TAEL 2.0: An Improved Optogenetic Expression System for Zebrafish

Jesselynn LaBelle[1], Adela Ramos-Martinez[1], Kyle Shen[1], Laura B. Motta-Mena[3], Kevin H. Gardner[2], Stefan C. Materna[1], and Stephanie Woo[1,4]

[1]Department of Molecular Cell Biology, University of California, Merced, CA 95343 USA. [2]Structural Biology Initiative, Advanced Science Research Center, CUNY; New York, NY, 10031, USA. [3]Optologix, Inc., Dallas, TX 75201 USA. [4]Author for correspondence: swoo6@ucmerced.edu

Key Words: Optogenetics, Zebrafish, Gene expression

Abstract

Inducible gene expression systems are valuable tools for studying biological processes. We previously developed an optogenetic gene expression system called TAEL that is optimized for use in zebrafish. When illuminated with blue light, TAEL transcription factors dimerize and activate gene expression downstream of the TAEL-responsive C120 promoter. By using light as the inducing agent, the TAEL/C120 system overcomes limitations of traditional inducible expression systems by enabling fine spatial and temporal regulation of gene expression. Here, we describe ongoing efforts to improve the TAEL/C120 system. We made modifications to both the TAEL transcriptional activator and the C120 regulatory element, collectively referred to as "TAEL 2.0." We demonstrate that TAEL 2.0 consistently induces higher levels of reporter gene expression and at a faster rate, but with comparable background and toxicity as the original TAEL system. With these improvements, we were able to create functional stable transgenic lines to express the TAEL 2.0 transcription factor either ubiquitously or with a tissue-specific promoter. We demonstrate that the ubiquitous line in particular can be used to induce expression at late embryonic and larval stages, addressing a major deficiency of the original TAEL system. We believe this improved optogenetic expression system will be a useful resource for the zebrafish community.

Introduction

Inducible gene expression systems are valuable tools for studying biological processes as they enable user-defined control over the timing, location, and level of expression. In zebrafish and other model organisms, the most widely used inducible expression systems fall into two broad categories – those that rely on the heat shock response[1] and those using small molecule inducing agents[2]. More recently, optogenetic approaches have been developed based on light-sensitive transcription factors[3-6]. One such system is based on EL222, a naturally occurring blue light-activated transcription factor found in the bacterium *Erythrobacter litoralis HTCC2594*. The endogenous transcription factor contains a light-oxygen-voltage-sensing (LOV) domain that in response to blue light (450 nm) undergoes a conformational change and dimerizes, allowing it to bind and initiate transcription from a regulatory element termed C120[7]. EL222 was the basis for an inducible expression system designed for mammalian cell culture[8]. Our group previously designed EL222 for use in zebrafish by fusing it to a KalTA4 transcriptional activation domain, which minimized toxicity in zebrafish embryos while still maintaining functionality[6]. We demonstrated that this KalTA4-EL222 fusion protein, which we termed TAEL, could be combined with C120-containing transgenes to achieve light-inducible expression of multiple genes of interest. We also validated multiple approaches for delivering patterned blue light illumination to spatially and temporally control induction in zebrafish embryos. However, we were unable to establish stable transgenic lines for TAEL expression that could induce expression from our C120 reporter lines, suggesting that TAEL and/or the C120 promoter could be further optimized.

In this study, we present ongoing efforts to improve the function of the TAEL/C120 system. We made changes to both the TAEL transcriptional activator and the C120 promoter, collectively termed TAEL 2.0, that produce significantly higher levels of light-induced expression at a faster rate. Importantly, these improvements allowed us to address a major deficiency of our

53 previously published system (referred to here as TAEL 1.0), namely the lack of functional, stable
54 transgenic lines for both TAEL and C120 components. Here, we describe the generation of
55 transgenic lines that express functional TAEL 2.0 components either ubiquitously or in the
56 developing endoderm. We demonstrate that the ubiquitous line in particular can be used to
57 induce expression at late embryonic and larval stages, extending the use of this system beyond
58 early embryo stages.
59
60 **Materials and Methods**
61
62 **Vector construction and mRNA synthesis**
63 *pµTol2 backbone.* For expression plasmids and transgenes created for this study, we generated
64 a minimal plasmid backbone called pµTol2, which can be used for both Tol2-based
65 transgenesis and in vitro mRNA synthesis. Its short length of 2520 base pairs enables
66 modification of inserts by PCR through the backbone, thus eliminating the need to subclone. In
67 brief, pµTol2 was constructed by Gibson assembly, fusing the Tol2 sites for genomic
68 integration[9] with the commonly used expression cassette of pCS2 including polylinkers and
69 SV40 polyadenylation site[10,11] and a plasmid backbone derived from pUC19[12]. To ensure
70 efficient protein synthesis, all plasmids newly constructed for this study contain the zebrafish-
71 optimized Kozak sequence 5'-GCAAACatgG-3', where the lower case "atg" denotes the start
72 codon[13].
73
74 *Expression plasmids.* pCS2-TAEL has been described previously[6]. To construct expression
75 plasmids pµTol2-N-TAEL, Optologix, Inc. (Dallas, TX) provided synthesized oligomers
76 containing the SV40 large T-antigen nuclear localization signal. We fused these to the 5′ end of
77 the TAEL ORF and to the pµTol2 backbone by Gibson assembly[14]. Similarly, pµTol2-TAEL-N
78 was constructed by fusing synthesized oligomers containing the nucleoplasmin nuclear

79 localization signal (also provided by Optologix, Inc.) to the 3′ end of the TAEL ORF by Gibson
80 assembly[14]. Capped messenger RNA was synthesized using mMESSAGE mMACHINE SP6 kit
81 (Ambion) with plasmids cut with NotI as linear template. For experiments in Fig. 1–3,
82 Tg(C120:mCherry;cryaa:Venus) or Tg(C120F:mCherry) males were crossed to wild-type
83 females and resulting embryos were each injected with ~50 pg of TAEL, N-TAEL, or TAEL-N
84 mRNA at the 1-cell stage.
85
86 *Transgene plasmids.* To construct pμTol2-C120F:mCherry, the mouse *Fos* basal promoter
87 sequence: 5'-CCAGTGACGTAGGAAGTCCATCCATTCACAGCGCTTC-
88 TATAAAGGCGCCAGCTGAGGCGCCTACTACTCCAACCGCGACTGCAGCGAGCAACT -3'[15]
89 was synthesized by Integrated DNA Technologies and the C120 sequence[6] was amplified by
90 PCR. These sequences were fused together and inserted into pμTol2 by Gibson assembly. The
91 transgene plasmid pμTol2-C120F:GFP was constructed by separate PCR amplification of the
92 C120F promoter and GFP ORF which were then cloned into pμTol2 by Gibson assembly.
93 pμTol2-sox17:TAEL-N was constructed by separate PCR amplification of the *sox17* promoter[16]
94 and TAEL-N ORF which were then cloned into pμTol2 by Gibson assembly. pμTol2-ubb:TAEL-
95 N was constructed by separate PCR amplification of the *ubb* promoter[17] and TAEL-N ORF,
96 which were then cloned into pμTol2 by Gibson assembly.
97
98 **Zebrafish Strains**
99 Adult Danio rerio zebrafish were maintained under standard laboratory conditions. Zebrafish in
100 an outbred AB, TL, or EKW background were used as wildtype strains.
101 *Tg(C120:mCherry;cryaa:Venus)^{sfc14}*, referred to here as *Tg(C120T:mCherry)*, has been
102 previously described[6]. *Tg(C120-Mmu.Fos:mCherry)^{ucm104}*, *Tg(C120-Mmu.Fos:GFP)^{ucm107}*,
103 *Tg(ubb:TAEL-N)^{ucm113}*, and *Tg(sox17:TAEL-N)^{ucm114}* were generated using standard

104 transgenesis protocols[9,18]. This study was performed with the approval of the Institutional Animal
105 Care and Use Committee (IACUC) of the University of California Merced.
106
107 **Global light induction**
108 Global light induction was provided by a MARS AQUA-165-55 110W LED aquarium hood.
109 Actual power of light received by embryos (lids of plates removed) was measured as ~1.6 mW/
110 cm^2 at 456 nm. For experiments in Fig. 1–2, 4 hpf (hours post-fertilization) embryos were
111 illuminated with constant blue light for 1–3 hours. For experiments in Fig. 3–6, a timer was used
112 to apply constant or pulsed light (NEARPOW Timer Switch). Dark controls were placed in a light
113 proof box in the same 28.5°C incubator as the light-treated samples.
114
115 **Real-time quantitative PCR**
116 To quantify light-induced expression, total RNA from 30–50 light-treated or dark control embryos
117 was extracted using the illustra™ RNAspin Mini kit (GE Healthcare). 1 µg total RNA was used
118 for reverse transcription with qScript XLT cDNA SuperMix (Quantabio). Each qPCR reaction
119 contained 2X PerfeCTa® SYBR green fast mix (Quantabio), 5-fold diluted cDNA and 325 nM
120 each primer. Reactions were carried out on a QuantStudio3 (Applied Biosystems) real time PCR
121 machine using the following program: initial activation at 95°C for 10 min, followed by 40 cycles
122 of 30 s at 95°C, 30 s at 60°C and 1 min at 72°C. Once the PCR was completed, a melt curve
123 analysis was performed to determine reaction specificity. Data represent averages from 3–5
124 biological replicates, each with three technical replicates. The housekeeping gene *ef1a* was
125 used as a reference. Fold change was calculated using the $2^{(-\Delta\Delta CT)}$ method[19]. Statistical
126 significance was determined using Prism software (GraphPad). qPCR primers used are:
127 mcherry forward: 5'-GACCACCTACAAGGCCAAGA-3'; mcherry reverse: 5'-
128 CTCGTTGTGGGAGGTGATGA-3'; ef1a forward 5'-CACGGTGACAACATGCTGGAG-3'; ef1a
129 reverse: 5'-CAAGAAGAGTAGTACCGCTAGCAT-3'

130

131 **Microscopy and image processing**

132 Fluorescence and brightfield images were acquired on a Leica dissecting stereomicroscope or
133 Olympus dissecting stereomicroscope. Dechorionated embryos or larvae were embedded in
134 1.5% low-melting agarose (ISC BioExpress) containing 0.01% tricaine (Sigmal-Aldrich) within
135 glass-bottom Petri dishes (MatTek Corporation). Standard filter settings were applied and
136 brightfield and fluorescence images were merged after acquisition. Identical exposure settings
137 for fluorescence images were used for all embryos from the same set of experiments. All image
138 processing and analysis was performed using ImageJ software[20]. Illustrations were created with
139 BioRender (https://biorender.com/).

140

141 **Results**

142

143 **TAEL-induced expression is increased by coupling the C120 regulatory element to a *Fos***
144 **basal promoter**

145 In our previously published system, the TAEL-responsive C120 regulatory sequence was
146 coupled to a minimal TATA box[6,8]. Because this minimal TATA box originated from a
147 mammalian expression vector, we reasoned that using a zebrafish-optimized basal promoter
148 instead would improve performance of the TAEL system. The basal promoter from the mouse
149 *Fos* gene was previously shown to function well in zebrafish transgenes, allowing for high
150 expression levels with minimal background[15,21]. Therefore, we constructed a new TAEL-
151 responsive promoter consisting of 5 repeats of the C120 regulatory sequence coupled to the
152 mouse *Fos* basal promoter (*C120-Mmu.Fos*, abbreviated throughout as *C120F*). We then
153 determined whether this new C120 promoter improves light-induced expression compared to
154 the previous TATA box-containing version (Fig. 1). First, we generated a stable transgenic
155 zebrafish line using *C120F* to control expression of an mCherry reporter (*Tg(C120F:mCherry)*)

156 to make direct comparisons to our previously published reporter line[6], referred to here as
157 *Tg(C120T:mCherry)*. We injected both *Tg(C120T:mCherry)* and *Tg(C120F:mCherry)* embryos
158 with ~50 pg TAEL mRNA then globally illuminated them with blue light starting at 3 hpf. qRT-
159 PCR analysis showed that compared to sibling control embryos kept in the dark, mCherry
160 expression was induced $43.5\ ^{+10.6}_{-8.5}$-fold in *Tg(C120F:mcherry)* embryos, which was significantly
161 higher than the $2.9\ ^{+1.2}_{-0.8}$-fold induction in *Tg(C120T:mCherry)* embryos (p=0.0009) (Fig. 1C).
162 Consistent with these results, mCherry fluorescence was qualitatively brighter in
163 *Tg(C120F:mCherry)* embryos compared to *Tg(C120T:mCherry)* embryos (Fig. 1D-E).
164 Importantly, we did not observe mCherry fluorescence in embryos kept in the dark for either
165 genotype (Fig. 1F-G). Together, these results suggest that coupling the C120 regulatory
166 element with a *Fos* basal promoter instead of a minimal TATA box significantly increases TAEL-
167 induced gene expression while maintaining low background expression.
168

169 **TAEL-induced expression is increased by adding a C-terminal nuclear localization signal**
170 **to TAEL**

171 Our original TAEL construct consists of a Kal-TA4 transcription activation domain, the light-
172 sensitive LOV domain, and a DNA-binding domain that recognizes the C120 sequence but does
173 not contain an explicit nuclear localization signal (NLS). Although TAEL can likely enter the
174 nucleus through diffusion, because of its relatively small size of 257 amino acids, we wanted to
175 test whether targeting TAEL specifically to the nucleus by adding an NLS would increase the
176 amplitude of induction and improve light-induced expression (Fig. 2).
177

178 We first generated a construct in which the SV40 large T-antigen NLS was fused to the amino
179 terminus of TAEL (N-TAEL). When delivered by mRNA injection into *Tg(C120F:mCherry)*
180 embryos, we were surprised to find that N-TAEL induced mCherry expression less strongly

181 (47.7 $^{+81.3}_{-30.0}$-fold) than the original TAEL protein (81.7 $^{+69.8}_{-37.7}$-fold; p=0.3398) (Fig. 2C). Consistent
182 with these qPCR results, mCherry fluorescence was more variable and often dimmer in
183 embryos injected with N-TAEL versus TAEL mRNA (Fig. 2D-E). We speculated that fusing the
184 NLS to the N-terminus of TAEL places it directly adjacent to the KalTA4 transcriptional
185 activation domain, which may negatively interfere with transactivation. Therefore, we generated
186 a construct in which the nucleoplasmin NLS was fused to the carboxy terminus of TAEL (TAEL-
187 N). By qPCR analysis, *Tg(C120F:mCherry)* embryos injected with TAEL-N mRNA showed
188 higher levels of mCherry induction (176.5 $^{+87.6}_{-58.5}$-fold) compared to both TAEL (p=0.053) and N-
189 TAEL (p=0.0392) (Fig. 2C). Correspondingly, mCherry fluorescence was brightest in embryos
190 injected with TAEL-N (Fig. 2F). We did not observe mCherry fluorescence in any injected
191 embryos kept in the dark (Fig. 2G-I). Together, these results demonstrate that adding a nuclear
192 localization signal at the C-terminus of TAEL further increases light-induced gene expression
193 with minimal background.
194
195 **TAEL 2.0 induces higher expression levels at a faster rate**
196 We next characterized the effects of combining the modifications we made to the C120
197 promoter and TAEL transcriptional activator. With our previously published TAEL system, we
198 found that peak expression levels were reached by 3 hours post-illumination and could be
199 sustained up to 8 hours when embryos were exposed to blue light pulsed at 1 hour on/off
200 intervals[6]. To determine if TAEL 2.0 improves the kinetics and/or range of light-induced
201 expression, we injected *Tg(C120T:mCherry)* embryos with TAEL mRNA ("TAEL 1.0") or
202 *Tg(C120F:mcherry)* embryos with TAEL-N mRNA ("TAEL 2.0"). Starting at approximately 3 hpf,
203 injected embryos were globally illuminated with pulsed blue light (1 hour on, 1 hour off) and
204 mCherry expression was measured by qRT-PCR at various timepoints up to 9 hours post-
205 illumination. Throughout the time course, we found that TAEL 2.0 induced significantly higher

206 mCherry expression compared to TAEL 1.0 (2-way ANOVA, $p<0.0001$). Induction kinetics also
207 improved. At 1 hour post-illumination, mCherry expression was induced 738.6 $^{+749.2}_{-379.9}$-fold with
208 TAEL 2.0, and this level of expression was maintained up to 9 hours post-illumination. In
209 contrast, with TAEL 1.0, mCherry expression at 1 hour post-illumination was induced 89.1 $^{+54.0}_{-33.6}$-
210 fold; induction peaked by 3 hours post-illumination to 221.6 $^{+222.2}_{-110.0}$-fold and then decreased to
211 56.8 $^{+138.47}_{-16.28}$-fold by 9 hours post-illumination. Together, these results demonstrate that the
212 combined modifications we made to the TAEL system improve both the range and induction
213 kinetics of this light-activated expression system.
214
215 **TAEL 2.0 modifications enable functional stable transgenic lines of TAEL components**
216 One notable deficiency of our previous TAEL system was the lack of functional stable
217 transgenic lines expressing the TAEL transcriptional activator. With its greatly increased
218 amplitude and kinetics of induction, we determined whether TAEL 2.0 could address this
219 previous limitation.
220
221 We generated a stable transgenic line, *Tg(sox17:TAEL-N)*, to express TAEL-N under the *sox17*
222 promoter, which drives expression in the endoderm and dorsal forerunner cells (DFCs)[16]. We
223 crossed this line with a *Tg(C120F:GFP)* reporter line. The resulting double transgenic embryos
224 were globally illuminated with pulsed blue light (1 hour on/off) or kept in the dark from 6–18 hpf
225 (Fig. 4A). We observed GFP fluorescence in derivatives of the endoderm such as the gut tube
226 and the pharyngeal endoderm as well as derivatives of the dorsal forerunner cells (DFCs) within
227 the tail mesoderm in illuminated embryos but not those kept in the dark (Fig. 4B-E). Because
228 activating blue light was applied globally, this result suggested that TAEL-N functions in, and is
229 restricted to, the *sox17* expression domain. Additionally, we observed that the intensity of GFP
230 fluorescence was brightest in the tail (Fig. 4B-C), again consistent with the known *sox17*

231 expression pattern, which is highest in the DFCs. Together, these results demonstrate
232 successful generation of a stable transgenic line for tissue-specific TAEL-N expression, which in
233 turn enables tissue-specific induction of a gene of interest even when activating blue light is
234 applied globally.
235
236 One consequence of the lack of functional stable transgenic lines for TAEL 1.0 is that its use is
237 limited to early embryonic stages. To determine if TAEL 2.0 modifications could expand the
238 range of accessible developmental stages, we generated a stable transgenic line,
239 *Tg(ubb:TAEL-N)*, to express TAEL-N under the *ubb* promoter, which has been shown to drive
240 ubiquitous expression at all developmental stages[17]. We crossed this line to *Tg(C120F:GFP)*
241 then exposed double transgenic embryos to activating blue light at several different time points
242 spanning embryonic to larval stages (Fig. 5A). In all cases, we observed increased GFP
243 fluorescence in illuminated embryos or larva but not in control siblings that had been kept in the
244 dark (Fig. 5B-G). At 4 days post-fertilization (dpf), we observed GFP fluorescence in the livers of
245 both illuminated and control larvae (arrows, Fig. 5D, G), which is likely due to insertional effects
246 of the Tg(C120F:GFP) transgene specific to this line; we did not observe similar liver
247 fluorescence in Tg(C120F:mCherry). Importantly, we could still detect light-dependent GFP
248 induction above this background expression at 4 dpf (Fig. 5D). Taken together, these results
249 demonstrate that TAEL 2.0 can be used to induce expression in a broad range of
250 developmental stages.
251
252 A recent study showed that blue light alone can increase expression of *Fos* and other activity-
253 dependent genes in cultured mouse cortical neurons[22]. Because the *C120F* promoter utilizes
254 the basal promoter from the mouse *Fos* gene, it is possible that there are endogenous factors,
255 especially in neural tissues, that can drive light-responsive expression from the *C120F* promoter
256 independent of TAEL-N and reduce the specificity of the TAEL 2.0 system. To determine

257 whether the *C120F* promoter can function in the absence of TAEL-N, we exposed
258 *Tg(C120F:GFP)* zebrafish to blue light at 2 dpf or 4 dpf; the latter time point was chosen as
259 light-driven neuronal activity likely increases over time. Apart from the ectopic liver expression at
260 4 dpf described above, we did not observe any appreciable GFP fluorescence either in
261 illuminated animals or controls kept in the dark (Fig. 6A-D).
262
263 We quantified GFP expression by qPCR in *Tg(C120F:GFP)* and *Tg(C120F:GFP);Tg(ubb:TAEL-*
264 *N)* double-transgenic animals with and without blue light illumination. At 2 dpf, we detected low
265 levels of GFP expression in both *Tg(C120F:GFP)* and *Tg(C120F:GFP);Tg(ubb:TAEL-N)*
266 embryos kept in the dark, suggesting there is a small amount of basal activity of the *C120F*
267 promoter (Fig. 6E). Upon blue light illumination, we detected strong induction of GFP expression
268 in *Tg(C120F:GFP);Tg(ubb:TAEL-N)* double transgenic embryos (73 $^{+81.2}_{-38.5}$-fold compared to
269 controls kept in the dark, p=0.0032) (Fig. 6F). We also observed a slight but statistically
270 significant increase in GFP expression in embryos containing only the *Tg(C120F:GFP)*
271 transgene (1.5 $^{+0.3}_{-0.5}$-fold compared to controls kept in the dark, p=0.0386). However, given that
272 we did not observe any GFP fluorescence in 2 dpf *Tg(C120F:GFP)* embryos (Fig. 6A-B), this
273 slight increase in GFP mRNA levels is likely not functionally significant.
274
275 At 4 dpf, qPCR analysis detected elevated background GFP expression from all larvae kept in
276 the dark (Fig. 6E), presumably due to the ectopic liver expression in this transgenic line.
277 However, even with this higher background expression, we could detect significant induction of
278 GFP in response to light (6.6 $^{+3.8}_{-0.2}$-fold compared to controls kept in the dark, p=0.0025) in
279 *Tg(C120F:GFP);Tg(ubb:TAEL-N)* double transgenic larvae but not in *Tg(C120F:GFP)* larvae
280 (Fig. 6F). These results suggest that in the absence of activated TAEL-N, basal activity of the

281 *C120F* promoter is low and negligibly responsive to light, demonstrating specificity of the TAEL
282 2.0 system.
283
284 **Discussion**
285
286 In this study, we describe improvements we have made to a zebrafish-optimized optogenetic
287 expression system called TAEL/C120. In the original TAEL/C120 system, a LOV domain-
288 containing transcription factor (TAEL) is used to drive expression of genes of interest
289 downstream of the C120 regulatory element in response to blue light. The improvements we
290 made include adding a C-terminal nuclear localization signal to TAEL (TAEL-N) and coupling
291 C120 regulatory elements with a basal promoter taken from the mouse *Fos* gene (*C120F*).
292 These improvements, collectively referred to as TAEL 2.0, significantly increased both the level
293 and rate of light-induced expression.
294
295 Importantly, these improvements allowed us to generate functional stable transgenic lines for
296 TAEL-N expression. Previously under TAEL 1.0, we had difficulties generating such transgenic
297 lines, possibly due to sub-optimal performance of the TAEL transcriptional activator and/or
298 sensitivity of the C120 promoter. We speculate that these deficiencies were overcome in TAEL
299 1.0 by transiently expressing TAEL by mRNA or plasmid injection, which can deliver many more
300 molecules of TAEL than can be achieved by transgene expression. However, this approach
301 limits the applications for TAEL 1.0 as injections are labor intensive, introduce experimental
302 variability, and often preclude use beyond early embryonic stages. In this study, the
303 improvements we made to both the transcriptional activator (TAEL-N) and promoter (*C120F*)
304 together allowed us to generate functional TAEL-N transgenic lines. Such lines can provide
305 additional spatiotemporal specificity to gene induction, as demonstrated with the
306 *Tg(sox17:TAEL-N)* line (Fig. 4). And, as shown with the *Tg(ubb:TAEL-N)* line (Fig. 5),

307 transgenesis enables usage beyond early embryonic stages, which is not possible with mRNA
308 delivery to the zygote.
309
310 The choice of a basal promoter is often overlooked when designing zebrafish transgenes even
311 though it can have profound effects on the function of a transgene. In our original TAEL 1.0
312 system, the C120 regulatory element is coupled to a minimal TATA box sequence taken from a
313 mammalian expression vector[6,8]. In this study, we replaced the minimal TATA box with the basal
314 promoter of the mouse *Fos* gene, which was previously used in zebrafish transgenesis[15,21]. This
315 modification alone resulted in more than 40-fold activation following illumination — a 15-fold
316 increase over the original TAEL system (Fig. 1C). Several different basal promoters have been
317 used in zebrafish transgene and enhancer trap constructs, each with different
318 characteristics[15,21,23,24]. The *Fos* basal promoter is derived from a gene well-known for its
319 activation in response to neuronal activity[25]. Our experiments indicate that coupling this basal
320 promoter to the C120 regulatory sequence imparts several desirable attributes to the TAEL
321 system (fast induction, low background, high amplitude) that extend to the whole organism. For
322 cell type-specific applications, further improvement may be possible by choosing a different
323 basal promoter optimized for that cell type.
324
325 With the improvements that we have made, the TAEL 2.0 system further expands the multitude
326 of different applications we envision, including lineage tracing and precise targeting (spatially
327 and temporally) of gene perturbations. One major advantage of TAEL 2.0 is the extension of
328 these applications beyond early embryonic stages through transgene-directed expression of the
329 TAEL-N transcription factor. This improved zebrafish-optimized light-gated gene expression
330 system should be a broadly useful resource for the zebrafish community.
331
332 **Acknowledgements**

We thank Anna Reade, Didier Stainier, and Orion Weiner for support and advice during early stages of this work. We thank members of the Woo and Materna labs for providing insightful comments throughout the work and especially on the manuscript. We thank Chris Amemiya for generously providing access to the Lecia stereomicroscope. We thank Roy Hoglund, Emily Slocum, and their staff for excellent zebrafish care. This work was funded by grants from the National Institutes of Health (NIH) (DK106358) and the University of California Cancer Research Coordinating Committee (CRN-20-636896) to S.W. K.S. received support from the NSF-CREST: Center for Cellular and Biomolecular Machines at the University of California, Merced (NSF-HRD-1547848). K.H.G. acknowledges support from NIH (R01 GM106239).

Competing Interests

L.B.M-M. and K.H.G were co-founders of Optologix, Inc., which developed light-gated transcription factors for research applications. As of September 2020, Optologix, Inc. has ceased business.

349 **References:**

350

351 1. Halloran, M. C., Sato-Maeda, M., Warren, J. T., Su, F., Lele, Z., Krone, P. H., Kuwada, J.
352 Y. & Shoji, W. Laser-induced gene expression in specific cells of transgenic zebrafish.
353 *Development* **127,** 1953–1960 (2000).

354 2. Knopf, F., Schnabel, K., Haase, C., Pfeifer, K., Anastassiadis, K. & Weidinger, G. Dually
355 inducible TetON systems for tissue-specific conditional gene expression in zebrafish.
356 *Proc. Natl. Acad. Sci. U.S.A.* **107,** 19933–19938 (2010).

357 3. Liu, H., Gomez, G., Lin, S., Lin, S. & Lin, C. Optogenetic control of transcription in
358 zebrafish. *PLoS ONE* **7,** e50738 (2012).

359 4. Wang, X., Chen, X. & Yang, Y. Spatiotemporal control of gene expression by a light-
360 switchable transgene system. *Nat. Methods* **9,** 266–269 (2012).

361 5. Chen, X., Wang, X., Du, Z., Ma, Z. & Yang, Y. Spatiotemporal control of gene expression
362 in mammalian cells and in mice using the LightOn system. *Curr Protoc Chem Biol* **5,** 111–
363 129 (2013).

364 6. Reade, A., Motta-Mena, L. B., Gardner, K. H., Stainier, D. Y., Weiner, O. D. & Woo, S.
365 TAEL: a zebrafish-optimized optogenetic gene expression system with fine spatial and
366 temporal control. *Development* **144,** 345–355 (2017).

367 7. Rivera-Cancel, G., Motta-Mena, L. B. & Gardner, K. H. Identification of natural and
368 artificial DNA substrates for light-activated LOV-HTH transcription factor EL222.
369 *Biochemistry* **51,** 10024–10034 (2012).

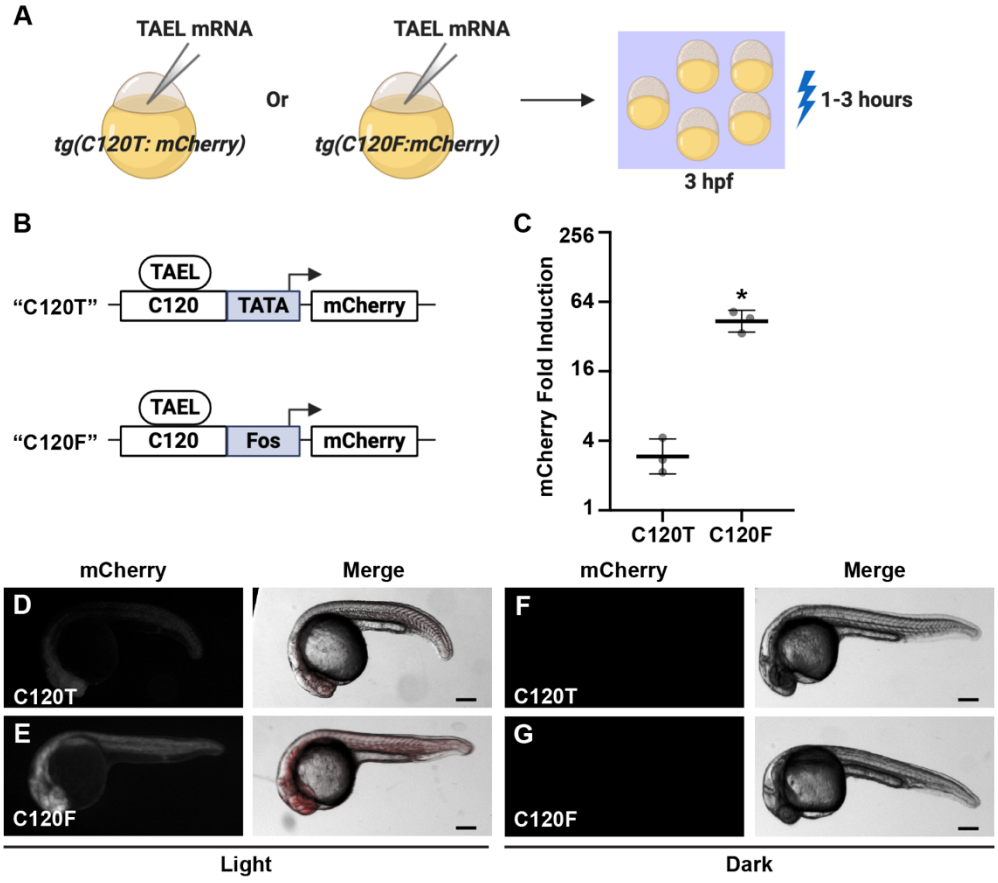
370 8. Motta-Mena, L. B., Reade, A., Mallory, M. J., Glantz, S., Weiner, O. D., Lynch, K. W. &
371 Gardner, K. H. An optogenetic gene expression system with rapid activation and
372 deactivation kinetics. *Nat. Chem. Biol.* **10,** 196–202 (2014).

373 9. Kawakami, K., Takeda, H., Kawakami, N., Kobayashi, M., Matsuda, N. & Mishina, M. A
374 transposon-mediated gene trap approach identifies developmentally regulated genes in
375 zebrafish. *Dev. Cell* **7,** 133–144 (2004).
376 10. Rupp, R. A., Snider, L. & Weintraub, H. Xenopus embryos regulate the nuclear
377 localization of XMyoD. *Genes Dev.* **8,** 1311–1323 (1994).
378 11. Turner, D. L. & Weintraub, H. Expression of achaete-scute homolog 3 in Xenopus
379 embryos converts ectodermal cells to a neural fate. *Genes Dev.* **8,** 1434–1447 (1994).
380 12. Norrander, J., Kempe, T. & Messing, J. Construction of improved M13 vectors using
381 oligodeoxynucleotide-directed mutagenesis. *Gene* **26,** 101–106 (1983).
382 13. Grzegorski, S. J., Chiari, E. F., Robbins, A., Kish, P. E. & Kahana, A. Natural variability of
383 Kozak sequences correlates with function in a zebrafish model. *PLoS ONE* **9,** e108475
384 (2014).
385 14. Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. A. & Smith, H. O.
386 Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6,**
387 343–345 (2009).
388 15. Scott, E. K. & Baier, H. The cellular architecture of the larval zebrafish tectum, as
389 revealed by gal4 enhancer trap lines. *Front Neural Circuits* **3,** 13 (2009).
390 16. Mizoguchi, T., Verkade, H., Heath, J. K., Kuroiwa, A. & Kikuchi, Y. Sdf1/Cxcr4 signaling
391 controls the dorsal migration of endodermal cells during zebrafish gastrulation.
392 *Development* **135,** 2521–2529 (2008).
393 17. Mosimann, C., Kaufman, C. K., Li, P., Pugach, E. K., Tamplin, O. J. & Zon, L. I.
394 Ubiquitous transgene expression and Cre-based recombination driven by the ubiquitin
395 promoter in zebrafish. *Development* **138,** 169–177 (2011).
396 18. Clark, K. J., Urban, M. D., Skuster, K. J. & Ekker, S. C. Transgenic zebrafish using
397 transposable elements. *Methods Cell Biol.* **104,** 137–149 (2011).

398 19. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time
399 quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25,** 402–408 (2001).
400 20. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of
401 image analysis. *Nat. Methods* **9,** 671–675 (2012).
402 21. Dorsky, R. I., Sheldahl, L. C. & Moon, R. T. A transgenic Lef1/beta-catenin-dependent
403 reporter is expressed in spatially restricted domains throughout zebrafish development.
404 *Dev. Biol.* **241,** 229–237 (2002).
405 22. Tyssowski, K. M. & Gray, J. M. Blue Light Increases Neuronal Activity-Regulated Gene
406 Expression in the Absence of Optogenetic Proteins. *eNeuro* **6,** ENEURO.0085–19.2019
407 (2019).
408 23. Scheer, N. & Campos-Ortega, J. A. Use of the Gal4-UAS technique for targeted gene
409 expression in the zebrafish. *Mech. Dev.* **80,** 153–158 (1999).
410 24. Scott, E. K., Mason, L., Arrenberg, A. B., Ziv, L., Gosse, N. J., Xiao, T., Chi, N. C.,
411 Asakawa, K., Kawakami, K. & Baier, H. Targeting neural circuitry in zebrafish using GAL4
412 enhancer trapping. *Nat. Methods* **4,** 323–326 (2007).
413 25. Schoenenberger, P., Gerosa, D. & Oertner, T. G. Temporal control of immediate early
414 gene induction by light. *PLoS ONE* **4,** e8185 (2009).
415

416 **Figures and Figure Legends**

417



418

419 **Figure 1. Coupling the C120 regulatory element to the Fos basal promoter significantly**

420 **increases light-induced expression. A.** Schematic of experimental design.

421 *Tg(C120T:mCherry)* or *Tg(C120F:mCherry)* embryos were injected with TAEL mRNA. mCherry

422 expression was induced by illuminating embryos with blue light starting at 3 hours post-

423 fertilization (hpf). **B.** Schematic comparing different C120-based reporter constructs in which

424 TAEL-responsive C120 sequences (C120) were coupled to either a minimal TATA box (TATA)

425 or the basal promoter from the mouse *Fos* gene (Fos) and used to drive expression of mCherry.

426 **C.** Comparison of light-induced mCherry expression in *Tg(C120T:mCherry)* and

427 *Tg(C120F:mCherry)* embryos injected with TAEL mRNA. mCherry transcript levels were

428 measured by qPCR from embryos illuminated with blue light for 1 hour and compared to sibling

429 embryos kept in the dark. Dots represent biological replicates. Solid lines represent mean. Error
430 bars represent S.D. *p<0.05. **D-G.** Representative images of mCherry fluorescence in
431 *Tg(C120T:mCherry)* (D, F) or *Tg(C120F:mCherry)* (E, G) embryos injected with TAEL mRNA
432 and illuminated with blue light for 3 hours (D, E) or kept in the dark (F, G). Images were
433 acquired between 20 and 24 hours post-illumination. Scale bars, 200 μm.
434

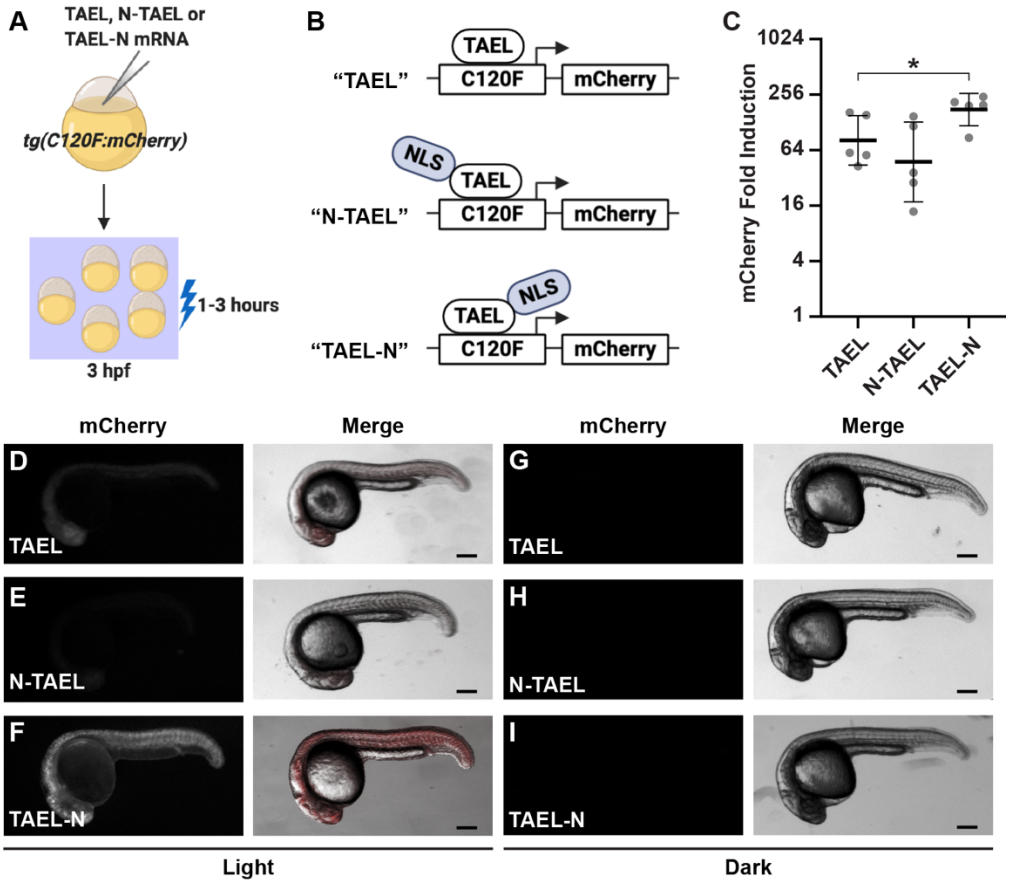


Figure 2. Adding a C-terminal nuclear localization signal (NLS) to TAEL significantly increases light-induced expression. A. Schematic of experimental design. *Tg(C120F:mCherry)* embryos were injected TAEL, N-TAEL, or TAEL-N mRNA. mCherry expression was induced by illuminating embryos with blue light starting at 3 hours post-fertilization (hpf). **B.** Schematic comparing different TAEL constructs containing no NLS (TAEL), one N-terminal NLS (N-TAEL), or one C-terminal NLS (TAEL-N). **C.** Comparison of light-induced mCherry expression in *Tg(C120F:mCherry)* embryos injected with TAEL, N-TAEL, or TAEL-N mRNA. mCherry transcript levels were measured by qPCR from embryos illuminated with blue light for 1 hour and compared to sibling embryos kept in the dark. Dots represent biological replicates. Solid lines represent mean. Error bars represent S.D. *$p<0.05$. **D-I.** Representative images of mCherry fluorescence in *Tg(C120F:mCherry)* embryos injected with TAEL (D, G), N-TAEL (E, H), or TAEL-N (F, I) mRNA and illuminated with blue light for 3 hours

448 (D-F) or kept in the dark (G-I). Images were acquired between 20 and 24 hours post-

449 illumination. Scale bars, 200 μm.

450

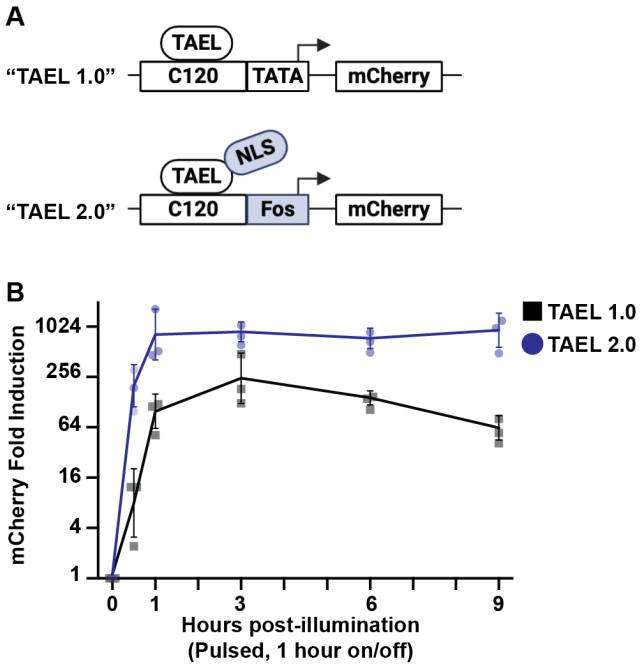


Figure 3. TAEL 2.0 modifications improve both the rate and level of light-induced expression. A. Schematic comparing TAEL 1.0 and TAEL 2.0. TAEL 1.0 consists of the TAEL transcription factor that lacks an NLS and the C120T promoter containing a minimal TATA box sequence. TAEL 2.0 consists of the TAEL-N transcription factor with a C-terminal NLS and the C120F promoter containing the basal promoter from the mouse *Fos* gene. **B.** Comparison of light-induced mCherry expression over time using TAEL 1.0 (black, dots) or TAEL 2.0 (blue, squares). *Tg(C120T:mCherry)* or *Tg(C120F:mCherry)* embryos were injected with mRNA for TAEL or TAEL-N, respectively. mCherry expression was induced by illuminating embryos with blue light (pulsed at a frequency of 1 hour on/1 hour off), starting at 3 hours post-fertilization. mCherry transcript levels were measured by qPCR at the indicated time points and normalized to 0 h post-illumination. Dots and squares represent biological replicates. Solid lines represent mean. Error bars represent S.D.

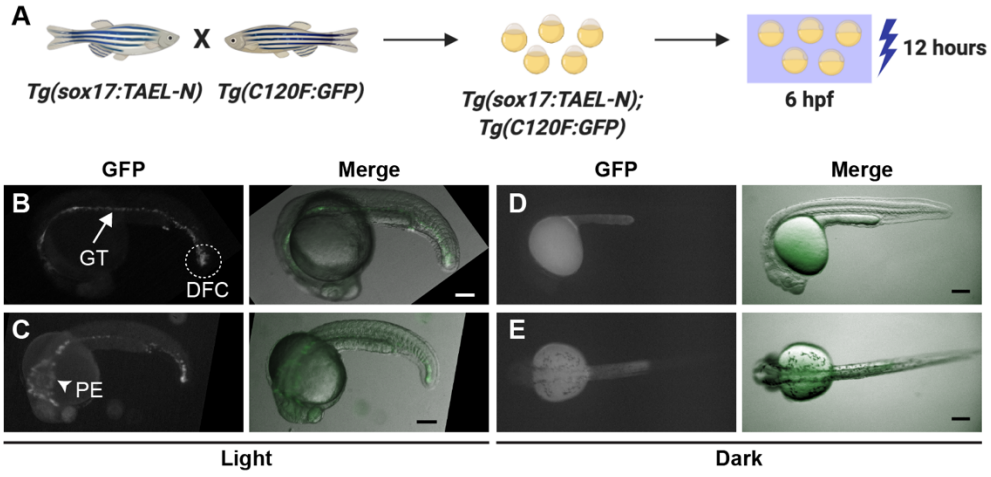


Figure 4. The stable transgenic line *Tg(sox17:TAEL-N)* restricts light-induced expression to endoderm-derived tissues. A Schematic depicting experimental design. *Tg(sox17:TAEL-N)* and *Tg(C120F:GFP)* adult zebrafish were crossed to produce double transgenic embryos. GFP expression was induced by illuminating embryos for 12 hours, starting at 6 hours post-fertilization (hpf), with blue light pulsed at a frequency of 1 hour on/1 hour off. **B-E.** Representative images of *Tg(sox17:TAEL-N);Tg(C120F:GFP)* embryos exposed to blue light (B-C) or kept in the dark (D-E). Images were acquired between 18 and 20 hours post-illumination. Arrow in (B) indicates gut tube (GT). Dashed lines in (B) indicate derivatives of the dorsal forerunner cells (DFC). Arrowhead in (C) indicates pharyngeal endoderm (PE). B, D are lateral views, anterior to the left. C, E are dorsal views, anterior to the left. Scale bars, 200 μm.

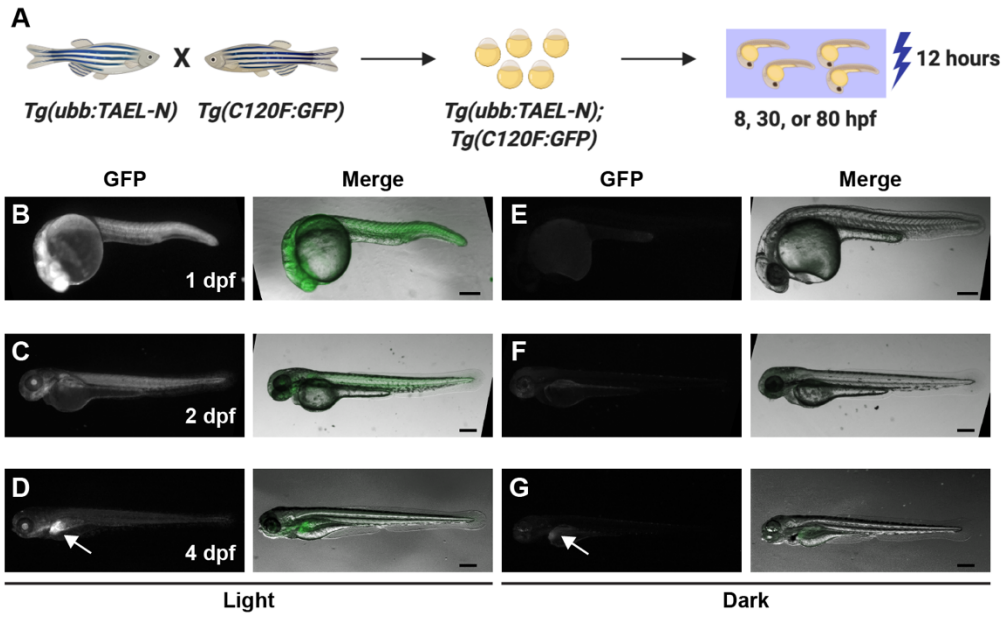


Figure 5. The stable transgenic line *Tg(ubb:TAEL-N)* enables light-induced expression at multiple developmental stages. A. Schematic depicting experimental design. *Tg(ubb:TAEL-N) and Tg(C120F:GFP)* adult zebrafish were crossed together to produce double transgenic embryos. GFP expression was induced at multiple time points by illuminating embryos for 12 hours with blue light pulsed at a frequency of 1 hour on/1 hour off. hpf, hours post-fertilization. **B-G.** Representative images of *Tg(ubb:TAEL-N);Tg(C120F:GFP)* embryos or larvae exposed to blue light (B-D) or kept in the dark (E-G). Images were acquired at the indicated stages between 18 and 20 hours post-illumination. dpf, days post-fertilization. Arrows in (D, G) point to ectopic liver expression of GFP. Scale bars, 200 nm.

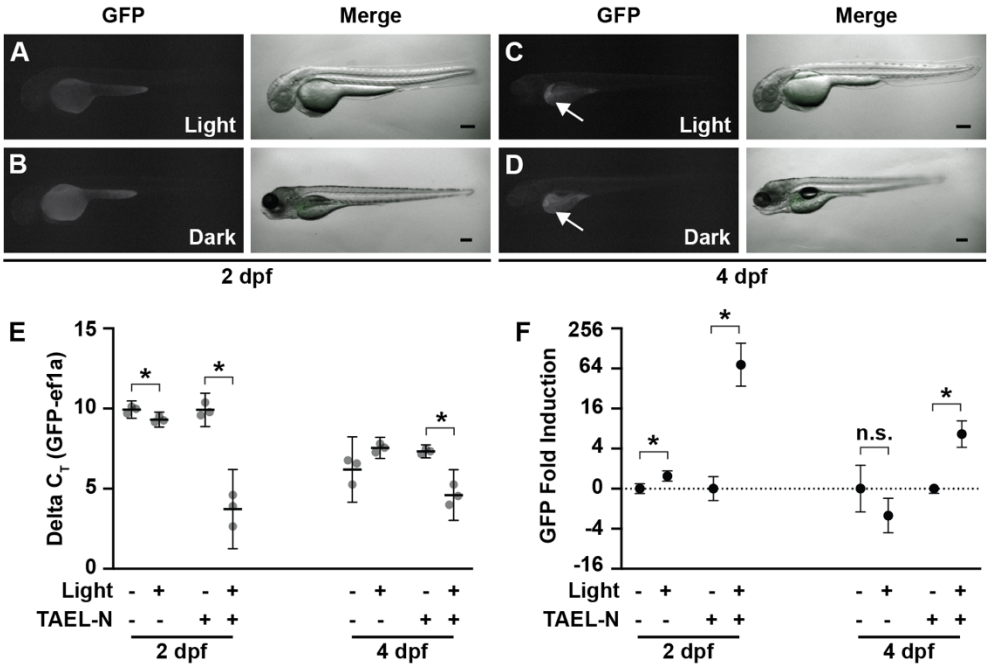


Figure 6. Basal expression from *Tg(C120F:GFP)* is not responsive to light. A-D. Representative images of *Tg(C120:GFP)* embryos at 2 days post-fertilization (dpf) (A-B) or larvae at 4 dpf (C-D). Embryos were illuminated for 12 hours with blue light pulsed at a frequency of 1 hour on/1 hour off (A, C) or kept in the dark (B, D). Images were acquired between 18 and 20 hours post-illumination. Arrows in (C, D) point to ectopic liver expression of GFP. Scale bars, 200 nm. **E.** qPCR analysis of GFP expression from *Tg(C120:GFP)* or *Tg(C120F:GFP);Tg(ubb:TAEL-N)* zebrafish at 2 or 4 dpf illuminated with constant blue light for 1 hour or kept in the dark. Data are presented as delta-C_T values normalized to the housekeeping gene *ef1a*. Dots represent biological replicates. Solid lines represent mean. Error bars represent S.D. *p<0.05. **F.** Fold induction of GFP expression in response to light calculated from the same qPCR analysis shown in (E). Data are presented as mean ± S.D. *p<0.05. n.s., not significant.