

1 **Towards the identification of the molecular toolkit involved in scale worm bioluminescence**
2 **(Polinoidea, Annelida).**

3

4 CARLOTA GRACIA-SANCHA¹, MARÍA CONEJERO¹, SERGIO TABOADA^{1,2}, DANIEL
5 MARTÍN³, ANA RIESGO^{1,2}, MANDĚ HOLFORD⁴, AIDA VERDES^{1*}

6

7 ¹ Biodiversity and Evolutionary Biology Department, National Museum of Natural History
8 (MNCN-CSIC), Madrid, Spain.

9 ² Invertebrate Division, Department of Life Sciences, Natural History Museum, London, UK.

10 ³ Centro de Estudios Avanzados de Blanes (CEAB-CSIC), Gerona, Spain.

11 ⁴ Department of Chemistry, Hunter College, City University of New York, New York, USA.

12

13 *Corresponding author: aida.verdes@mncn.csic.es

14 **Abstract**

15 **Background:** Bioluminescence, or the ability of a living organism to produce light, has evolved
16 independently in numerous taxa inhabiting a panoply of ecosystems, although it is more frequent
17 among marine animals. Scale worms are a group of marine polynoid annelids characterized by
18 having dorsal scales, known as elytra, capable of emitting bioluminescent light by a mostly
19 unknown molecular mechanism that may involve a photoprotein called polynoidin. Here, we used
20 RNA-seq data to characterize the expression of genes potentially involved in light production in
21 the polynoid species *Harmothoe imbricata* (Linnaeus, 1767) and *Harmothoe areolata* (Grube,
22 1860) across tissues of the specimens. We also compared the transcriptomes of the selected
23 species with other bioluminescent and non-bioluminescent polynoids, to identify shared
24 orthologous genes potentially involved in light production. In addition, we investigated the
25 disposition of the photocytes on the elytra using confocal microscopy and histological analyses.

26 **Results:** Our results showed a total of 16 candidate genes, 15 orthologous genes and 12 enriched
27 GO terms potentially involved in bioluminescence, including genes related with oxidative stress,
28 cytoskeleton, nervous system, stress response, wounding response, eye constituents and metabolic
29 pathways. We also confirmed the presence of photocytes in both species, which appeared
30 distributed around the elyrophore.

31 **Conclusions:** Among the genes found potentially implicated in bioluminescence we suggest that
32 the oxidoreductase protein, peroxidasin, could be a polynoidin candidate since it appears
33 overexpressed in the elytra of both species and it is located in the endoplasmic reticulum, where
34 this photoprotein has been described to be found.

35 **Keywords:** bioluminescence, oxidoreductase, photocyte, polynoids, polynoidin.

36

37 **Background**

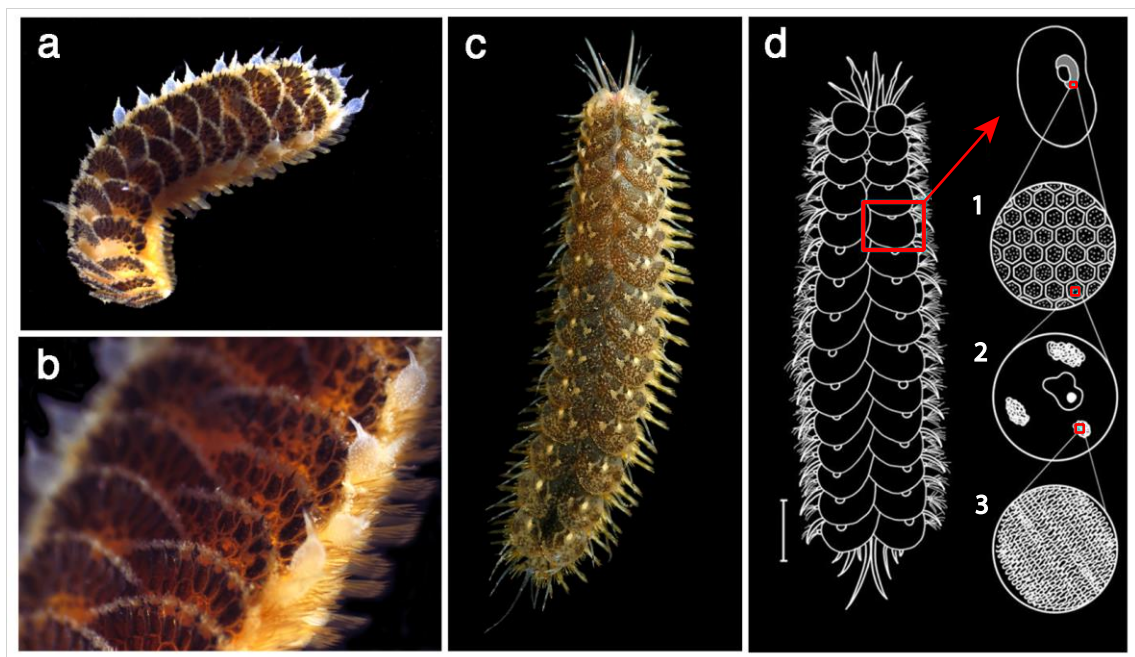
38 Bioluminescence, the ability to produce light by a living organism, is widespread through the Tree
39 of Life. The production of bioluminescence normally involves the oxidation of a light emitting
40 photoprotein, that in some organisms is called luciferin, in a chemical reaction catalysed by an
41 enzyme, called luciferase, resulting in the emission of a photon [1, 2]. Biological mechanisms
42 behind light emission have evolved independently in both terrestrial and aquatic organisms from
43 polar to tropical regions and from coastal waters to the open ocean [3], being more frequent among
44 marine dwellers [4]. There are more than 10,000 species that have disparate mechanisms for the
45 production of light [5], the majority of which are still unknown. The ability to produce light has
46 evolved independently in the different organism lineages, resulting unique mechanisms through
47 which the different taxa produce light. This has led to diversity not only in how each taxon
48 produces light but also in the wavelengths emitted and the specific functions associated with light
49 emission [4–11]. Although there is diversity in the chemical nature of the light-emitting proteins
50 responsible for the production of light, the generic terms of luciferase for the enzymes involved
51 and luciferin for the substrate have become widely accepted [3]. The most studied luciferin was
52 originally described from fireflies, and since then only a few have been characterized, some of

53 which are shared by different lineages [1, 12, 13]. The chemical reaction involved in
54 bioluminescence should be sufficiently energetic to produce an excited molecule that will
55 generate a visible photon when it relaxes, and oxidation reactions fit well into this description,
56 with the breakdown of peroxide bonds among the most widespread mechanism [3]. In this sense,
57 most animal lineages possess different kinds of luciferases for light production [5].

58 Only in the phylum Annelida, there are more than 100 bioluminescent species described
59 to date [11]. Among these, there are important differences in the chemistry of the bioluminescence
60 systems used by marine and terrestrial taxa, which suggests the independent origin of
61 bioluminescence [14–18]. Polynoids, also known as scale worms, are a group of marine annelids
62 with several bioluminescent species reported, including members of the genera *Gattyana*,
63 *Malmgrenia*, *Polynoe* and *Harmothoe* [11, 16, 19–21] and possibly *Neopolynoe* [10]. Polynoids
64 are characterized by a series of paired structures denominated scales or elytra which are arranged
65 in two rows covering entirely or partially the length of the dorsum [22] (Fig. 1). The elytra are
66 dorsal cirri modified into flat, circular disc-like plates that are attached to the body by a thin
67 layered structure called elyrophore (Fig. 1). The scales serve a wide variety of functions,
68 including egg brooding, facilitating water circulation and also enabling defensive behaviors [23,
69 24]. When threatened, luminous polynoids emit flashes of light to startle predators that come from
70 the ventral epithelium of the elytra, which has a layer of bioluminescent cells called photocytes
71 not found in non-luminous species [19, 25]. If the stimulus is strong, one or more glowing elytra
72 might be detached from the insertion zone of the elyrophore while the animal swims away [16,
73 19, 24–26]. The elytra, despite not being attached to the body anymore, continue producing
74 flashes of light for some time, catching the attention of the predator that follows the light source,
75 allowing the polynoid time to escape [24]. Under extreme danger, polynoids might autotomize
76 the entire posterior end of the body. When this happens, the elytra of the posterior end emits light
77 that attracts the predator while the rest of the animal remains dark and escapes to later regenerate
78 the missing segments [19, 27]. Therefore, bioluminescence in scale worms is thought to serve
79 mainly as a warning or distracting mechanism for defense, as elytra are easily autotomized and

80 emit bright flashes of light for some time, acting as a sacrificial lure and allowing the animal to
81 escape avoiding potential threats [1, 16, 19, 24, 28].

82 Morphologically, elytra are composed of a unicellular epidermis, which is interrupted by
83 the insertion of a muscular elyrophore, covered by a cuticular layer. A main nerve trunk traverses
84 the elyrophore branches into nerves that cover the whole structure. The efferent nerve fibres from
85 the main nerve trunk are connect to the photogenic cells, known as photocytes [19] located
86 ventrally, around the elyrophore zone. Each photocyte contains 30 to 50 fluorescent photosomes,
87 the bioluminescent organelles, arranged around the nucleus of the photocytes [29] and linked by
88 various junctions [30]. The polygonal-shaped photocytes exhibit a granular structure composed
89 of undulating microtubules (Fig. 1d), considered to produce secretions implicated in
90 bioluminescence production [29].



91

92 **Figure 1. Photographs of *H. areolata* and *H. imbricata* with a diagram representing the**
93 **bioluminescence system. (a-b).** Images of *Harmothoe areolata* with a close-up of the dorsum
94 with the elytra in detail. (c). Dorsal view of *Harmothoe imbricata*. (d). Diagram representing an
95 individual of the polynoidae family showing the disposition of the bioluminescent system. The
96 shaded zone of the elytra corresponds to the bioluminescent area. (1). Zoom of this area where
97 the photocytes are arranged in a mosaic pattern. (2). A magnified photocyte with three

98 photosomes and a nucleus. (3). The paracrystalline structure of the endoplasmic reticulum of the
99 photosomes. (Scale = 5 μ m). Modified from *Ouldali et al., 2018* [31].

100

101 Light production in scale worms involves a membrane photoprotein of about 65 kDa
102 called polynoidin, located in the membrane of the photosomes [32] and triggered by superoxide
103 or hydroxyl radicals in the presence of Ca^{2+} and Fe^{2+} [30, 32]. The oxidation of the reduced flavin
104 present in the photosomes is hypothesized to produce the superoxide or hydroxyl radicals [16, 29]
105 to trigger light production. Interestingly, polynoidin is also present in non-luminescent scale
106 worms suggesting that bioluminescence might have originated from a mechanism able to quench
107 superoxide radicals [25]. Previous studies of polynoid bioluminescence have been mostly focused
108 on the morphology and ultrastructure of the elytra, the biochemistry of light emission by
109 polynoidin, and the electrophysiology of the photogenic epithelium [19, 25, 30, 32–35]. However,
110 the genetic machinery behind bioluminescence in this group of marine annelids is currently
111 unknown. Here, we take advantage of RNA-Seq based differential gene expression analysis to
112 identify genes potentially involved in scale worm bioluminescence and investigate the
113 morphology of the elytra and the structure of the bioluminescent system in two polynoid species:
114 *Harmothoe imbricata* (Linnaeus, 1767) [36] and *Harmothoe areolata* (Grube, 1860) [37]. For this
115 purpose, we generated RNA-Seq libraries of both elytra and the posterior end of the body and
116 performed differential expression analyses to identify candidate genes potentially involved in
117 light production. Our results suggest that a potential candidate for polynoidin could be a
118 peroxidase homolog that has been found overexpressed in the elytra of both studied
119 bioluminescent species. Furthermore, numerous transcripts were identified in connection with
120 secondary aspects of the bioluminescent mechanism, encompassing those linked to a response to
121 injury and stress, as well as others correlated with the composition of the photosomes.

122

123 **Results**

124 **Transcriptomic and differential expression analyses**

125 The sequencing of the six libraries prepared for *Harmothoe imbricata* resulted in a total of 96.9
126 million raw reads. The filtered reads examined with FastQC were assembled *de novo* producing
127 a reference transcriptome of 384,174 transcripts including 303,091,081 assembled base pairs (bp)
128 and an N50 value of 1,439 [see Additional file 1]. The sequencing of the six libraries prepared for
129 *H. areolata* resulted in a total of 153.4 million raw reads. The filtered reads examined with FastQC
130 were assembled *de novo* producing a reference transcriptome of 513,553 transcripts including
131 377,970,784 assembled base pairs (bp) and a N50 value of 807 [see Additional file 1]. The
132 completeness of the transcriptomes based on BUSCO values was high compared to that of other
133 transcriptomic analyses of annelids [38].

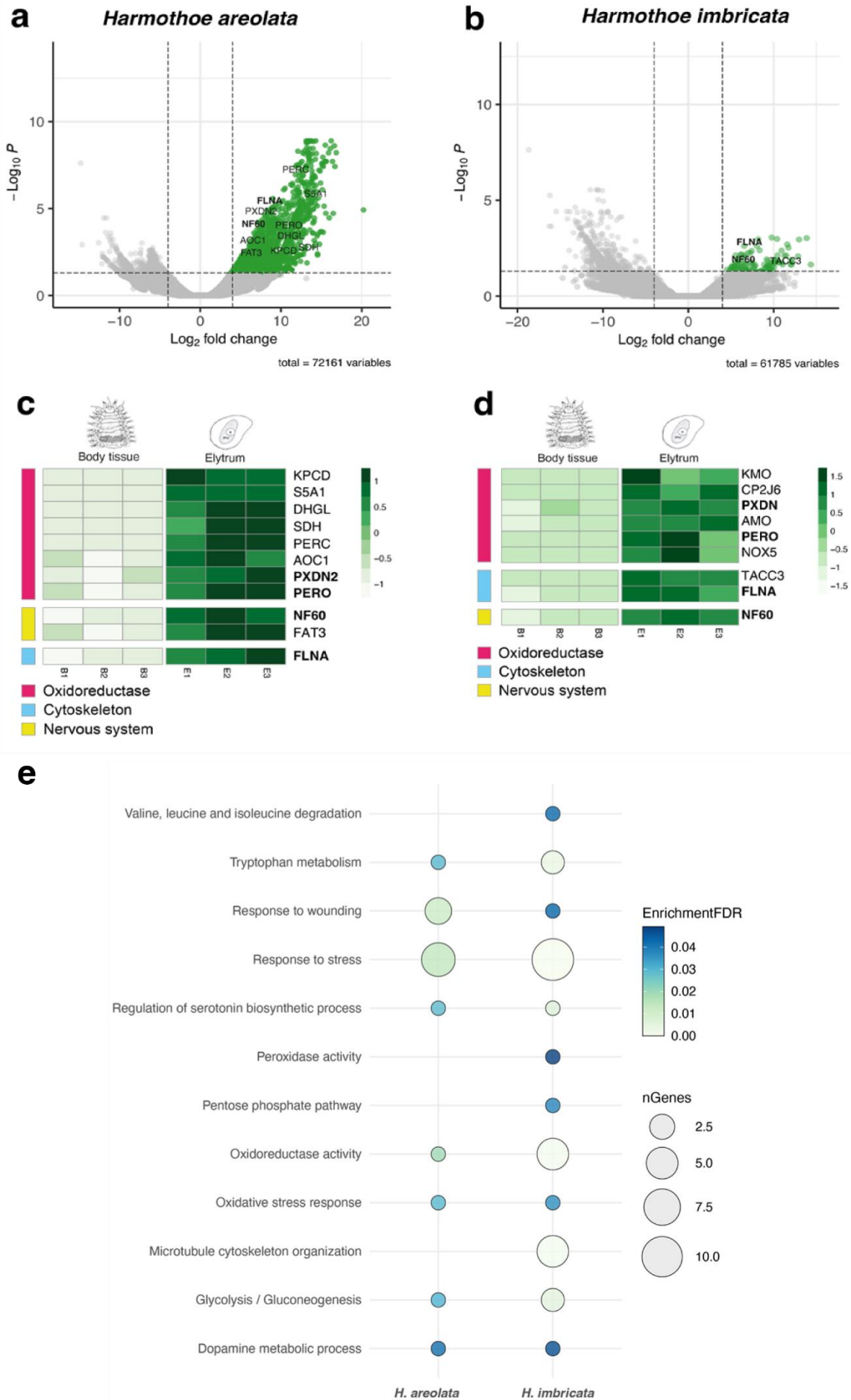
134 The differential gene expression (DGE) analysis for *H. imbricata* resulted in a total of
135 61,785 transcripts differentially expressed, of which 331 were upregulated in the elytra (Fig. 2a).
136 Of the upregulated transcripts, only 111 could be functionally annotated [see Additional file 2].
137 To identify genes potentially involved in bioluminescence production, we focused our
138 downstream analyses only on the annotated transcripts differentially expressed in the elytra. The
139 DGE analysis for *H. areolata* resulted in a total of 72,161 transcripts differentially expressed [see
140 Additional file 2], of which 475 were upregulated in the elytra (Fig. 2b). Of the upregulated
141 transcripts, only 65 could be functionally annotated, and again, we focused on those to identify
142 transcripts potentially involved in bioluminescence.

143 Of the 331 transcripts upregulated in the elytra of *H. imbricata*, only 3 were identified as
144 potentially involved in the bioluminescence production. The annotation of two of these transcripts
145 was related with the cytoskeleton: TACC3 (transforming acidic coiled-coil-containing protein 3)
146 [39] and FLNA (filamin-A) [40]. The remaining transcript blasted against a gene with functions
147 associated with the nervous system, NF60 (neurofilament protein)[41]. In *H. areolata* 475
148 transcripts were upregulated in the elytra and we selected 11 as possible candidates involved in
149 the bioluminescence process. Several of these transcripts had annotations related to
150 oxidoreductase function: AOC1(amiloride-sensitive amine oxidase) [42], DHGL (Glucose
151 dehydrogenase) [43], PERC (Chorion peroxidase)[44], PERO (Peroxidase) [45], PXDN2
152 (Peroxidasin homolog) [46], S5A1 (3-oxo-5-alpha-steroid 4-dehydrogenase) [47] and SDH (L-

153 sorbose 1-dehydrogenase) [48]. Another transcript was related with oxygen radical production,
154 KPCD (Protein kinase C delta type) [49]. We also found transcripts annotated to functions in the
155 nervous system, NF60 (neurofilament protein), shared by both species, and FAT3 (Protocadherin
156 Fat 3) [50]. Finally, we found a cytoskeleton related gene, FLNA, which also appeared
157 upregulated in *H. imbricata*.

158 Heat maps in Figure 2 (c–d) illustrate the expression levels of the transcripts selected in
159 the elytra and the body of both species. We found small differences in the expression profiles of
160 the replicates of each tissue but large differences between tissues, with higher levels of expression
161 in the elytra. For *imbricata*, 6 transcripts were additionally included, given their high expression
162 in the elytra (although they were not differentially expressed). These additional transcripts were
163 annotated to oxidoreductase function: AMO (Putative amine oxidase) [51], KMO (Kynurenine 3-
164 monooxygenase) [52], NOX5 (NADPH oxidase 5) [53] and CP2J6 (Cytochrome P450) [54].
165 Inside this group we include PERO and PXDN which are also found upregulated in the elytra of
166 *areolata*.

167 Besides oxidation-reduction processes, in our analyses, we retrieved functional
168 information from 1,149 transcripts of those found upregulated in the elytra in *H. imbricata* and
169 594 in *H. areolata*, which allowed us to identify various enriched GO terms potentially relevant
170 for bioluminescence (Fig. 2e). The shared pathways with a possible relation with bioluminescence
171 between the two species were: tryptophan metabolism, response to wounding, response to stress,
172 regulation of serotonin biosynthetic process, oxidoreductase activity, oxidative stress response,
173 glycolysis/gluconeogenesis and dopamine metabolic process (Fig. 2e). For *H. imbricata* we found
174 four additional pathways that could be involved in bioluminescence: valine, leucine and
175 isoleucine degradation, peroxidase activity, pentose phosphate pathway and microtubule
176 cytoskeleton organization. For more information of the enriched GO terms [see Additional file
177 3].



179 **Figure 2. Differential expression and gene enrichment in the elytra of *Harmothoe areolata***
180 **and *H. imbricata*. (a–b).** Volcano plots displaying the $-\log_{10}$ p value (false discover rate [FDR])
181 as a function of fold change in the scale and the body of *H. areolata* (a) and in *H. imbricata* (b).
182 Grey dots in the left represent upregulated genes in the body while green dots at the right
183 correspond to upregulated genes in the elytrum. (c–d). Heatmaps showing selected upregulated
184 genes in the elytrum with their associated functions in *H. areolata* (c) and in *H. imbricata* (d).
185 (e). Gene Ontology enrichment analysis of the upregulated annotated genes in the elytra of both
186 species.

187

188 **Identification of shared orthologs between bioluminescent species**

189 The reference transcriptomes of the elytra of *H. imbricata* and *H. areolata* were compared with
190 other bioluminescent and non-bioluminescent polynoid species. With this analysis, we identified
191 orthologous genes shared by bioluminescent species, and present in the elytra of the selected
192 species, but not present in non-bioluminescent species. We focused on the clusters of orthologous
193 genes shared only between the bioluminescent species to identify genes potentially involved in
194 bioluminescence.

195 All the polynoid species included in the analysis shared a total of 2,994 clusters while *H.*
196 *extenuata*, *H. imbricata* (GenBank), and the elytra of *H. imbricata* and *H. areolata* shared 400
197 clusters of orthologous genes [see Additional file 4]. Of those 400 shared clusters, we identified
198 15 that could be potentially related to bioluminescence and are listed in Table 1. Interestingly, the
199 Orthovenn analyses revealed a total of 7 clusters related with an oxidoreductase function shared
200 among the bioluminescent polynoids (i.e. Amine metabolic process, hydroquinone, 3-
201 hydroxybutyrate dehydrogenase activity, positive regulation of lipid biosynthetic process, l-
202 ascorbic acid biosynthetic process, thioredoxin-disulfide reductase activity and octopamine
203 biosynthetic process). One cluster exhibited a connection to microtubule organization (i.e. Actin
204 binding), while two clusters were linked to the nervous system, sharing the same GO annotation
205 (i.e. Synaptic transmission, cholinergic), and another two were associated with response to
206 oxidative stress (i.e. Sulfur amino acid metabolic process and regulation of heme biosynthetic

207 process). Furthermore, three clusters demonstrated similarities to genes involved in visual sense
208 (i.e. Estructural constituent of eye lens, retinol metabolic process and visual perception).

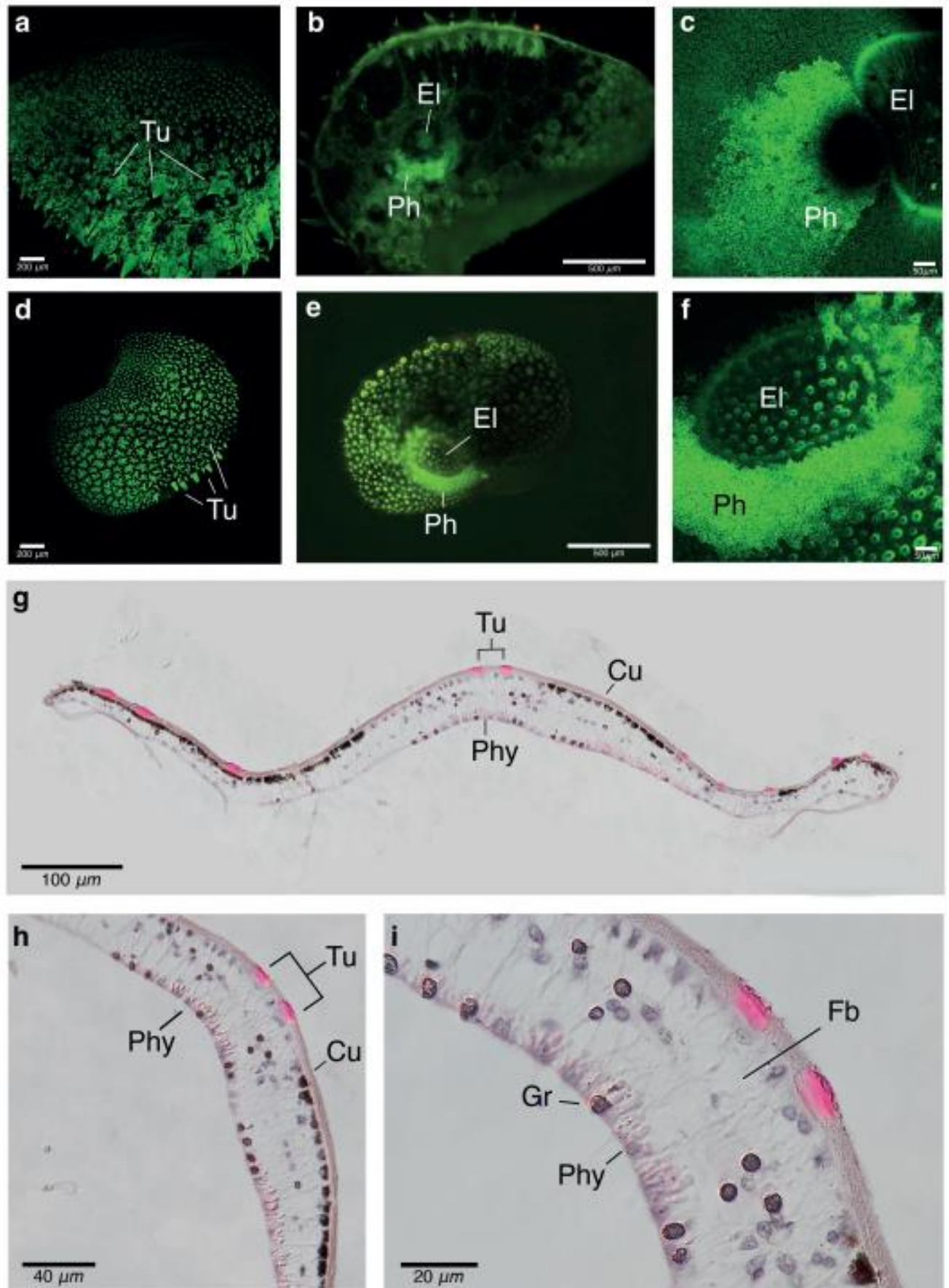
209

210 **Histology and tissue organization of elytra**

211 The excitation light used to image the fluorescent photocytes was 405 nm and the fluorescent
212 structures of the elytra emitted within a range spectrum of 517 nm and 570 nm. Figure 3 shows
213 an image of the complete structure of the elytra in *H. imbricata* (Fig. 3a) and *H. areolata* (Fig.
214 3d) taken from a ventral position. It is known that tubercles covering the surface are
215 autofluorescent, as well as the photocytes [25]. Although the tubercles were located in the dorsal
216 part of the elytra, their autofluorescence was visible from the ventral part. The tubercles covering
217 the dorsum in *H. areolata* were much more conspicuous and with spines, while in *H. imbricata*
218 are rounded (Fig. 3a–f). In figures the ventral side, surrounding the elyrophore we observed a
219 dense mass of brighter fluorescent cells (Fig. 3b–c, e–f). These cells corresponded to the
220 fluorescent organelles of the photocytes, the photosomes, which are distributed within the
221 photogenic area (Fig. 3b–c, e–f). Normally there are 30-50 photosomes per photocyte, with a
222 diameter of 1 - 5 micrometres [29].

223 In the transversal histological sections of the elytra of *H. imbricata* (Fig. 3) we observed
224 several tubercles in the dorsal side (Fig. 3g). Beneath the cuticle we found a continuous upper
225 and lower monostratified epithelium. On the upper layer of the epithelium, we found cell bodies,
226 some of which contained pigment granules (Fig. 3i), nerves (*Fb*) connecting the cuticle with the
227 lower epithelium in a ganglion near the elytral stalk [55], although we could not identify the
228 ganglia in our sections. The lower epithelium was similar to the upper one, with the exception of
229 the presence of photocytes, which were modified epithelial cells, arranged near the elyrophore
230 (Fig. 3i). The insertion of the elyrophore was not clearly identified, although we observed
231 photocytes not following a continuous pattern (Fig. 3h). Photocytes appeared as calyx-shaped
232 cells (Fig. 3i) similar to what has been observed in other species [10].

233



234

235

236 **Figure 3. Confocal and light microscopy images of the elytra of *H. imbricata* (a–c) and *H.***
237 ***areolata* (d–f).** Absorption spectrum between 517 and 570 nm. (a). Dorsal picture of the whole
238 elytrum of *H. areolata* where dorsal tubercules (Tu) are observed as spines. (b). Ventral picture
239 of the whole elytrum of *H. areolata* with photosomes (Ph) arranged concentrically around the
240 elyrophore (El). (c). Detail of the photogenic area with the photosomes (Ph) in *H. areolata*
241 surrounding the elyrophore (El). (d). Dorsal picture of the whole elytrum of *H. imbricata* where
242 dorsal tubercules (Tu) are observed as small circles. (e). Ventral picture of the whole elytrum of
243 *H. imbricata* with photosomes (Ph) arranged concentrically around the elyrophore (El). (f).
244 Detail of the photogenic area with the photosomes (Ph) surrounding the elyrophore (El). (g–i).
245 Transversal Histological section of an elytrum of *H. imbricata* stained with haematoxylin-eosin
246 and mounted in DPX. (g). Picture of the elytrum the cuticle (Cu) and tubercules (Tu) appearing
247 dorsally and with photocytes (Phy) appearing ventrally. (h). Same section at a higher
248 magnification showing photocytes (Phy) at the ventral side breaking their continuity due to the
249 elyrophore. (i). Detail of the photocytes (Phy) where the calix shape of these cells is observed.
250 Pigmented granules (Gr) and fibres nerves (Fb).

251

252 **Discussion**

253 Our study aimed to identify the genetic machinery involved in the bioluminescence of the
254 polynoid worms *Harmothoe imbricata* and *Harmothoe areolata*. The search of candidate genes
255 with the different analyses led us to the identification of various transcripts in both species that
256 could potentially be involved in light production and additional processes associated with
257 bioluminescence. The DGE analysis allowed us to identify genes that were upregulated in the
258 elytra, the structures where the photogenic cells are located, and thus may be related to
259 bioluminescence production. Additionally, our comparative transcriptomic analysis allowed us to
260 identify orthologous genes shared only by bioluminescent species, which might also be involved
261 in light production. Finally, the morphological analyses with light and confocal microscopy
262 revealed the location and structure of the bioluminescent system within the elytra of both polynoid
263 species.

264 **Genes potentially involved in the bioluminescence chemical reaction**

265 Bioluminescence in polynoid worms takes place in the elytra. More specifically, this process
266 happens in the photosomes of the bioluminescent cells present in the photogenic area, thanks to
267 the photoprotein polynoidin which is activated by the emission of superoxide or hydroxyl radicals,
268 Ca^{2+} and Fe^{2+} [30, 32]. We identified several candidate transcripts and pathways involved in
269 bioluminescence in *H. imbricata* and *H. areolata*, sharing a total of 9 pathways between the two
270 species (Fig. 2e). A total of 11 genes were found differentially expressed and upregulated in the
271 elytra of both species involved in an **oxidoreductase** function. These were identified as the most
272 plausible candidates for luciferase function, given that the primary mechanism operating in
273 bioluminescence involves the breakdown of a peroxide bond [11]. In this sense, one of the most
274 accepted theories of the origin of bioluminescence is the antioxidative hypothesis, and it refers to
275 the evolutionary history of the bioluminescent substrate. Since luciferin has antioxidative
276 properties and high reactivity with reactive oxygen species (ROS), the antioxidative hypothesis
277 suggests that originally, the luciferin was a molecule implicated in reducing oxidative stress [56].
278 In marine ecosystems, organisms living in shallow waters attempting to escape from predators,
279 might have migrated to the deep ocean where oxidative stress is reduced, leading to a functional
280 change from the antioxidative to the chemiluminescent properties of this molecule [56]. These
281 proteins involved in oxidoreductase processes included Putative Amine Oxidase, Kynurenine 3-
282 monooxygenase, NADPH Oxidase 5 and Cytochrome P450 for *H. imbricata*, while for *H.*
283 *areolata* these were Amiloride-sensitive Amine Oxidase, Glucose dehydrogenase, Chorion
284 peroxidase, 3-oxo-5- α -steroid 4-dehydrogenase, L-sorbose 1-dehydrogenase and Protein
285 Kinase C (Table 2, Additional file 2). Interestingly, both species shared two genes, a peroxidase
286 and a peroxidasin homolog. Oxidoreductase enzymes are involved in controlling the amount of
287 oxygen radicals or reduced compounds that appear in the cells [57]. It is known that the
288 bioluminescent reaction in polynoids is triggered by superoxide radicals that result from the
289 oxidation of riboflavin [16]. Many of the genes selected have NAPH oxidase/dehydrogenase
290 function which are considered one of the major sources of reactive oxygen species [53].
291 Oxidoreductases are also involved in many other bioluminescent systems including fireflies [58],

292 bacteria [59], fungi [60] and other annelids [61]. Oxygen has been proved to be essential for the
293 luminescent reaction in polynoids [32], therefore, we hypothesize that the oxidoreductase
294 homologs identified here represent candidate polynoidins, the protein responsible of catalyzing
295 the bioluminescent reaction in polynoids. The two only genes that appeared shared by both species
296 are peroxidase (PERO) and peroxidasin (PXDN). The appearance of these genes in the two
297 bioluminescent species studied reinforces the possibility that one of them may be the elusive
298 polynoidin. Moreover, given that the peroxidasin homolog shared is found in endoplasmic
299 reticulum, where polynoidin has been described to be found, this gene appears as the most
300 possible candidate to catalyze the light production in *Harmothoe imbricata* and *H. areolata* [32].

301 Interestingly, two of the orthologous clusters shared by bioluminescent species and one
302 of the upregulated genes in the elytra of both species (NADPH oxidase 5) are involved in
303 **oxidative stress responses**. Oxidative stress is generated by reactive oxygen species [62],
304 between others, which could finally produce bioluminescence in polynoids. It is possible that
305 initially, bioluminescence occurred as an antioxidant mechanism in order to avoid oxidative stress
306 [56]. This response appearing in the genes overexpressed in the elytra confirms the idea that there
307 are products in the cells of the polynoids capable of acting against this oxidative stress and thus,
308 potentially leading to the production of light.

309

310 **Genes potentially involved in additional aspects of the bioluminescence process**

311 Besides the potentially direct production of light of the genes commented above, we identified a
312 battery of genes indirectly related to the bioluminescence process, including those potentially
313 involved in the formation of bioluminescent structures, or with the nervous signal triggering the
314 bioluminescence. Among them, we found two genes differentially expressed in the elytra of the
315 two worms and one shared orthologous cluster associated with the **cytoskeleton**. For both species
316 Filamin-A was upregulated in the elytra. But also, *H. imbricata* showed another gene upregulated
317 in their elytra, Transforming Acidic coiled-coil-containing Protein 3. The cytoskeleton plays an
318 important role in the formation of the photosomes as it is its main structural component. The
319 granules of the photosomes are composed of undulating microtubules that appear regularly

320 arranged in a paracrystalline organization continuous with the endoplasmic reticulum [29]. It has
321 been described that the photocytes of fireflies are also connected to the endoplasmic reticulum,
322 suggesting that this organelle plays a role in the light emission [63, 64]. The task carried out by
323 the gene found homolog to the Transforming Acidic coiled-coil-containing Protein 3 consists in
324 stabilizing the microtubules in the kinetochore fibers of the mitotic spindle [65]. This gene,
325 therefore, could be related with the formation of the paracrystalline organization which conforms
326 the photosomes where the bioluminescent reaction takes part. Filamin-A and the Actin binding
327 cluster, which are shared by both polynoids, are not related with the microtubule organization,
328 but with the actin cytoskeleton. Although the main structural component of the photosomes is
329 indeed the microtubule cytoskeleton, the intermediate filaments of the cytoskeleton, that are
330 formed of actin, are one of the principal elements of the desmosome junctions that join the
331 photosomes inside the photocytes [30]. It is also remarkable that in dinoflagellates, filaments of
332 the actin cytoskeleton are involved in the mechanotransduction of bioluminescence [66].
333 Consequently, these genes could potentially play an important role in the formation of the
334 photosomes as they have been proved to have relationships with their main structure and with the
335 junctions between them.

336 One of the most enriched GO categories in both species was **stress response**. The
337 mechanism of bioluminescence in polynoids has hypothesized to be a defensive strategy to escape
338 from predators [67]. It seems obvious that there must be a toolkit to perceive stress and respond
339 creating an impulse ending in light emission to escape. In other studies, similar analyses have also
340 revealed genes overexpressed in the luminal tissue with stress response functions in Terebellidae
341 species [68]. Our study suggests that these genes present upregulated in the elytra of the polynoid
342 worms might work under stress conditions when the animal is trying to escape from its predators.
343 [16][31, 33] We also found several genes related to the **nervous system** that are potentially
344 involved in the bioluminescence process. Bioluminescence in many taxa is controlled by the
345 nervous system, and this includes polynoids. Each elytrum has a central ganglion located near the
346 elytophore from which nerves branch into nerve fibers in various directions covering the
347 periphery of the elytra [19]. In its upper part, sensory fibers connect with the tubercles or papillae

348 of the cuticle and in its lower part, efferent fibers travel until they reach the photocytes at the base
349 of the elytrum. The stimulus, thus, starts at the upper part and it is transmitted to the central
350 ganglion to reach the photocytes [69]. We identified two upregulated genes in the elytra,
351 Protocadherin Fat 3 and Neurofilament protein and as well as several shared orthologous clusters
352 involved in cholinergic synaptic transmission. Protocadherins homolog, present only in *areolata*,
353 are transmembrane proteins specialized in signaling. Specifically, fat cadherins are defined to
354 enable influence multiple features of neuronal development [50]. This Protocadherin may be
355 involved in the establishment of neuronal connections related to the bioluminescence signal, as
356 these proteins normally binds calcium. Calcium is one of the ions starting the action potential of
357 the luminous process [70] and it could be related to the transmission of the nerve impulse
358 triggering bioluminescence in *H. areolata*. The neurofilament protein, found upregulated in the
359 elytra of both species, is commonly present in neuronal tissue as it is the cytoskeletal constituent
360 of neurons [71]. As it appears overexpressed in both species, this gene may be in fact, one of the
361 genes being responsible of the formation of neurons performing in the elytra. A cholinergic
362 synapse is a chemical synapse that use acetylcholine molecules as neurotransmitter. Nicolas and
363 collaborators in 1978 specified that a cholinergic mechanism was indirectly involved in the
364 luminescence control, and previously it was confirmed a positive acetylcholinesterase
365 histochemical reaction of the elytral ganglion and nerves [72]. All the above mentioned genes
366 may be involved in the neurotransmission necessary for emission of light in polynoids.

367 Another function indirectly related with the process of bioluminescence is wounding,
368 which its related GO term, **wounding response**, was enriched in the upregulated component of
369 the elytra. When light is being produced, the polynoids detach their luminescent elytra while they
370 are glowing. When this occurs, the luminescent elytra keep on glowing as the same time that the
371 animal tries to scape [24]. The autotomization of the elytra leaves a wound in the worms, that is
372 repaired in the next 15 days, when a new elytrum is grown [73]. Genes capable of responding to
373 this loss of the elytra must be in the elytrum itself, as it must be repaired in order to serve the
374 polychaeta a defensive function.

375 Three orthologous clusters appeared shared by the bioluminescent species related to
376 **visual perception**. At first it seems strange that genes related to vision appear in the elytra of the
377 polynoids. We must remember that orthologous genes were obtained with the transcriptome of
378 the elytra of our target species and the whole transcriptome of *H. extenuata*. These results suggest
379 that there are homologous genes that are expressed both in the body of the bioluminescent species
380 and in their elytra. With these results we can verify genetically the study carried out by Bassot
381 and Nicolas in 1978 where they confirm that the paracrystalline endoplasmic reticulum
382 characteristic of the photogenic cells in polynoids appears also in the eyes, where it constitutes a
383 greater part of the lens. The relationship that exists in the appearance of the same structures in the
384 photoreceptors as in the photosomes of polinoid species is a mystery to be studied.

385 Two metabolic pathways related with glycolysis and tryptophan were enriched in the
386 elytra, and even though they have not been reported in bioluminescent systems of polynoids, they
387 have been proved to be involved in other bioluminescent species. The process of obtaining energy
388 out of the glycolysis metabolism, provides energy to an inhibitory mechanism that maintains the
389 photophores in a non-luminescent state in the teleost fish *Porichthys* [74]. In a more closely
390 related species, the earthworm *Lampito mauritii* [75], this metabolic pathway has also been shown
391 to activate a greenish light emission, when the celomic cells where exposed to stimulants and
392 some products of gluconeogenesis [76]. This could indicate that the glucose pathway may
393 interfere in the system of activation and deactivation of light production by polynoids. Also,
394 tryptophan metabolism is central to produce tryptophan compounds, which are one of the main
395 elements of the sea firefly luciferin, cypridinid luciferin [77, 78]. Polynoid's luciferin has not
396 been discovered yet and since the tryptophan metabolic pathway has been found to be
397 overexpressed in the elytra of polynoids, we propose that tryptophan or its derivatives may be
398 part of the composition of polynoid luciferin.

399 Finally, we retrieved two more enriched GO categories that were only present in *H.*
400 *imbricata*: valine, leucine and isoleucine degradation and pentose phosphate metabolism. The
401 generation of NADPH molecules is essential for the bioluminescent process in Polynoidae, as
402 these coenzymes provide reactive oxygen species [16]. Earlier, we indicated the importance of

403 oxidative stress in the reaction of luminescence. Pentose phosphate pathway and valine, leucine
404 and isoleucine degradation are both metabolic pathways in which NADPH is generated [79, 80].
405 Our study suggests that these pathways may be functioning to generate compounds necessary for
406 beginning the bioluminescent reaction in *H. imbricata*.

407

408 **Morphology of the bioluminescent system**

409 Light production in polynoid worms takes place in the photosomes, organelles located in
410 specialized cells called photocytes and found in the elytra. The photosomes are autofluorescent
411 and thus are responsible for the fluorescent properties of elytra, having the ability of emit light
412 under specific excitation wavelengths. In addition, the tubercules also show this autofluorescence.
413 First, it was thought that this property was only activated after the bioluminescent reaction was
414 triggered [81] but it has been proved that this fluorescence is permanently present in these
415 structures [25, 35]. Figure 3 shows images taken from individuals preserved in KINFix after three
416 months, confirming therefore that they effectively do not lose their autofluorescence and that it is
417 present independently of bioluminescence light emission. We confirmed that the excitation
418 wavelength that works best to see the arrangement of bioluminescent structures is UV light, as
419 suggested in previous studies [25, 82]. The green light emission spectrum (517-570 nm) is
420 consistent with that reported for this 20 and other species in previous articles (510-525 nm) [10,
421 16, 32, 83]. We observed two distinctive green autofluorescent zones: the papilla and tubercules
422 and the area around the insertion zone of the elytraphore. The disposition of the photosomes can
423 be observed more accurately on figure 3c and 3f, where they appear as a dense, tangled mass,
424 which may be due to the paracrystalline form of the endoplasmic reticulum from which they are
425 formed. Our histological sections show the disposition and structure of the photosomes, arranged
426 in the centre of the ventral side of the elytra (Fig. 3g, h, i). They are elongated cells that appear
427 together and close to one another. Each photocyte is connected by efferent fibres to the central
428 ganglion of the elytra, responsible for the transmission of the impulse that triggers
429 bioluminescence. These nerves are visualized as the thin transversal lines connecting both
430 epithelia (Fig. 3i). This innervation can also be seen in non-bioluminescent species, although they

431 lack photosomes [84]. In general, our morphological study is consistent with the organization of
432 light producing organs in other polynoids [10, 19, 25, 35, 82, 85].

433 Although the light production ability of the polynoid photoprotein system could not be
434 assessed, we identified an array of proteins involved in oxidoreductase processes that point to
435 both luciferase and luciferin proteins. Peroxidasin might correspond to the elusive polynoidin, the
436 photoprotein responsible of catalyzing bioluminescence reaction in polynoids. We have seen that
437 the oxidoreductase most likely to be polynoidin is a **peroxidasin homolog**, since it is
438 overexpressed in the elytra of both species and is found in the endoplasmic reticulum, which is
439 an essential part of the structure of photocytes and where this photoprotein has been described to
440 be found. Functional validation analysis must be done to confirm the role of this protein in the
441 bioluminescent system of *Harmothoe*. It will be necessary to search for this supposed polynoidin
442 in other bioluminescent polynoids to verify its role in the bioluminescence process within this
443 genus.

444

445 **Methods**

446 **Sample collection and preservation**

447 Six specimens of *Harmothoe imbricata* and six of *Harmothoe areolata* were collected by SCUBA
448 diving in O’Grove (42.498448, -8.865719, Galicia, Spain; 12 m depth) and near Blanes
449 (41.673245, 2.802646, NW Mediterranean Sea; 10 m depth) respectively, and kept alive in
450 seawater for transport to the laboratory. Taxonomic identification was done following the original
451 description of the species and relevant additional literature [36, 86] under a Zeiss Stemi 2000
452 stereoscope up to 35x magnification by experts on annelid taxonomy. Before sample preservation,
453 bioluminescence emission was confirmed in a dark room by disturbing the individual with
454 tweezers. For transcriptomic analyses, we preserved the specimens in RNAlater (Life
455 Technologies), during 24h at 4 °C, replaced it once, and stored samples at -80 °C until further
456 processing. For histological analyses, samples were preserved in 2.5% glutaraldehyde in 1M PBS
457 and 0.34M NaCl and KINFix, which allows the sample to be used in examination by histological
458 or immunohistochemical techniques [87].

459

460 **RNA extraction, cDNA library preparation, and sequencing**

461 For RNA extraction, dorsal elytra and a portion of the posterior end of the body were dissected
462 from three specimens of each of the two species, totalling three replicates per tissue and per
463 species. Total RNA was extracted from the elytra and from the posterior end of the body with
464 TRIzol (Invitrogen), following the manufacturer's instructions. Total RNA from each of the six
465 biological replicates of *H. areolata* was sent to New York University Center for Genomics and
466 Systems Biology (New York, USA) for cDNA library preparation and sequencing. A total of
467 twelve cDNA libraries were generated with Illumina TruSeq library prep kit v2, corresponding to
468 the three biological replicates of each tissue type. We verified the quality and integrity of the total
469 RNA and the cDNA library with an Agilent BioAnalyzer (RNA and high sensitivity DNA assay
470 respectively). Libraries were then sequenced using Illumina HiSeq 2500 v4 technology, at 150
471 base-pairs (bp) paired-end reads.

472

473 **De novo transcriptome assembly, annotation, and differential expression analysis**

474 The quality of the raw reads generated from each of the six libraries was evaluated using FastQC
475 v0.11.5 [88] and then, adapter sequences and low-quality bases (phred score < 30) were removed
476 using Trimmomatic [89]. Processed reads of all replicates were then pooled together and
477 assembled *de novo* using the software Trinity v2.4.0 ([90, 91] to generate a reference
478 transcriptome. Functional annotation of the reference transcriptomes was performed first by
479 blasting the transcripts with DIAMOND [92] against the UniprotKB database [93, 94] with a cut-
480 off E value of 1e-5. Transcripts with blast hits were further annotated using Gene Ontology (GO)
481 terms with BLAST2GO PRO [95].

482 To identify the genes significantly upregulated in the elytra and the posterior end of the
483 body, a differential gene expression (DGE) analysis was carried out with the Trinity module [90],
484 which incorporates Bowtie and RSEM [96], to map and estimate transcript abundance, and edgeR
485 [97] to identify differentially expressed transcripts [91] in both species. We later selected only
486 those genes showing a minimum of 2-fold change, and *p*-value cut off for FDR of 0.01. A GO

487 enrichment analysis was performed using the GO annotations of the significant differentially
488 expressed transcripts using gProfiler [98] with a FDR cutoff 0.05.

489

490 **Comparative transcriptomics of bioluminescent and non-bioluminescent polynoid species.**

491 To assess whether bioluminescent polynoids possess a unique set of proteins shared across the
492 family, we compared the transcriptomes of bioluminescent (including our reference
493 transcriptomes) and non-bioluminescent species using OrthoVenn2 [99]. For this, we assembled
494 the reference transcriptomes of the non-bioluminescent polynoids *Branchipolynoe pettibone*
495 (Miura & Hashimoto, 1991) [100] and *Lepidonotopodium* sp. [100, 101] and the bioluminescent
496 *Harmothoe extenuata* (Grube, 1840) [102], using the available reads in SRA (Table 2) and
497 applying the same bioinformatic pipeline as for our reference transcriptomes. Predicted protein
498 datasets from the transcriptomes were obtained with TransDecoder [90] and implemented in
499 OrthoVenn2 where we set a cut-off E value of 1e-5 for the comparisons. This served us to select
500 only those orthologous gene clusters shared by the luminous polynoid species (“bioluminescent
501 clusters”) for downstream analyses. Functional annotation of the shared bioluminescents clusters
502 was automatically performed in OrthoVenn2 against UniProtKB [93, 94] as well an additional
503 GO enrichment analysis.

504

505 **Histology analysis of elytra**

506 To identify the presence of photocytes, histological analyses were performed on the preserved
507 elytra of *H. imbricata*. Briefly, samples were rinsed in distilled water, dehydrated through an
508 ascending series of ethanol, bathed in xylene and embedded in paraffin overnight at 60°C. Serial
509 7µm-thick sections were then made with a Leitz 1512 microtome and stained with a standard
510 haematoxylin - eosin protocol and mounted with DPX. The resulting preparations were scanned
511 using an Olympus© BX51-P microscope and photographed with an Olympus DP-23 camera.

512 Confocal and light microscopy of elytra preserved in formalin was also performed taking
513 advantage of the autofluorescence of the photosomes present in the photocytes of the elytra of
514 both species. The excitation light used was of 405 nm in order to observe the disposition and

515 arrangement of the photocytes within the elytra. The confocal microscope used was an LSCM
516 Olympus FV1200 at Universidad Complutense de Madrid and the light microscope was a
517 Olympus BX53 at Museo Nacional de Ciencias Naturales (MNCN).

518

519 **Acknowledgements**

520 We acknowledge Ana Sánchez and Eduardo Roldán for help with fluorescent microscopy and the
521 staff of the imaging center of Universidad Complutense de Madrid for help with the confocal
522 microscope. We are also thankful to Belén Arias for help amplifying the cytochrome c oxidase I
523 of *Harmothoe* from Galicia to confirm IDs. This work was funded by the Camille and Henry
524 Dreyfus Teacher-Scholar Award and National Institutes of Health (NIH-NIMHD grant 8-G-12-
525 MD007599) to MH, and the European Union's Horizon 2020 Research and Innovation program
526 through a Marie Skłodowska-Curie Individual Fellowship (grant agreement 841576) and the
527 Spanish Ministry of Science MCIN/AEI/ 10.13039/501100011033, European Union
528 NextGenerationEU/PRTR (grant IJC2020-045256-I) to AV.

529

530 **References**

531

- 532 1. Oba Y, Stevani C V., Oliveira AG, Tsarkova AS, Chepurnykh T V., Yampolsky I V. Selected
533 Least Studied but not Forgotten Bioluminescent Systems. *Photochemistry and Photobiology*.
534 2017;93:405–15.
- 535 2. Fajardo C, De Donato M, Rodulfo H, Martinez-Rodriguez G, Costas B, Mancera JM, et al.
536 New perspectives related to the bioluminescent system in dinoflagellates: *Pyrocystis lunula*, a
537 case study. *International Journal of Molecular Sciences*. 2020;21.
- 538 3. Widder EA. Bioluminescence in the ocean: Origins of biological, chemical, and ecological
539 diversity. *Science*. 2010;328:704–8.
- 540 4. Oba Y, Schultz DT. Eco-evo bioluminescence on land and in the sea. *Adv Biochem Eng*
541 *Biotechnol*. 2014;144:3–36.

- 542 5. Haddock SHD, Moline MA, Case JF. Bioluminescence in the sea. *Ann Rev Mar Sci.*
543 2010;2:443–93.
- 544 6. Kubodera T, Koyama Y, Mori K. Observations of wild hunting behaviour and bioluminescence
545 of a large deep-sea, eight-armed squid, *Taningia danae*. *Proceedings of the Royal Society B:*
546 *Biological Sciences.* 2007;274:1029–34.
- 547 7. Tsuji FI, Lynch R V, Haneda Y. Studies on the Bioluminescence of the Marine Ostracod
548 Crustacean *Cypridina serrata*. *Biological Bulletin.* 1970;139:386–401.
- 549 8. Shimomura O, Masugi T, Johnson FH, Hanedal Y. Properties and Reaction Mechanism of the
550 Bioluminescence System of the Deep-sea Shrimp *Oplophorus gracilorostris*P. *Biochemistry.*
551 1978;17:994–8.
- 552 9. Young RE, Roper CFE, Mangold K, Leisman G, Hochberg FG. Luminescence from Non-
553 Bioluminescent Tissues in Oceanic Cephalopods. *Mar Biol.* 1979;53:69–77.
- 554 10. Taboada S, Silva AS, Díez-Vives C, Neal L, Cristobo J, Ríos P, et al. Sleeping with the enemy:
555 unravelling the symbiotic relationships between the scale worm *Neopolynoe chondrocladiae*
556 (Annelida: Polynoidae) and its carnivorous sponge hosts. *Zool J Linn Soc.* 2020;:1–24.
- 557 11. Verdes A, Gruber DF. Glowing Worms: Biological, Chemical, and Functional Diversity of
558 Bioluminescent Annelids. *Integrative and Comparative Biology.* 2017;57:18–32.
- 559 12. Viviani VR, Silva JR, Amaral DT, Bevilaqua VR, Abdalla FC, Branchini BR, et al. A new
560 brilliantly blue-emitting luciferin-luciferase system from *Orfelia fultoni* and *Keroplantinae*
561 (Diptera). *Sci Rep.* 2020;10.
- 562 13. Tsarkova AS, Kaskova ZM, Yampolsky I V. A Tale of Two Luciferins: Fungal and
563 Earthworm New Bioluminescent Systems. *Acc Chem Res.* 2016;49:2372–80.
- 564 14. Lau ES, Oakley TH. Multi-level convergence of complex traits and the evolution of
565 bioluminescence. *Biological Reviews.* 2021;96:673–91.
- 566 15. Wampler JE, Jamieson BGM. Earthworm bioluminescence: comparative physiology and
567 biochemistry. *Biochem Physiol.* 1980;66:43–50.
- 568 16. Bassot J-M, Nicolas M-T. Bioluminescence in scale-worm photosomes: the photoprotein
569 polynoidin is specific for the detection of superoxide radicals. 1995.

- 570 17. Francis WR, Powers ML, Haddock SHD. Characterization of an anthraquinone fluor from the
571 bioluminescent, pelagic polychaete *Tomopteris*. *Luminescence*. 2014;29:1135–40.
- 572 18. Rawat R, Deheyn DD. Evidence that ferritin is associated with light production in the mucus
573 of the marine worm *Chaetopterus*. *Sci Rep*. 2016;6.
- 574 19. Nicol JAC. Luminescence in polynoid worms. *Journal of the Marine Biological Association*
575 of the United Kingdom. 1953;32:65–83.
- 576 20. Nicol JAC. Luminescence in polynoids II. Different modes of response in the elytra. 1957.
- 577 21. Moraes G V., Hannon MC, Soares DMM, Stevani C V., Schulze A, Oliveira AG.
578 Bioluminescence in Polynoid Scale Worms (Annelida: Polynoidea). *Frontiers in Marine Science*.
579 2021;8.
- 580 22. Gonzalez BC, Martínez A, Borda E, Iliffe TM, Eibye-Jacobsen D, Worsaae K. Phylogeny and
581 systematics of Aphroditiformia. *Cladistics*. 2018;34:225–59.
- 582 23. Norlinder E, Nygren A, Wiklund H, Pleijel F. Phylogeny of scale-worms (Aphroditiformia,
583 Annelida), assessed from 18SrRNA, 28SrRNA, 16SrRNA, mitochondrial cytochrome c oxidase
584 subunit I (COI), and morphology. *Mol Phylogenet Evol*. 2012;65:490–500.
- 585 24. Livermore J, Perreault T, Rivers T. Luminescent defensive behaviors of polynoid polychaete
586 worms to natural predators. *Mar Biol*. 2018;165.
- 587 25. Plyuscheva M, Martin D. Zoosymposia 2: On the morphology of elytra as luminescent organs
588 in scale-worms (Polychaeta... *Zoosymposia*. 2009;2:379–89.
- 589 26. Shimomura O. *Bioluminescence: Chemical principles and methods*. World Scientific
590 Publishing Company; 2012.
- 591 27. Osborn KJ, Haddock SHD, Pleijel F, Madin LP, Rouse GW. Deep-Sea, Swimming Worms
592 with Luminescent “Bombs.” *Science* (1979). 2009;325:964.
- 593 28. Herring PJ. *Bioluminescence in Action*. London: Academic Press; 1978.
- 594 29. Bassot JM. Une forme microtubulaire et paracrystalline de reticulum endoplasmique dans les
595 photocytes des annelides Polynoiaae. *Cell Biology*. 1966;31:135–58.
- 596 30. Bilbaut A. Cell junctions in the excitable epithelium of bioluminescent scales on a polynoid
597 worm: a freeze-fracture and electrophysiological study. *J Cell Sci*. 1980;41:341–68.

- 598 31. Ouldali M, Maury V, Nicolas G, Lepault J. Photosome membranes merge and organize
599 tending towards rhombohedral symmetry when light is emitted. *J Struct Biol.* 2018;202:35–41.
- 600 32. Nicolas M-T, Bassot J-M, Shlmomura O. Polynoidin: a membrane photoprotein isolated from
601 the bioluminescent system of scale-worms. *Photochem Photobiol.* 1982;35:201–7.
- 602 33. Bassot J-M, Nicolas G. An Optional Dyadic Junctional Complex Revealed by Fast-Freeze
603 Fixation in the Bioluminescent System of the Scale Worm. *J Cell Biol.* 1987;105:2245–56.
- 604 34. Miron MJ, LaRivière L, Bassot JM, Anctil M. Immunohistochemical and radioautographic
605 evidence of monoamine-containing cells in bioluminescent elytra of the scale-worm *Harmothoe*
606 *imbricata* (Polychaeta). *Cell Tissue Res.* 1987;249:547–56.
- 607 35. Aneli NB, Shunkina K V., Vays VB, Plyuscheva M V. Ultrastructure and morphology of the
608 elytrum of scaleworm *Lepidonotus squamatus* Linnaeus, 1767 (Polychaeta, Polynoidae).
609 *Invertebrate Zoology.* 2017;14:99–107.
- 610 36. Linnaeus C. *Harmothoe imbricata* (Linnaeus, 1767). In: *Systema Naturae*. 12th edition. 1767.
611 p. 533–1327.
- 612 37. Grube AE. *Beschreibung neuer oder wenig bekannter Anneliden*. 5th edition. Berlin: Archivio
613 de *Historia Natural*; 1860.
- 614 38. Ribeiro RP, Ponz-Segrelles G, Bleidorn C, Aguado MT. Comparative transcriptomics in
615 Syllidae (Annelida) indicates that posterior regeneration and regular growth are comparable,
616 while anterior regeneration is a distinct process. *BMC Genomics.* 2019;20.
- 617 39. Hood FE, Williams SJ, Burgess SG, Richards MW, Roth D, Straube A, et al. Coordination of
618 adjacent domains mediates TACC3-ch-TOG-clathrin assembly and mitotic spindle binding.
619 *Journal of Cell Biology.* 2013;202:463–78.
- 620 40. Vadlamudi RK, Li F, Adam L, Nguyen D, Ohta Y, Stossel TP, et al. Filamin is essential in
621 actin cytoskeletal assembly mediated by p21-activated kinase 1. *Nat Cell Biol.* 2002;4:681–90.
- 622 41. Drake PF, Lasek2 RJ. Regional differences in the neuronal cytoskeleton. *The Journal of*
623 *Neuroscience.* 1984;4:1173–86.
- 624 42. Schwelberger HG. Structural organization of mammalian copper-containing amine oxidase
625 genes. *Inflammation Research.* 2010;59 SUPPL. 2.

- 626 43. Krasney PA, Carr C, Cavener DR. Evolution of the Glucose Dehydrogenase Gene in
627 *Drosophila*. *Mol Biol Evol.* 1990;7:155–77.
- 628 44. Kovtuna K, Strazdina I, Bikerniece M, Galinina N, Rutkis R, Martynova J, et al. Improved
629 Hydrogen Peroxide Stress Resistance of *Zymomonas mobilis* NADH Dehydrogenase (*ndh*) and
630 Alcohol Dehydrogenase (*adhB*) Mutants. *Fermentation.* 2022;8.
- 631 45. Konstandi OA, Papassideri IS, Stravopodis DJ, Kenoutis CA, Hasan Z, Katsorchis T, et al.
632 The enzymatic component of *Drosophila melanogaster* chorion is the Pxd peroxidase. *Insect*
633 *Biochem Mol Biol.* 2005;35:1043–57.
- 634 46. Kumaran S, Ngo ACR, Schultes FPJ, Saravanan VS, Tischler D. In vitro and in silico analysis
635 of Brilliant Black degradation by Actinobacteria and a Paraburkholderia sp. *Genomics.* 2022;114.
- 636 47. Uemura M, Tamura K, Chung S, Honma S, Okuyama A, Nakamura Y, et al. Novel 5 α -steroid
637 reductase (*SRD5A3*, type-3) is overexpressed in hormone-refractory prostate cancer. *Cancer Sci.*
638 2008;99:81–6.
- 639 48. Schauder S, Schneider K-H, Giffhorn F. Polyol metabolism of *Rhodobacter sphaeroides* :
640 biochemical characterization of a short-chain sorbitol dehydrogenase. *Microbiology (N Y).*
641 1995;141:1857–63.
- 642 49. Gopalakrishna R, Jaken S. Protein kinase C signaling and oxidative stress. *Free Radic Biol*
643 *Med.* 2000;28:1349–61.
- 644 50. Avilés EC, Goodrich L v. Configuring a robust nervous system with Fat cadherins. *Seminars*
645 *in Cell and Developmental Biology.* 2017;69:91–101.
- 646 51. Sayavedra-Soto LA, Hommes NG, Arp DJ. Characterization of the Gene Encoding
647 Hydroxylamine Oxidoreductase in *Nitrosomonas europaea*. *Journal of Bacteriology .* 1994;:504–
648 10.
- 649 52. Sorci L, Kurnasov O, Rodionov D, Osterman A. Genomics and Enzymology of NAD
650 Biosynthesis. In: Hung-Wen L, Mander L, editors. *Comprehensive Natural Products II.* Elsevier;
651 2010. p. 213–57.

- 652 53. Qian J, Chen F, Kovalenkov Y, Pandey D, Moseley MA, Foster MW, et al. Nitric oxide
653 reduces NADPH oxidase 5 (Nox5) activity by reversible S-nitrosylation. *Free Radic Biol Med.*
654 2012;52:1806–19.
- 655 54. Guengerich FP, Waterman MR, Egli M. Recent Structural Insights into Cytochrome P450
656 Function. *Trends in Pharmacological Sciences.* 2016;37:625–40.
- 657 55. Miron MJ, LaRivière L, Bassot JM, Anctil M. Immunohistochemical and radioautographic
658 evidence of monoamine-containing cells in bioluminescent elytra of the scale-worm *Harmothoe*
659 *imbricata* (Polychaeta). *Cell Tissue Res.* 1987;249:547–56.
- 660 56. Rees J-F, de Wergifosse B, Noiset O, Dubuisson M, Janssens B, Thompson EM. The origins
661 of marine bioluminescence: turning oxygen defence mechanisms into deep-sea communication
662 tools. *J Exp Biol.* 1998;201:1211–21.
- 663 57. Kareem H. Oxidoreductases: Significance for Humans and Microorganism. In: Ahmed M,
664 editor. *Oxidoreductase.* London: IntechOpen; 2020. p. 115–6.
- 665 58. Oba Y, Ojika M, Inouye S. Firefly luciferase is a bifunctional enzyme: ATP-dependent
666 monooxygenase and a long chain fatty acyl-CoA synthetase. *FEBS Lett.* 2003;540:251–4.
- 667 59. Baldwin TO, Christopher JA, Raushel FM, Sinclair JF, Ziegler MM, Fisher AJ, et al. Structure
668 of bacterial luciferase. *Curr Opin Struct Biol.* 1995;5:798–809.
- 669 60. Kotlobay AA, Sarkisyan KS, Mokrushina YA, Marcet-Houben M, Serebrovskaya EO,
670 Markina NM, et al. Genetically encodable bioluminescent system from fungi. *Proc Natl Acad Sci*
671 *U S A.* 2018;115:12728–32.
- 672 61. Rodionova NS, Rota E, Tsarkova AS, Petushkov VN. Progress in the Study of Bioluminescent
673 Earthworms. *Photochemistry and Photobiology.* 2017;93:416–28.
- 674 62. Lushchak VI. Free radicals, reactive oxygen species, oxidative stress and its classification.
675 *Chemico-Biological Interactions.* 2014;224:164–75.
- 676 63. Yasutake S, Okita S, Takeshige Y. The Fine Structure of the Light Organ of Firefly (*Luciola*
677 *Crudata*), Especially the Photocyte Granules. *Journal of Electronmicroscopy.* 1963;12:240–53.
- 678 64. Hanna CH, Hopkins TA, Buck J. Peroxisomes of the Firefly Lantern. 1976.

- 679 65. Booth DG, Hood FE, Prior IA, Royle SJ. A TACC3/ch-TOG/clathrin complex stabilises
680 kinetochore fibres by inter-microtubule bridging. *EMBO Journal*. 2011;30:906–19.
- 681 66. McDougall CA. Bioluminescence and the actin cytoskeleton in the dinoflagellate *Pyrocystis*
682 *fusiformis*: an examination of organelle transport and mechanotransduction. University of
683 California; 2002.
- 684 67. Livermore J, Perreault T, Rivers T. Luminescent defensive behaviors of polynoid polychaete
685 worms to natural predators. *Mar Biol*. 2018;165.
- 686 68. Kanie S, Miura D, Jimi N, Hayashi T, Nakamura K, Sakata M, et al. Violet bioluminescent
687 *Polycirrus* sp. (Annelida: Terebelliformia) discovered in the shallow coastal waters of the Noto
688 Peninsula in Japan. *Sci Rep*. 2021;11.
- 689 69. Nicol JAC. Luminescence in polynoid worms. *Journal of the Marine Biological Association*
690 *of the United Kingdom*. 1953;32:65–84.
- 691 70. Herrera AA. A Electrophysiology of Bioluminescent Excitable Epithelial Cells in a Polynoid
692 Polychaete Worm. *J Comp Physiol*. 1979;129:67–78.
- 693 71. Bomont P. The dazzling rise of neurofilaments: Physiological functions and roles as
694 biomarkers. *Current Opinion in Cell Biology*. 2021;68:181–91.
- 695 72. Pavans de Ceccaty M, Bassot J, Bilbaut A, Nicolas M. Bioluminescence des élytres
696 d’*Acholoe*. *Biol Cell*. 1977;64:28–57.
- 697 73. Nicolas M. Bioluminescence des élytres d’*Acholoe*. V Les principales étapes de la
698 régénération. *Archives de zoologie expérimentale et générale*. 1977;118:103–20.
- 699 74. Mallefet J, Baguet F. Metabolic control of luminescence in isolated photophores of
700 porichthys: effects of glucose on oxygen consumption and luminescence. *Journal of Experimental*
701 *Biology*. 1993;181:279–93.
- 702 75. Kingberg J. *Annulata Nova* [Continuatio. Fam. Lumbricina (Sav.). Öfversigt af Kongl
703 Vetenskaps- Akadamiens Förhandlingar. 1866;23:97–103.
- 704 76. Haram NS, Rao L, Santhanam KS v. Glucose activated bioluminescence of coelomic cells of
705 *Lampito mauritii*: a novel assay of glucose. 1996.

- 706 77. Oba Y, Kato S-I, Ojika M, Inouye S. Biosynthesis of luciferin in the sea firefly, *Cypridina*
707 *hilgendorffii*: L-tryptophan is a component in *Cypridina* luciferin. *Tetrahedron Lett.*
708 2002;43:2389–92.
- 709 78. Kato SI, Oba Y, Ojika M, Inouye S. Identification of the biosynthetic units of *Cypridina*
710 luciferin in *Cypridina* (*Vargula*) *hilgendorffii* by LC/ESI-TOF-MS. *Tetrahedron.* 2004;60:11427–
711 34.
- 712 79. Stincone A, Prigione A, Cramer T, Wamelink MMC, Campbell K, Cheung E, et al. The return
713 of metabolism: Biochemistry and physiology of the pentose phosphate pathway. *Biological*
714 *Reviews.* 2015;90:927–63.
- 715 80. Wanders RJA, Duran M, Loupatty FJ. Enzymology of the branched-chain amino acid
716 oxidation disorders: The valine pathway. *J Inherit Metab Dis.* 2012;35:5–12.
- 717 81. Bassot J-M. A Transient Intracellular Coupling Explains the Facilitation of Responses in the
718 Bioluminescent System of Scale Worms. *J Cell Biol.* 1987;105:2235–43.
- 719 82. Goñi A, Plyuscheva M. Fluorescence of scale-worms (*Polychaeta*, *Polynoidae*) under the
720 different photonic conditions Bioluminescence and fluorescence in scale-worms View project
721 Mediterranean Sea View project Masha Plyuscheva Centre for Genomic Regulation. *Interface*
722 *Focus* (Under review). 2020.
- 723 83. Nicol JAC. Luminescence in Polynoids IV. Measurements of light intensity. *J mar biol Ass*
724 *UK.* 1958;37:33–41.
- 725 84. Aneli NB, Shunkina K v., Vays VB, Plyuscheva M v. Ultrastructure and morphology of the
726 elytrum of scaleworm *Lepidonotus squamatus* Linnaeus, 1767 (*Polychaeta*, *Polynoidae*).
727 *Invertebrate Zoology.* 2017;14:99–107.
- 728 85. Plyushcheva M V, Pereira Da Silva K, Aneli NB, Vays VB, Kondrashov FA, Goñi AR. Two-
729 color fluorescence in elytra of the scale-worm *Lepidonotus squamatus* (*Polychaeta*, *Polynoidae*):
730 in vivo spectral characteristic. *†. Mater Today Proc.* 2017;4:4998–5005.
- 731 86. Barnich Ruth, Fiege Dieter. Revision of the genus *Harmothoe* Kinberg, 1856
732 (*Polychaeta:Polynoidae*) in the Northeast Atlantic. *Magnolia Press*; 2009.

- 733 87. Groelz D, Sobin L, Branton P, Compton C, Wyrich R, Rainen L. Non-formalin fixative versus
734 formalin-fixed tissue: A comparison of histology and RNA quality. *Exp Mol Pathol.*
735 2013;94:188–94.
- 736 88. www.bioinformatics.babraham.ac.uk.
- 737 89. Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for Illumina sequence
738 data. *Bioinformatics.* 2014;30:2114–20.
- 739 90. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-length
740 transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol.*
741 2011;29:644–52.
- 742 91. Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, et al. De novo
743 transcript sequence reconstruction from RNA-seq using the Trinity platform for reference
744 generation and analysis. *Nat Protoc.* 2013;8:1494–512.
- 745 92. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. *Nature*
746 *Methods.* 2014;12:59–60.
- 747 93. Bairoch A. Serendipity in bioinformatics, the tribulations of a Swiss bioinformatician through
748 exciting times! 2000.
- 749 94. Bateman A, Martin MJ, Orchard S, Magrane M, Agivetova R, Ahmad S, et al. UniProt: the
750 universal protein knowledgebase in 2021. *Nucleic Acids Res.* 2021;49:D480–9.
- 751 95. Conesa A, Madrigal P, Tarazona S, Gomez-Cabrero D, Cervera A, McPherson A, et al. A
752 survey of best practices for RNA-seq data analysis. *Genome Biology.* 2016;17.
- 753 96. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or
754 without a reference genome. *BMC Bioinformatics.* 2011;12:323.
- 755 97. Robinson MD, McCarthy DJ, Smyth GK. edgeR: A Bioconductor package for differential
756 expression analysis of digital gene expression data. *Bioinformatics.* 2009;26:139–40.
- 757 98. Raudvere U, Kolberg L, Kuzmin I, Arak T, Adler P, Peterson H, et al. G:Profiler: A web
758 server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic*
759 *Acids Res.* 2019;47:W191–8.

- 760 99. Xu L, Dong Z, Fang L, Luo Y, Wei Z, Guo H, et al. OrthoVenn2: A web server for whole-
 761 genome comparison and annotation of orthologous clusters across multiple species. *Nucleic Acids*
 762 *Res.* 2019;47:W52–8.
- 763 100. Miura T, Hashimoto J. Two new branchiate scale-worms (Polynoidae: Polychaeta) from the
 764 hydrothermal vent of the Okinawa Trough and the volcanic seamount off Chichijima Island.
 765 *Proceedings of the Biological Society of Washington.* 1991;104:166–74.
- 766 101. Pettibone MH. A new scale worm (Polychaeta: Polynoidae) from the hydrothermal rift-area
 767 off western Mexico at 21N. *Proceedings of the Biological Society of Washington.* 1983;96:392–
 768 9.
- 769 102. Grube AE. Actinien, Echinodermen und Würmer des Adriatischen- und Mittelmeers nach
 770 eigenen Sammlungen beschrieben. Königsberg: JH Bon. 1840;:92.

771

772 **Tables**

773 **Table 1: List of clusters shared between the bioluminescent species**

774

Cluster	Protein ID	Go Annotation	Relation with bioluminescence
cluster13850	Q9C0H6	Actin binding	Microtubule organization
cluster10051	P17644	Synaptic transmission, cholinergic	Nervous system
cluster11826	Q5IS76	Synaptic transmission, cholinergic	
cluster3310	H2A0M3	Amine metabolic process	Oxidoreductase
cluster5878	J9VQZ4	Hydroquinone:oxygen oxidoreductase activity	
cluster6532	O86034	3-hydroxybutyrate dehydrogenase activity	
cluster8278	Q7Z5P4	Positive regulation of lipid biosynthetic process	
cluster16223	Q8HXW0	L-ascorbic acid biosynthetic process	

cluster5790	Q6PHY8	Thioredoxin-disulfide reductase activity	
cluster9970	Q6UVY6	Octopamine biosynthetic process	
cluster5785	P08761	Sulfur amino acid metabolic process	Response to oxidative stress
cluster4890	Q9CQN6	Regulation of heme biosynthetic process	
cluster2261	Q9EPF3	Estructural constituent of eye lens	Visual sense
cluster10748	Q9QYF1	Retinol metabolic process	
cluster11855	Q9GLM3	Visual perception	

775

776 The table shows the identification protein label (Protein ID), the annotation made by Orthovenn
 777 (Go Annotation), and the corresponding function related with bioluminescence
 778 (Bioluminescence).

779

780 **Table 2 Accession numbers.**

Species	Accession number NCBI
Harmothoe areolata	SAMN38153224
Harmothoe imbricata	SAMN38153218
Harmothoe imbricata elytra	SRR4841788
Harmothoe extenuata	SRR1237766
Branchipolynoe pettiboneae	SRR4419842
Lepidonotopodium sp.	SRR4419843

781

782 NCBI Reference numbers of the transcriptomes of the species used in the Orthovenn analyses.

783

784 **Ethics declarations**

785

786 - Ethics approval and consent to participate: Not applicable.

787 - Consent for publication: Not applicable.

788 - Competing interests: The authors declare no conflict of interests.

789

790 **Availability of data and materials**

791 The datasets generated and/or analyzed during the current study are available in the [Additional
792 files] repository,

793 [https://www.dropbox.com/sh/rqx5l06cglk7ojp/AAAdKFH_i9xSo_2_yeYJNpFQa?dl=0]

794

795 **Additional files:**

796 **Additional file 1 Sequencing information.** *De novo* transcriptome assembly metrics and
797 transcriptome completeness statistics based on BUSCO analysis with the Metazoa gene set.

798

799 **Additional file 2 Scale up subset.** Table showing the results of the Differential Expression
800 Analyses performed for both species (one species per sheet).

801

802 **Additional file 3 GO terms.** Results of the Gen Ontology analyses performed with the whole
803 transcriptome of the selected species as background. The results for each species are found on
804 separate sheets.

805

806 **Additional file 4 Orthovenn results.** Orthologous genes shared by the selected bioluminescent
807 species.