in the seabream testis, ED and SD.** (A-D) Paraffin  
1164 sections were hybridized with antisense DIG-labeled riboprobes specific for different *pdgf* paralogs  
1165 (upper panels) as indicated. Control sections (lower panels), hybridized with sense probes, were  
1166 negative. Scale bars, 50  $\mu$ m. SPG, spermatogonia; SPC, spermatocytes; SPD, spermatids; SPZ<sub>I</sub>,  
1167 intratesticular spermatozoa. The arrowheads indicate epithelial cells of the ED and SD.

#### 1168 **Figure 6-Supplement 1**

1169 Localization of *pdgfaa* and *pdgfc* transcripts in the seabream testis and efferent and spermatic ducts.  
1170 (A-B) Paraffin sections from the testis, efferent duct (ED) and two regions of the spermatic duct  
1171 (SD) were hybridized with antisense DIG-labeled riboprobes specific for *pdgfaa* (A) and *pdgfc* (B)  
1172 (upper panels) or with specific sense probes (lower panels, negative controls). Scale bars, 50  $\mu$ m.  
1173 Abbreviations: SPG, spermatogonia; SPC, spermatocytes; SPD, spermatids; SPZ<sub>I</sub>, intratesticular  
1174 spermatozoa; SC, Sertoli cells. The arrowheads indicate epithelial cells of the efferent and spermatic  
1175 ducts.

1176

1177 **Figure 7. Transcriptional regulation of seabream sperm motility by GnRH and PDGF.** (A)  
1178 Percentage of motility (%MOT) and progressivity (%PROG), and curvilinear velocity (VCL), of  
1179 spermatozoa from the efferent duct (SPZ<sub>ED</sub>) or ejaculated (SPZ<sub>EJ</sub>) determined at 5 or 30 s post  
1180 activation. (B and C) RT-PCR detection of mRNAs encoding GnRH receptors (*gnrhr1*, *gnrhr2* and  
1181 *gnrhr3*) and PDGF receptors b (*pdgfra* and *pdgfrb*) in SPZ<sub>ED</sub> or SPZ<sub>EJ</sub>. The Neg. line is the negative  
1182 control (absence of RT during cDNA synthesis). The arrows indicate the specific transcripts, and  
1183 the size (kb) of molecular markers are indicated on the left. (D) The % MOT and VCL of SPZ<sub>ED</sub>  
1184 exposed to 100 nM of sbGnRH or sGnRH, 40 nM of mouse recombinant PDGF (rPDGF-BB), or to  
1185 each hormone vehicle, determined at 5 or 30 s post activation. (E) Inhibition of motility of SPZ<sub>ED</sub>  
1186 induced by sbGnRH and rPDGF-BB by 100  $\mu$ g/ml actinomycin D (AcD) or chloramphenicol (CP)  
1187 at 5 s post activation. (F) Quantitative RT-PCR analysis of the expression of selected genes  
1188 potentially involved in water and ion transport, signaling, flagellar motility and glucose metabolism  
1189 in SPZ, after sbGnRH or rPDGF-BB stimulation. The Log<sub>2</sub> fold change in the expression of each  
1190 gene in the RNA-seq analysis is indicated on the right. In A, D and E, all data points are presented  
1191 as box and whisker plots/scatter dots with horizontal line (inside box) indicating median and  
1192 outliers. One ejaculate from each male was measured from  $n = 5-7$  males. In F, data are the mean  $\pm$   
1193 SEM ( $n = 3-5$  fish). Data were statistically analyzed by an unpaired Student's *t*-test (A and F), or by

1194 one-way ANOVA (D and E). \*,  $P < 0.05$ ; \*\*,  $P < 0.01$ ; \*\*\*,  $P < 0.001$ ; with respect to spermatozoa  
1195 incubated with the hormone vehicles, or as indicated in brackets.

1196 **Figure 7-source data 1**

1197 Data on sperm motility shown in A.

1198 **Figure 7-source data 2**

1199 Uncropped gels of the RT-PCR of mRNAs encoding seabream GnRH receptors (*gnrhr1*, *gnrhr2*  
1200 and *gnrhr3*) in SPZ<sub>ED</sub> or SPZ<sub>EJ</sub>. The Neg. line is the negative control (absence of RT during cDNA  
1201 synthesis). The arrows indicate the specific transcripts, and molecular markers are on the left.

1202 **Figure 7-source data 3**

1203 RT-PCR detection of mRNAs encoding seabream PDGF receptors b (*pdgfra* and *pdgfrb*) in SPZ<sub>ED</sub>  
1204 or SPZ<sub>EJ</sub>. The Neg. line is the negative control (absence of RT during cDNA synthesis). The arrows  
1205 indicate the specific transcripts, and the molecular markers are on the left.

1206 **Figure 7-source data 4**

1207 Data on sperm motility shown in D.

1208 **Figure 7-source data 5**

1209 Data on sperm motility shown in E

1210 **Figure 7-Supplement 1**

1211 Sperm motion kinetics of seabream SPZ<sub>ED</sub>. (A) Percentage of progressivity (PROG) of SPZ<sub>ED</sub>  
1212 exposed to 100 nM of sbGnRH or sGnRH, 40 nM of recombinant PDGF (rPDGF-BB), or to each  
1213 hormone vehicle, determined at 5 or 30 s postactivation. (B) Inhibition of PROG and curvilinear  
1214 velocity (VCL) of SPZ<sub>ED</sub> induced by sbGnRH and rPDGF-BB by 100 µg/ml actinomycin D (ActD)  
1215 or chloramphenicol (CP) at 5 s postactivation. All data points are presented as box and whisker  
1216 plots/scatter dots with horizontal line (inside box) indicating median and outliers. Data were  
1217 statistically analyzed by one-way ANOVA. \*,  $P < 0.05$ ; \*\*,  $P < 0.01$ ; \*\*\*,  $P < 0.001$ ; with respect  
1218 to spermatozoa incubated with the hormone vehicles, or as indicated in brackets.

1219 **Figure 7-Supplement 1-source data 1**

1220 Data on sperm motility shown in Figure 7-Supplement 1.

1221 **Figure 7-source data 6**

1222 Quantitative RT-PCR analysis of the expression of selected genes shown in F.

1223

1224 **Figure 8. Transcription-dependent regulation of zebrafish sperm motility by GnRH and**

1225 **PDGF. (A)** Paraffin histological sections of the zebrafish testis and extratesticular ducts (ETD)

1226 stained with hematoxylin and eosin. Scale bars, 10 and 100 µm. **(B)** Immunostaining of GnRH

1227 peptides (red, right panel) in the surface epithelium of the ETDs (arrowheads) and corresponding



1228 brightfield (BF) image (left panel). Control sections incubated with the secondary antibody only  
1229 were negative (**Figure 8-Supplement 1**). Scale bar, 200  $\mu\text{m}$ . (**C** and **D**) Paraffin sections the ETDs  
1230 hybridized with antisense DIG-labeled riboprobes specific for *pdgfaa* (**C**) and *pdgfb* (**D**) mRNAs  
1231 as indicated. The arrowheads indicate expression in the epithelial cells of the ETDs. The right  
1232 panels show the absence of signals in sections hybridized with sense probes. SPZ<sub>ETD</sub>, spermatozoa  
1233 from the ETDs. (**E**) RT-PCR detection of mRNAs encoding GnRH (*gnrhr1*, *gnrhr2*, *gnrhr3* and  
1234 *gnrhr4*) and PDGF (*pdgfra* and *pdgfrb*) receptors in SPZ<sub>ETD</sub> and SPZ<sub>EJ</sub>. The Neg. line is the  
1235 negative control (absence of RT during cDNA synthesis). The arrows indicate the specific  
1236 transcripts, and the size (kb) of molecular markers are indicated on the left. (**F**) Percentage of  
1237 motility (%MOT) and progressivity (%PROG), and curvilinear velocity (VCL), at 5 s postactivation  
1238 of SPZ<sub>ETD</sub> exposed to 100 nM of sbGnRH, 40 nM of mouse recombinant PDGF (rPDGF-BB), or to  
1239 each hormone vehicle, in the presence or absence of actinomycin D (AcD; 100  $\mu\text{g}/\text{ml}$ ). Data are  
1240 presented as box and whisker plots/scatter dots with horizontal line (inside box) indicating median  
1241 and outliers ( $n = 6\text{-}12$  fish), and were statistically analyzed by an unpaired Student's *t*-test. \*,  $P <$   
1242 0.05; \*\*,  $P < 0.01$ ; \*\*\*,  $P < 0.001$ ; with respect to spermatozoa incubated with the hormone  
1243 vehicles, or as indicated in brackets.

#### 1244 **Figure 8-Supplement 1**

1245 Control section of the zebrafish (ETDs) incubated with the secondary antibody only. The left panel  
1246 shows the brightfield (BF) image, whereas the right panel show the epifluorescence image. The  
1247 arrowheads indicate the ED epithelium. Scale bar, 200  $\mu\text{m}$ . Abbreviations: SPZ<sub>ETD</sub>, sperm from the  
1248 ETD.

#### 1249 **Figure 8-Supplement 2**

1250 Kinematic properties of SPZ<sub>ETD</sub> and SPZ<sub>EJ</sub> from zebrafish. Percentage of motility (MOT) and  
1251 progressivity (PROG), and curvilinear velocity (VCL), of zebrafish spermatozoa from  
1252 extratesticular ducts (SPZ<sub>ETD</sub>) or ejaculated (SPZ<sub>EJ</sub>) determined at 5 s postactivation. All data points  
1253 are presented as box and whisker plots/scatter dots with horizontal line (inside box) indicating  
1254 median and outliers. One ejaculate from  $n = 7$  males was measured. Data were statistically analyzed  
1255 by an unpaired Student's *t*-test. \*\*,  $P < 0.01$ ; \*\*\*,  $P < 0.001$ ; with respect to SPZ<sub>ED</sub>.

#### 1256 **Figure 8-Supplement 2-source data 1**

1257 Data on sperm kinetics shown in Figure 8-Supplement 2.

#### 1258 **Figure 8-source data 1**

1259 Uncropped gels from RT-PCR detection of mRNAs encoding zebrafish GnRH (*gnrhr1*, *gnrhr2*,  
1260 *gnrhr3* and *gnrhr4*) and PDGF (*pdgfra* and *pdgfrb*) receptors in SPZ<sub>ETD</sub> and SPZ<sub>EJ</sub>. The Neg. line is  
1261 the negative control (absence of RT during cDNA synthesis). The arrows indicate the specific

1262 transcripts, and the molecular markers are on the left.

1263 **Figure 8-source data 2**

1264 Data on sperm motility shown in F.

1265

1266 **Supplementary file 1.** Nucleotide sequences of the primers employed for ISH probe synthesis and  
1267 alignment of probes.

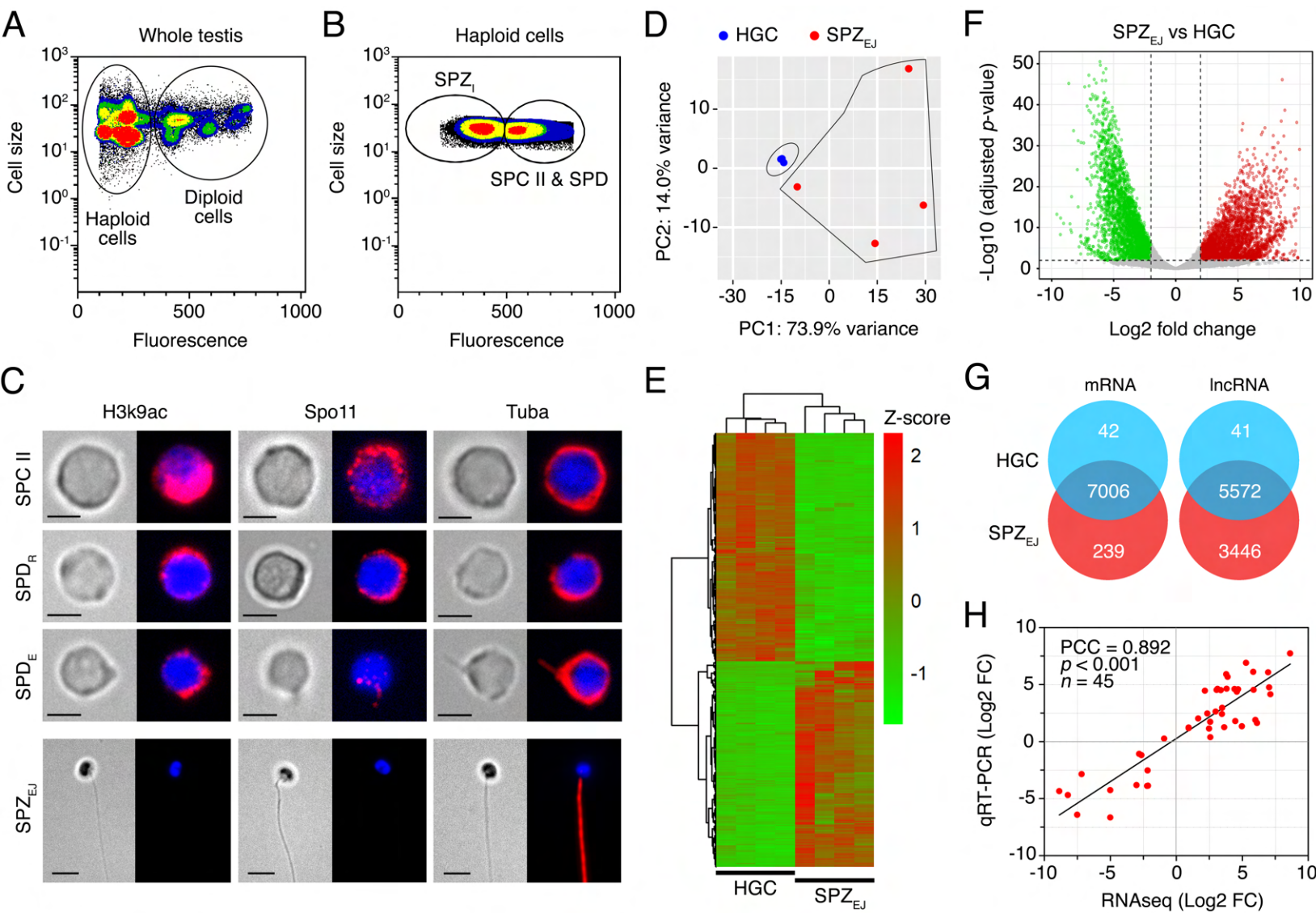
1268

1269 **Supplementary file 2.** Nucleotide sequences of the primers employed for RT-PCR and qRT-PCR  
1270 analyses.

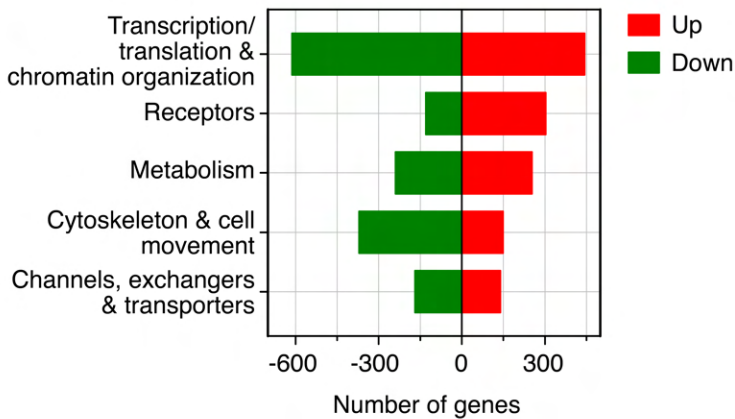
1271

1272

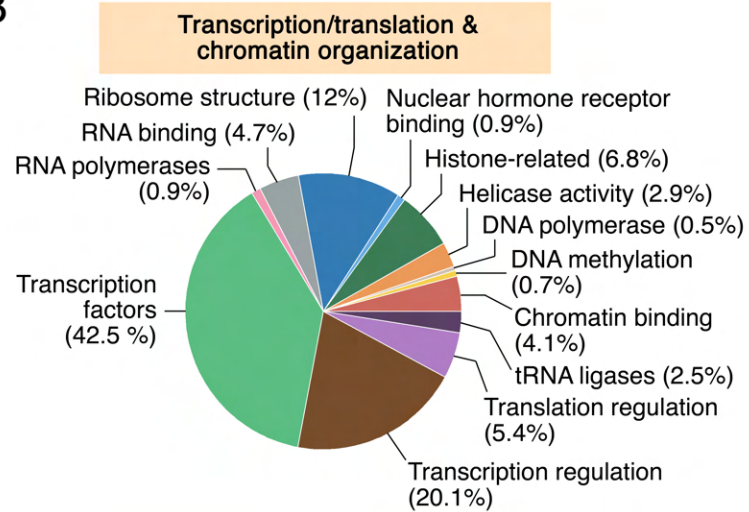
1273



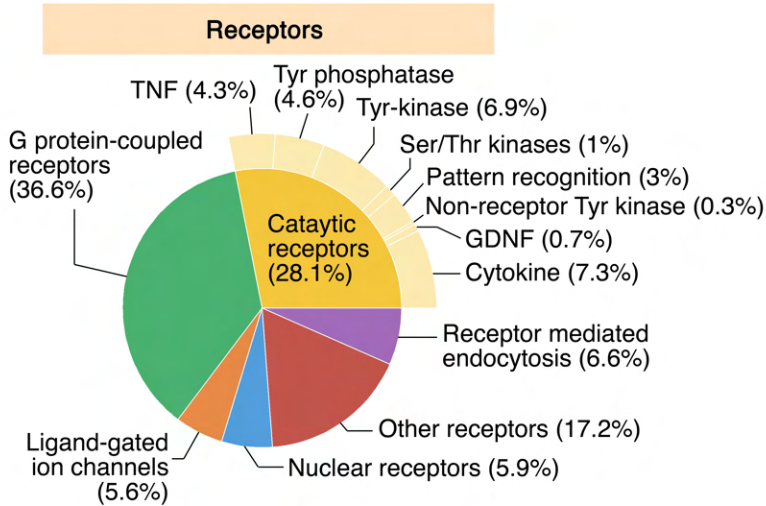
A



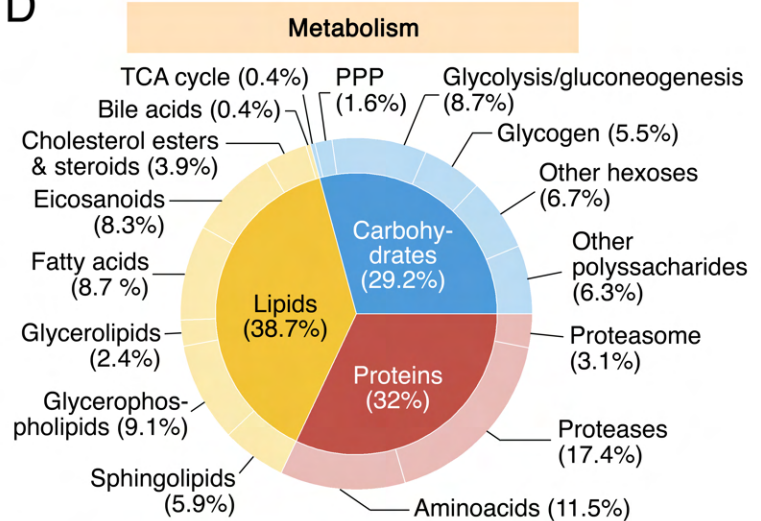
B



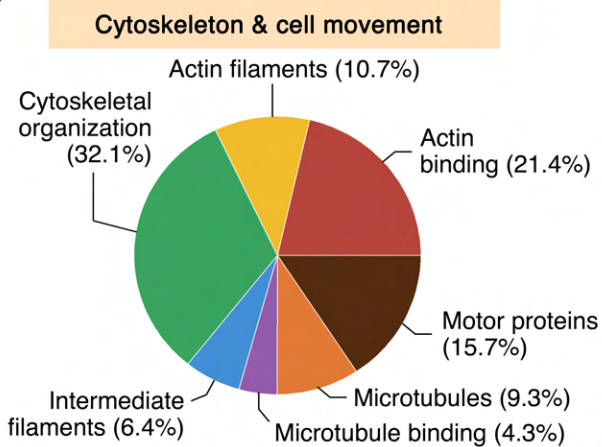
C



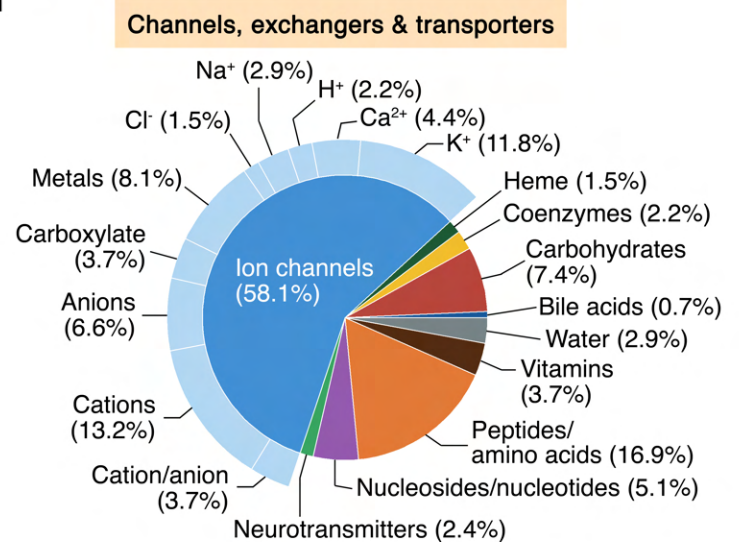
D



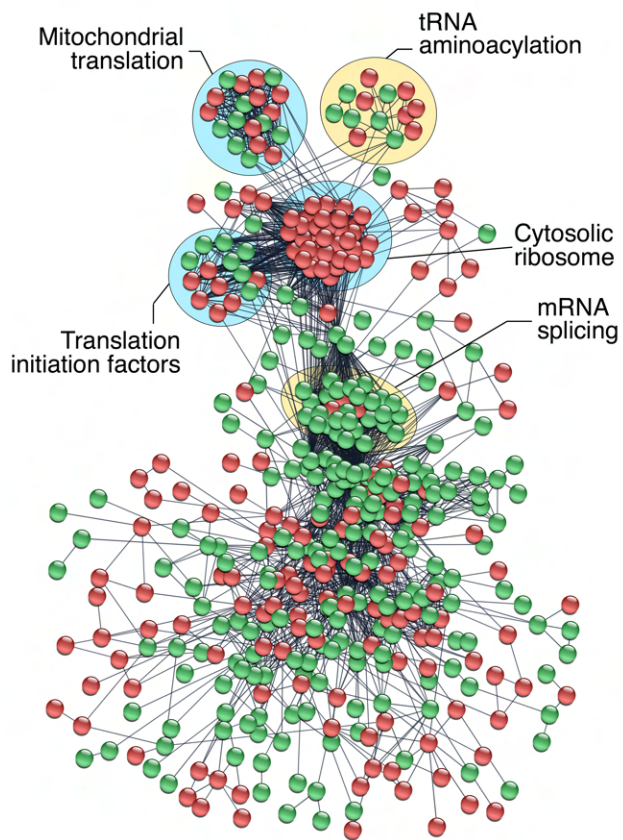
E



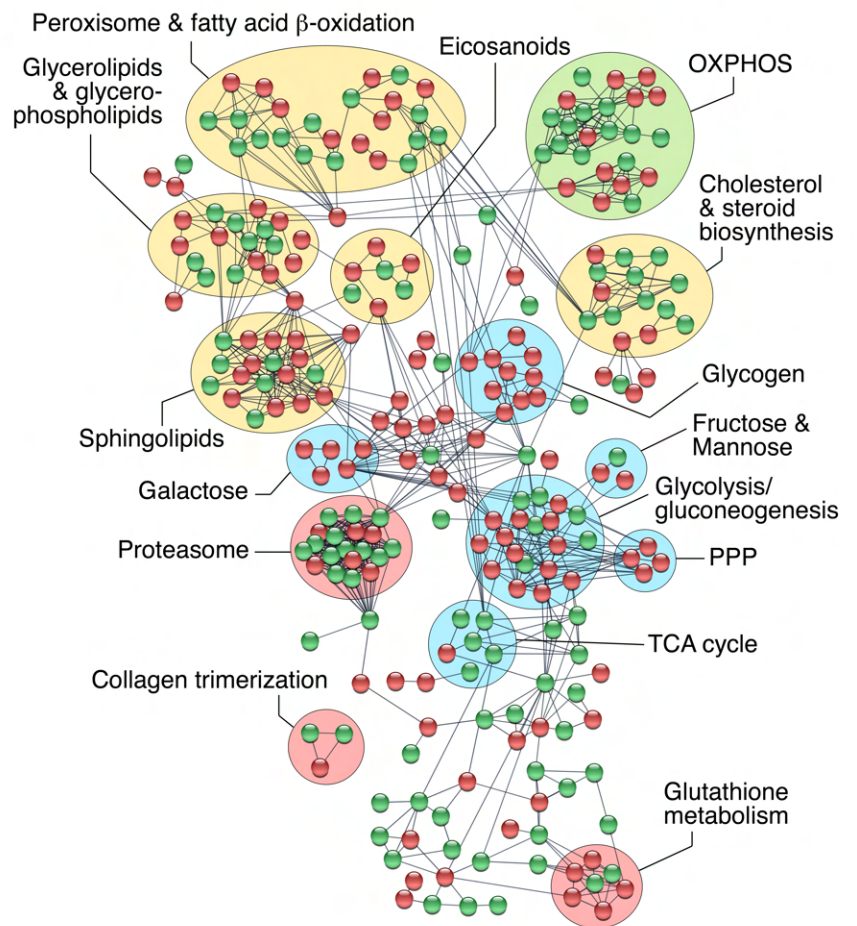
F

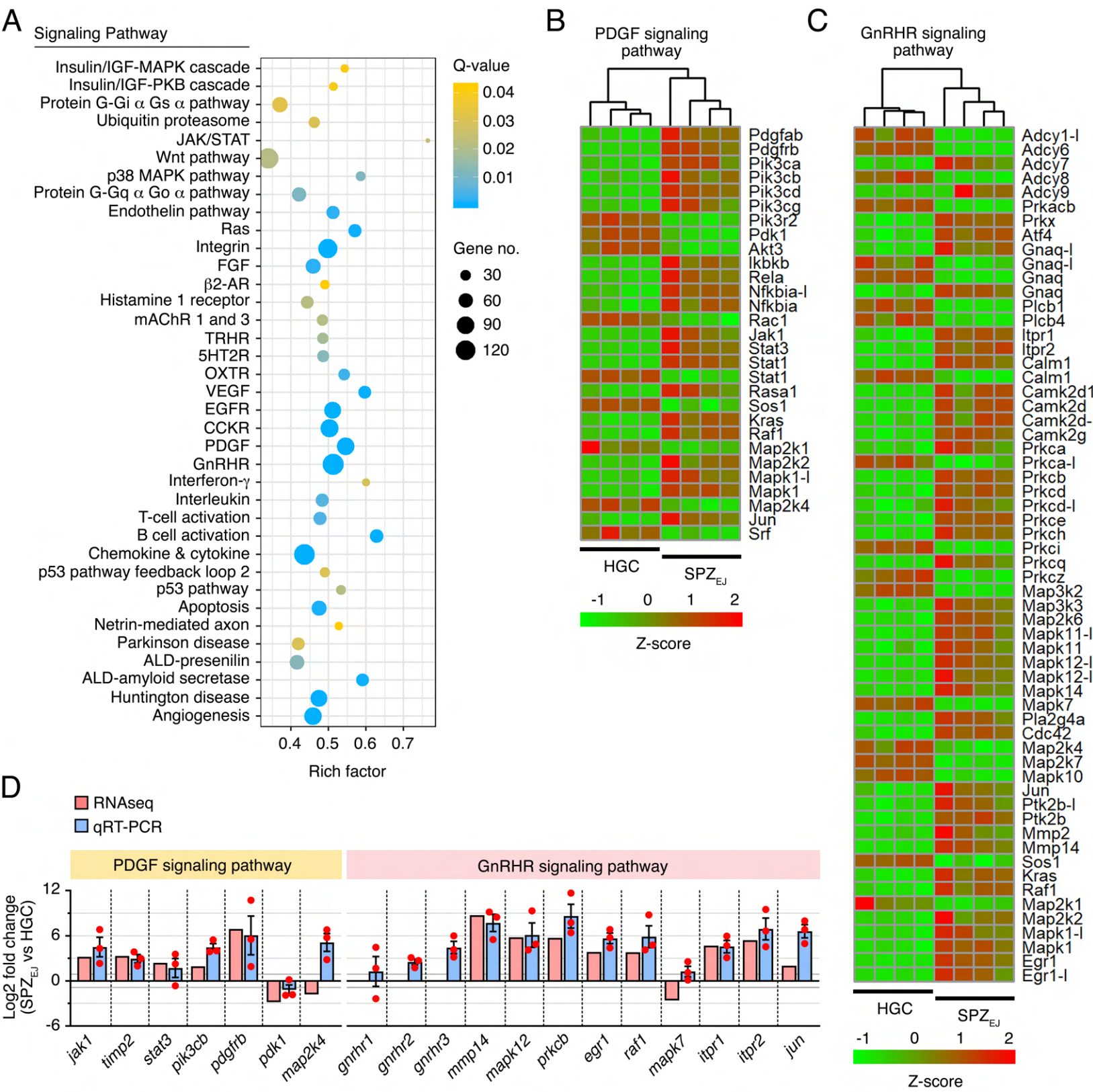


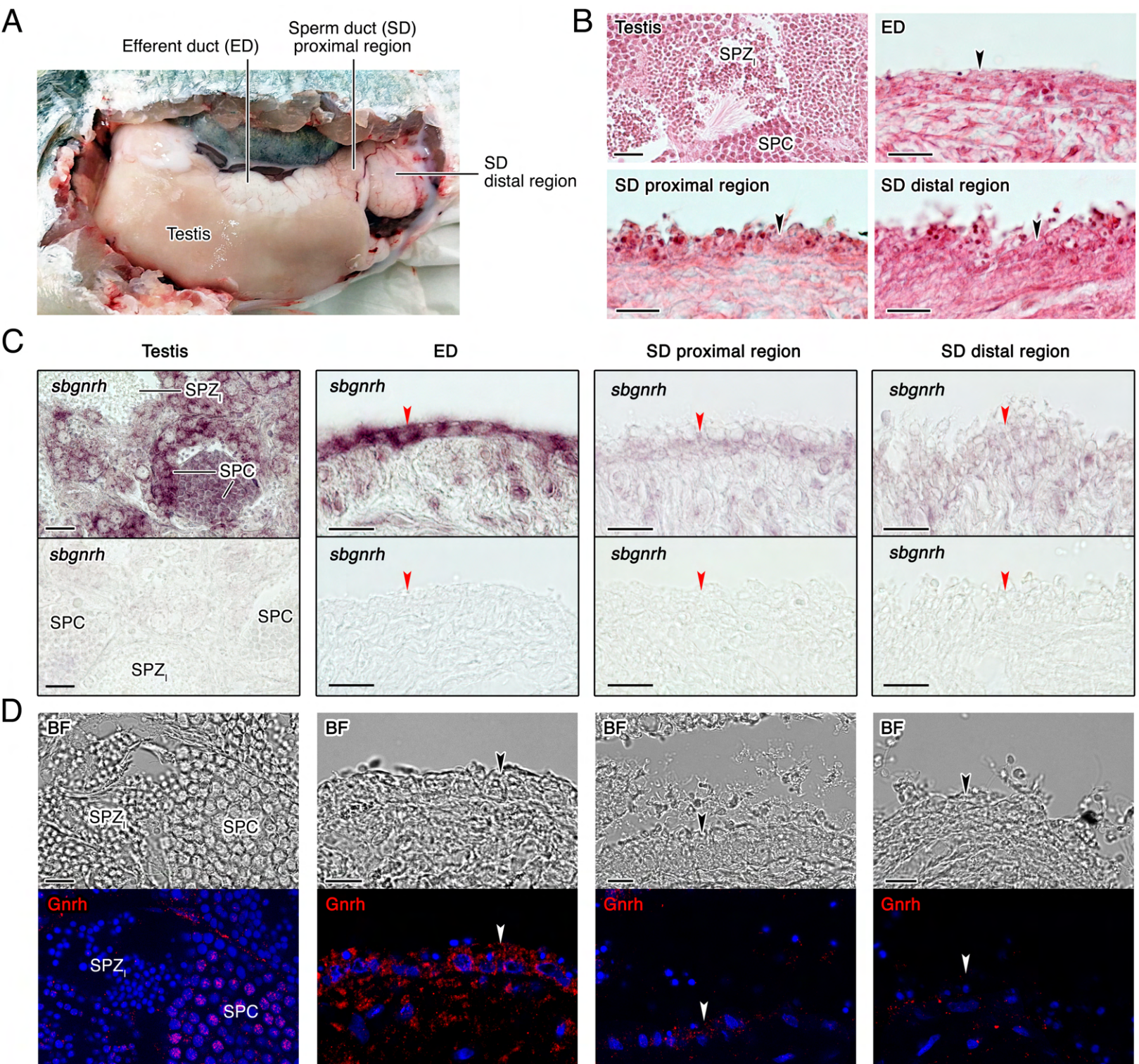
A

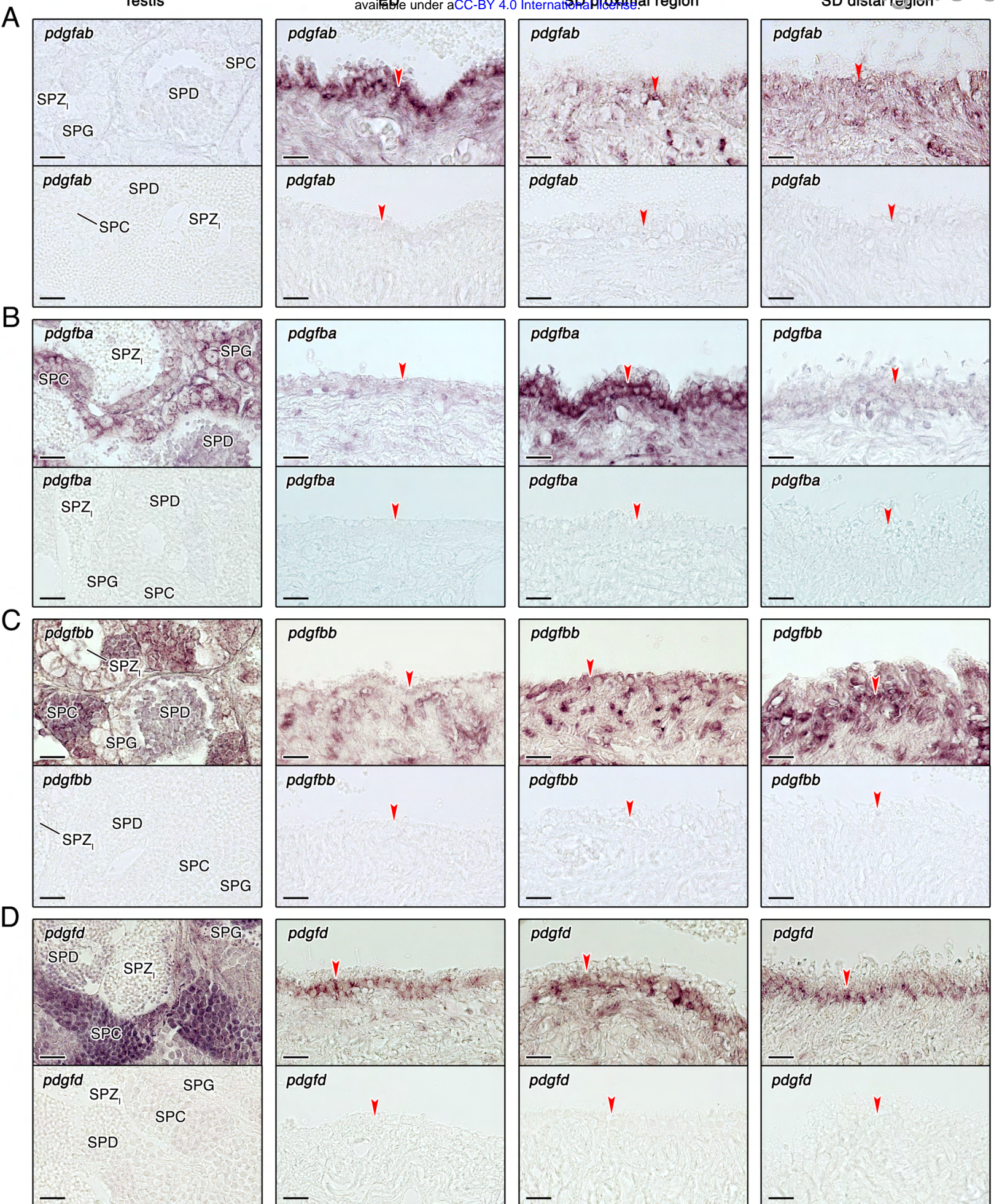


B

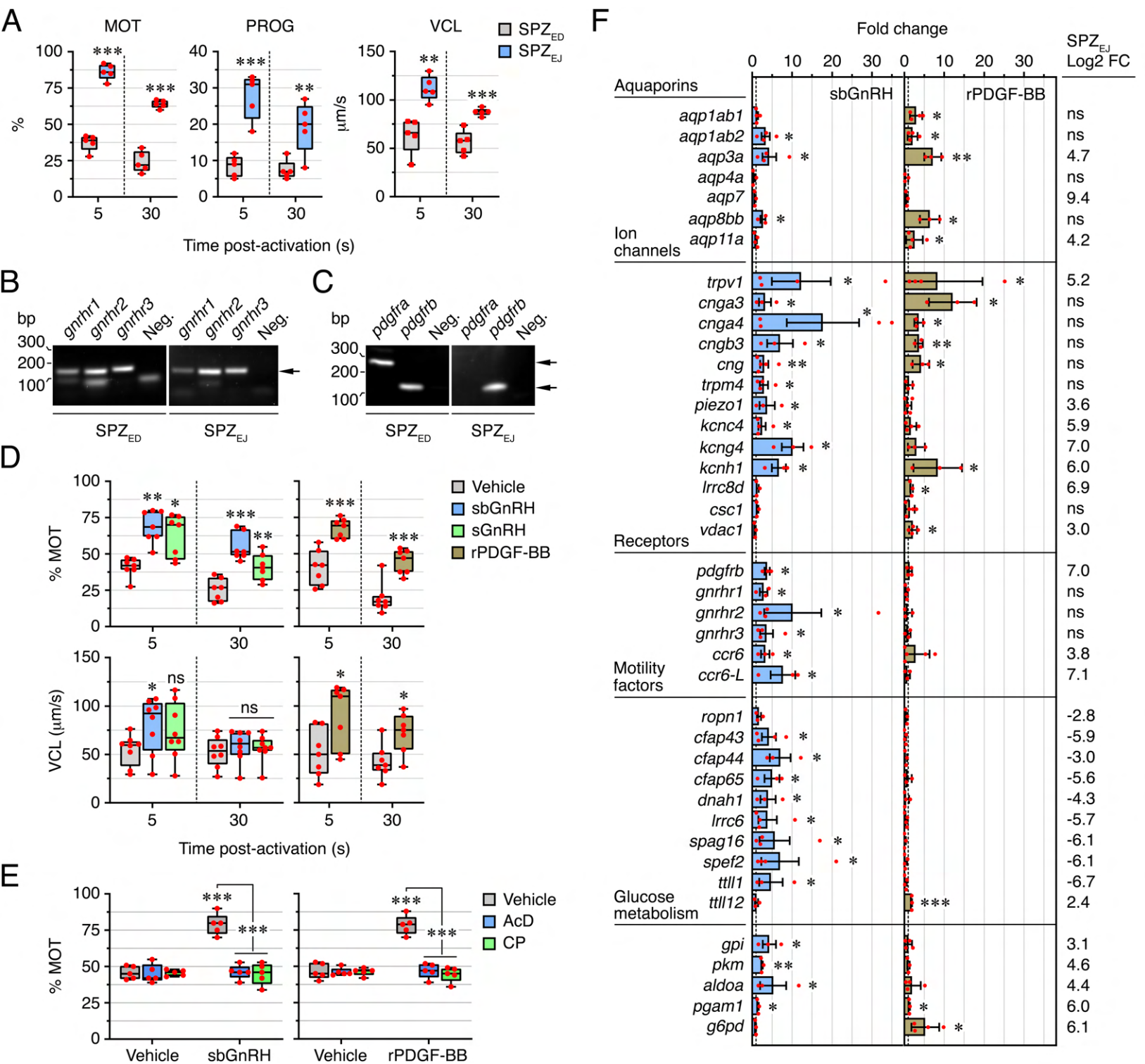


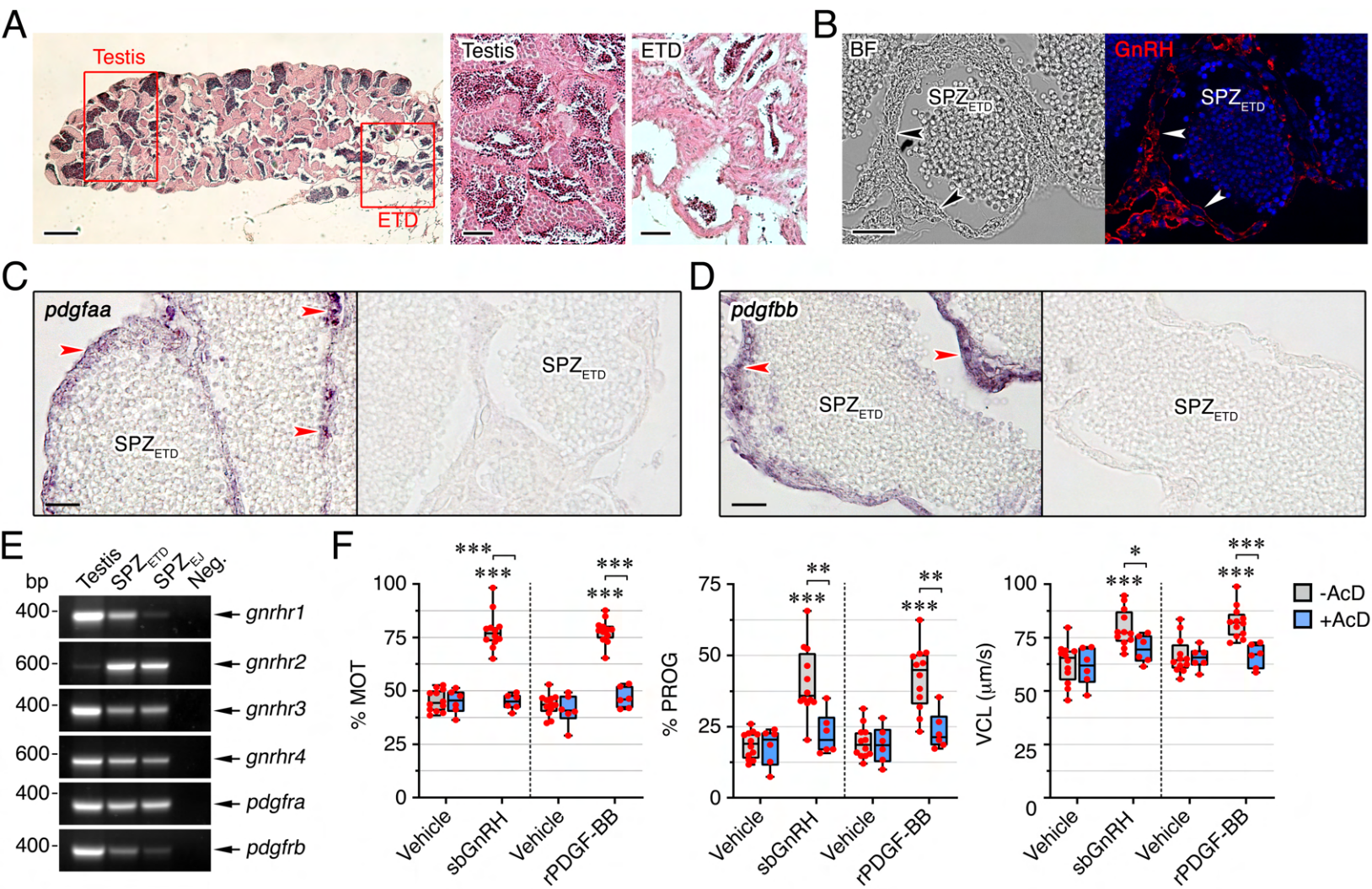




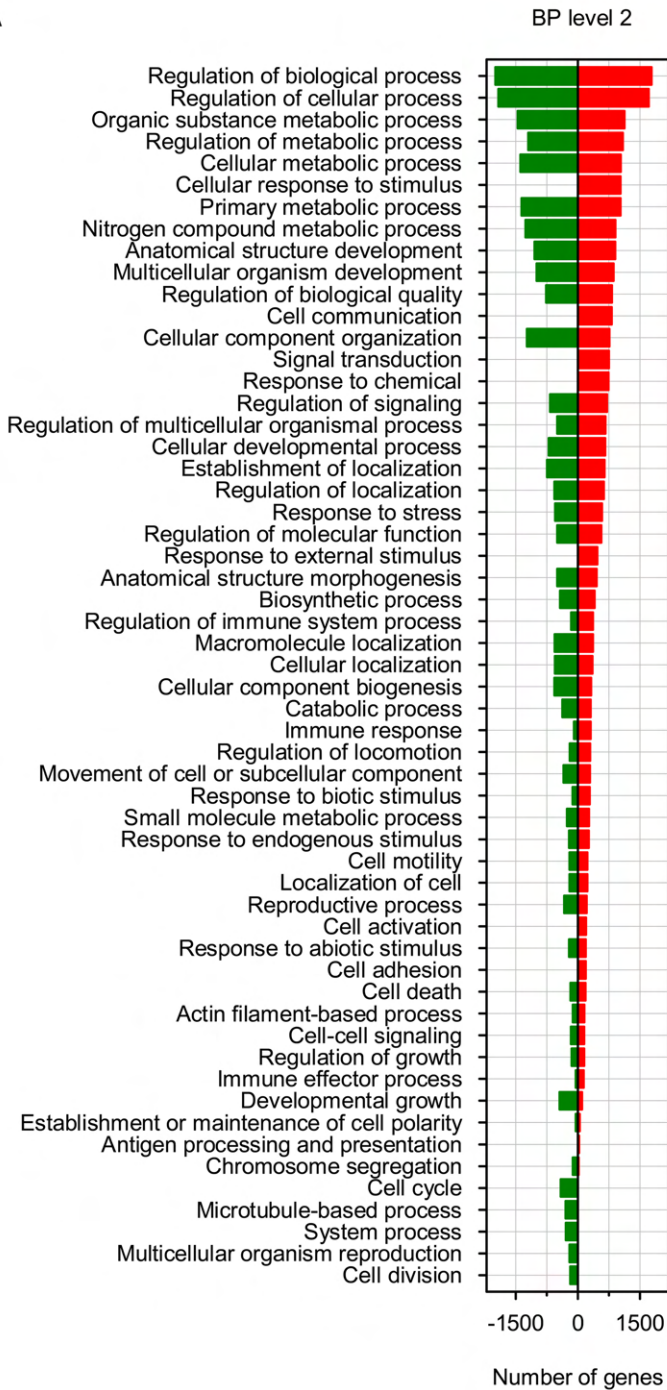




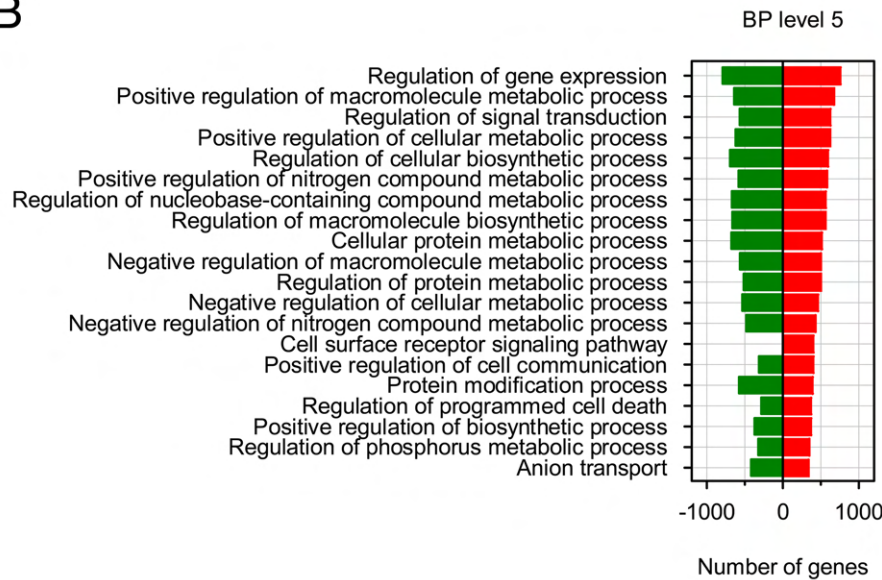




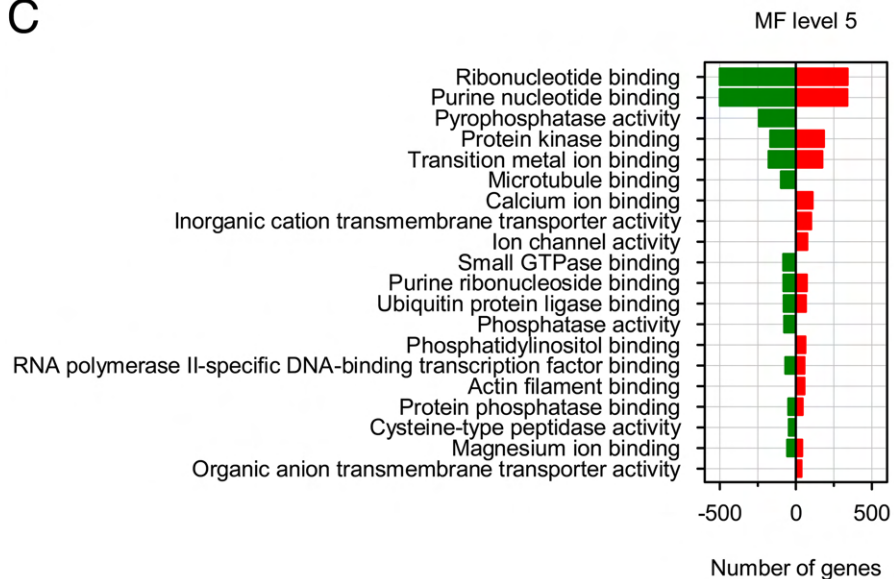
A

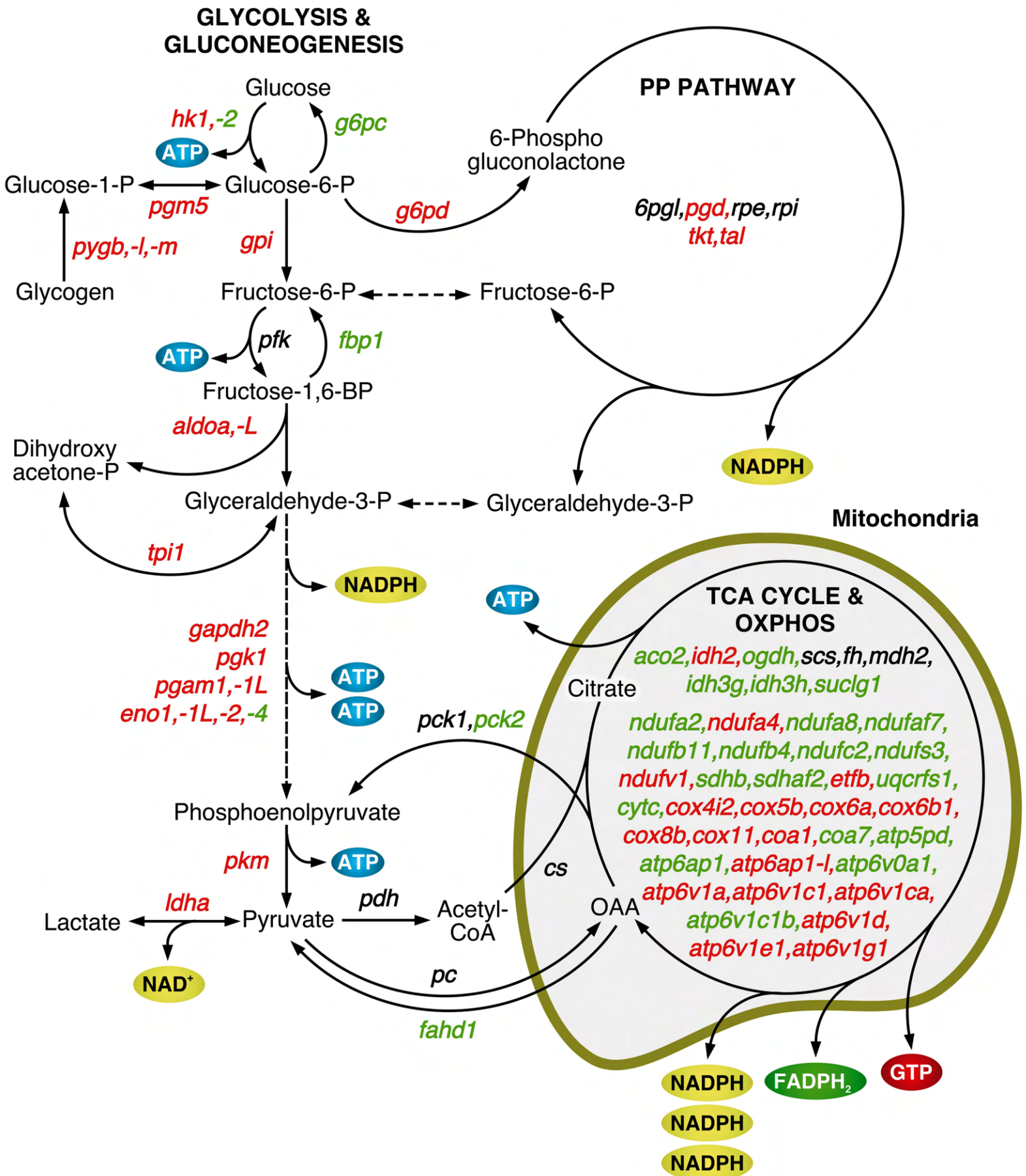


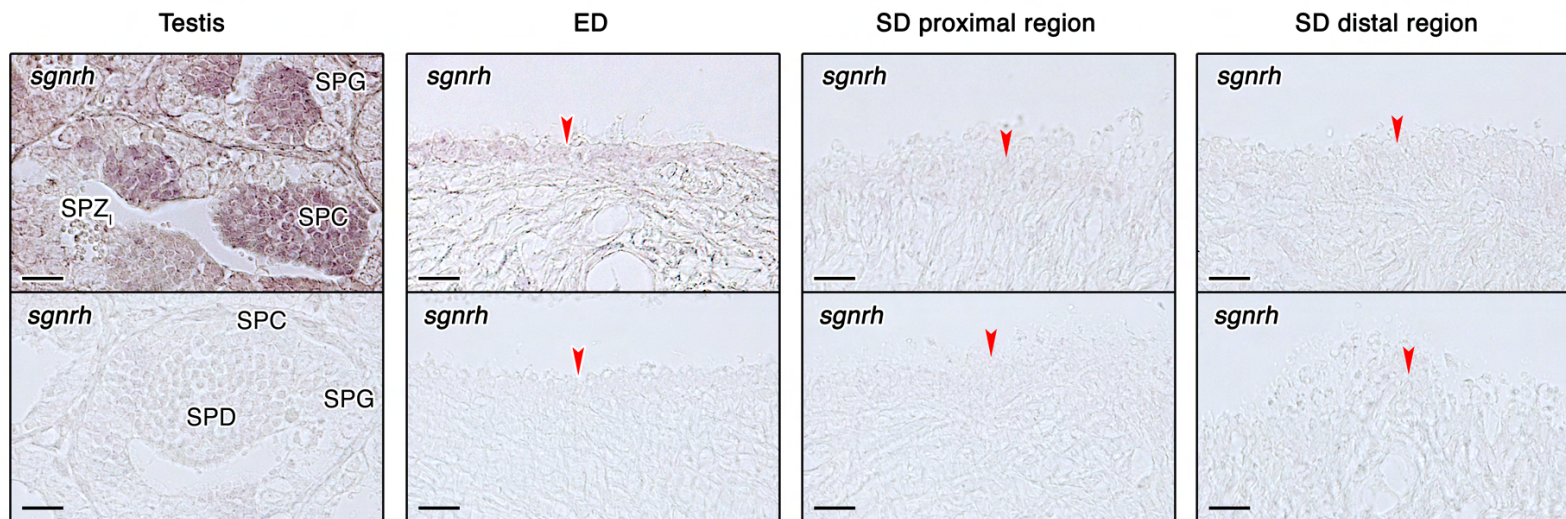
B



C





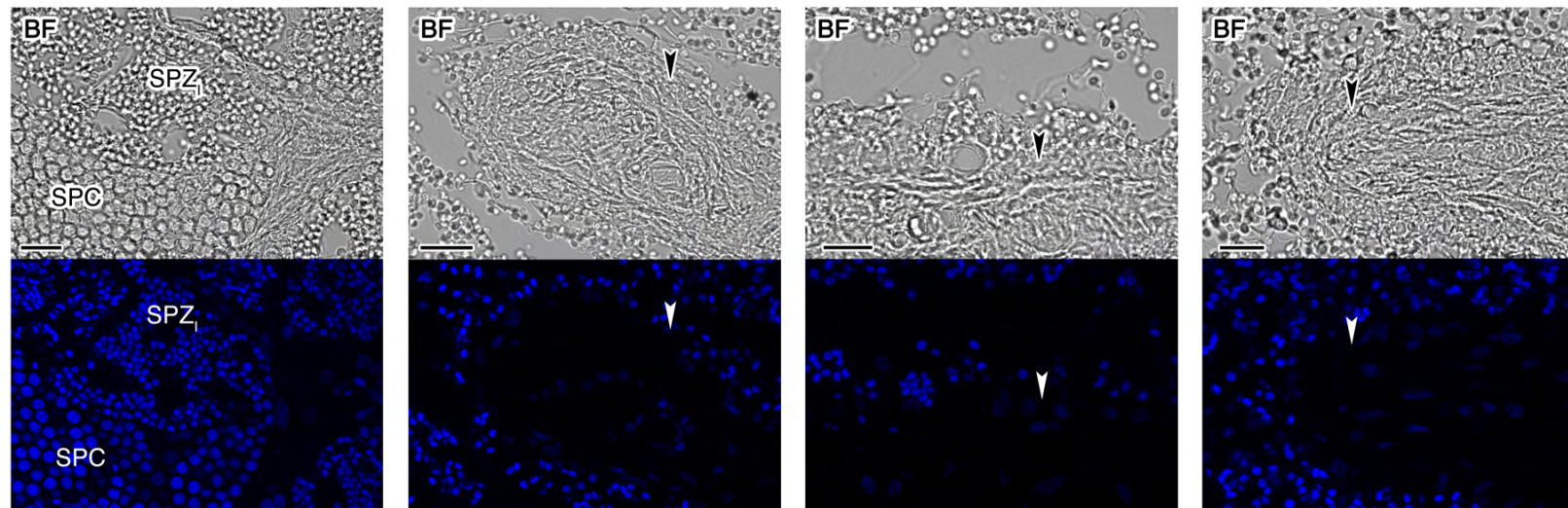


Testis

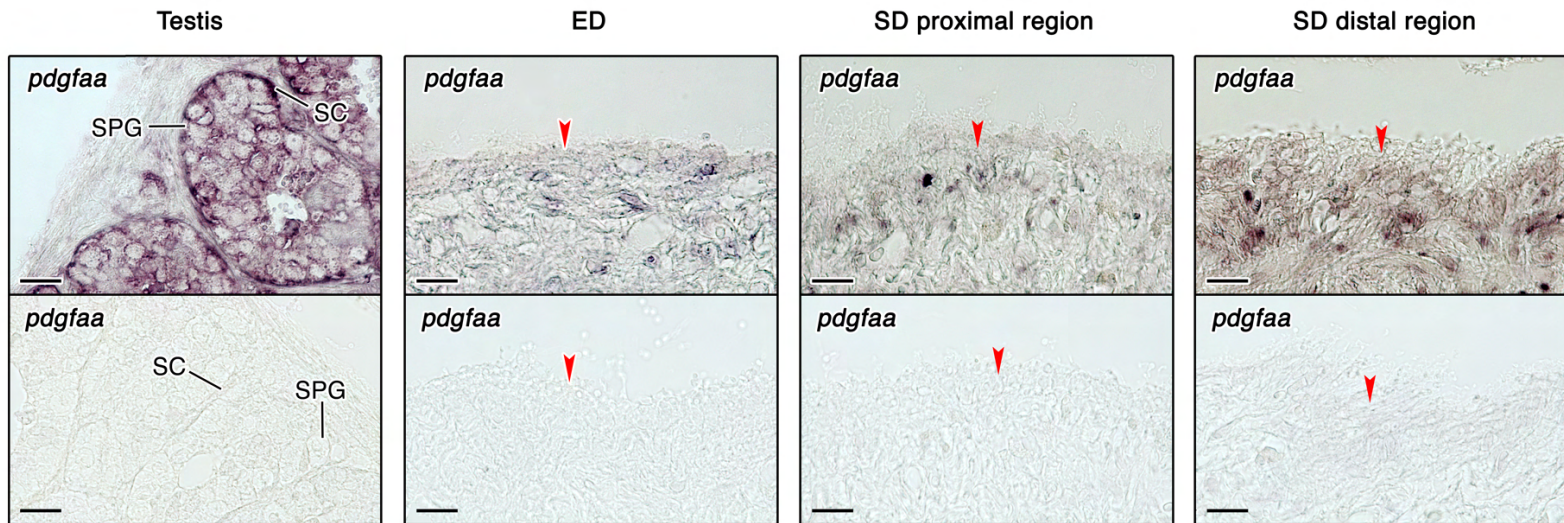
ED

SD proximal region

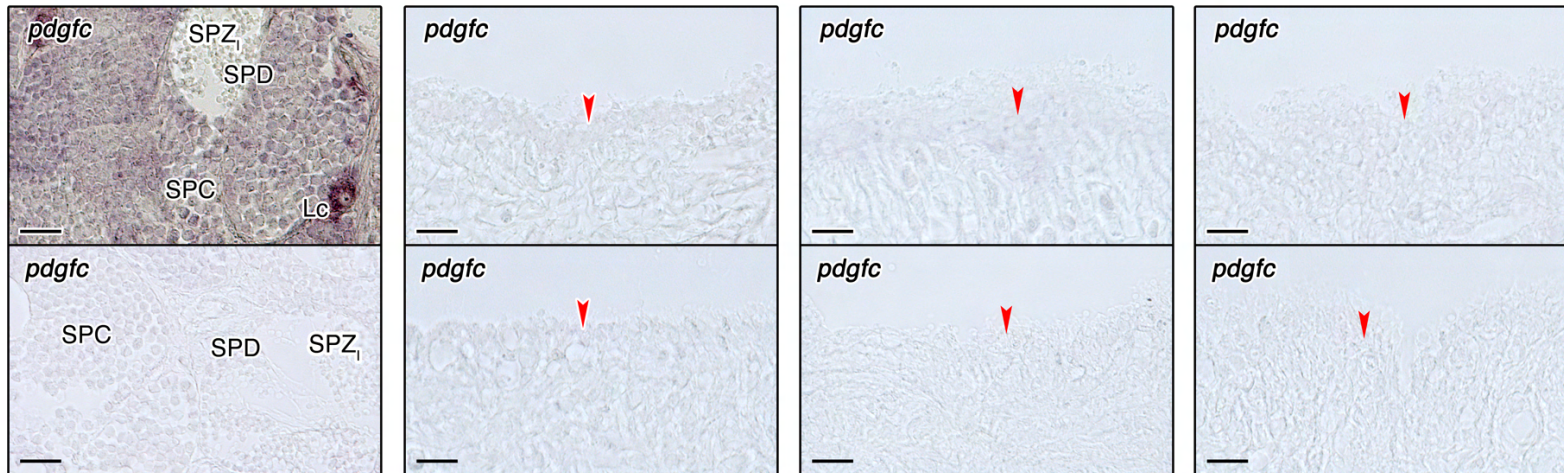
SD distal region



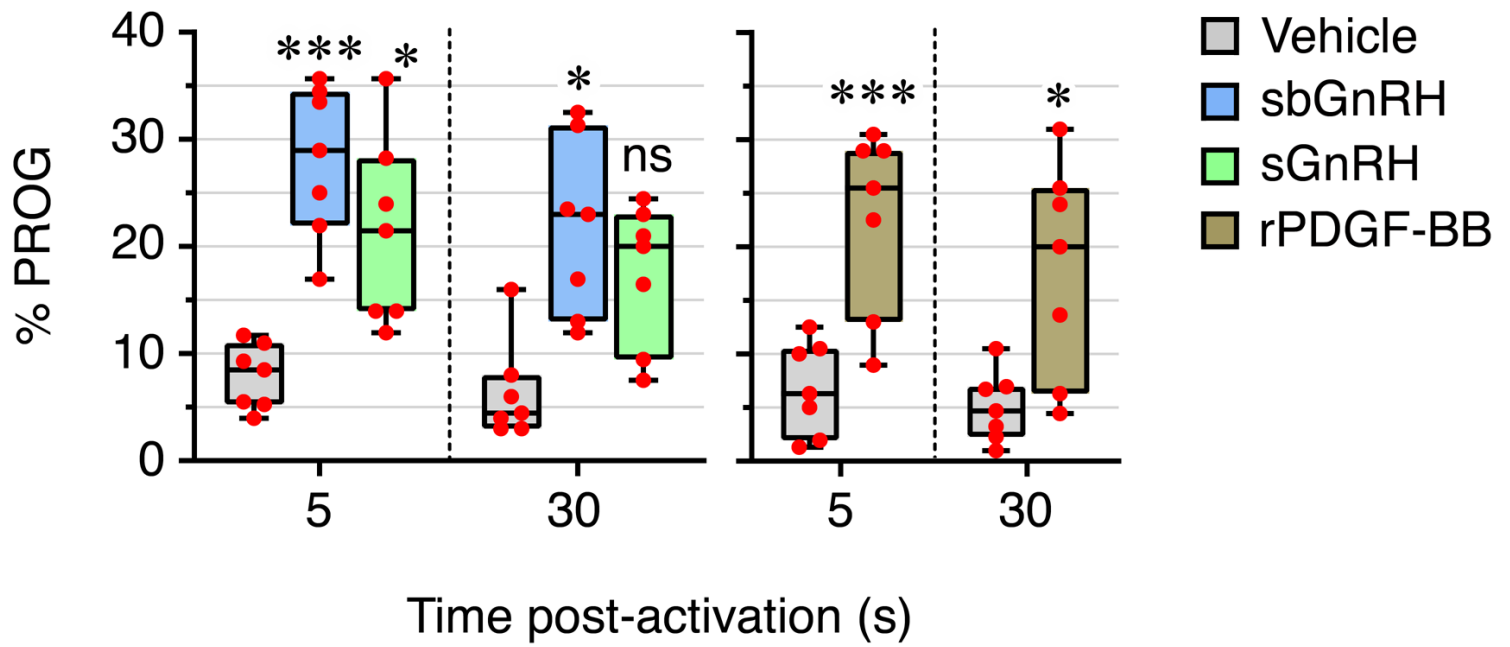
A



B



**A**



**B**

