

Characterizing a standardized BioPart for BAG-specific expression in *C. elegans*

Xinglin Qian¹, Sarah AlHarbi¹, Christian Frøkjær-Jensen^{1§}

¹Bioscience, BESE, King Abdullah University of Science and Technology (KAUST), Thuwal, 23955-6900, Kingdom of Saudi Arabia

[§]To whom correspondence should be addressed: cfjensen@kaust.edu.sa

Abstract

Biological parts (BioParts) are modular and standardized DNA sequences that encode biological functions and contribute to the efficient biological engineering of complex systems. Here, we characterize a short BioPart ($Pf!\underline{p-17}$, 300 bp) for bright multicopy and single-copy BAG-specific expression starting from the gastrula stage in hermaphrodite and male *C. elegans*. We have generated standardized $Pf!\underline{p-17}$ cloning vectors for BAG-specific *gfp* and *mScarlet* expression compatible with extra-chromosomal arrays and for single-copy transgene insertion. The short $Pf!\underline{p-17}$ promoter is easy to generate by gene synthesis and has been incorporated into our online transgene design tool (www.wormbuilder.org/transgenebuilder).



ľ	Neuron	Promoter (gene)	Length (bp)	Expression verified	Expression (onset)	Expression (multicopy)	Expression (single-copy)	Expression CenGEN	Log2 specificity (2nd highest neuron)	gfp vector	mScarlet vector
1	BAG (2 cells)	flp-17	300	Herm & male	Gastrula stage	High (5/6)	High (5/6)	10 ^{-1.5}	10.0 (VA12)	pJRM4	pJRM3

Figure 1. Expression and Characterization of Pflp-17 in BAG Neurons of C. elegans:

A. In hermaphrodites (top panel) and males (bottom panel) a Pflp-17::mScarlet transgene is expressed exclusively in BAG neurons (20x magnification, scale bars = 100 μ m). **B**. Single-copy P<u>flp-17</u>::mScarlet expression in a hermaphrodite (20x bar = μm). characteristics. magnification, scale 20 С. <u>flp-17</u> promoter CenGEN data from https://cengen.shinyapps.io/CengenApp/ and https://cengen.textpressolab.com/. **D**. Schematic of standard *mScarlet* and *qfp* vectors for BAG-specific expression and identification. BsaI restriction sites allow exchanging the fluorescent reporter for a gene of interest by Golden Gate Assembly; ApaI (N-terminal) and SmaI (C-terminal) allow fluorophore fusions. The vectors are compatible with single-copy insertions (MosTI) and expression from extra-chromosomal arrays. E. Illustration of the transgenebuilder workflow. The short 300 bp P_{flp-17} was added as a standard promoter in our online tool for optimizing transgene design and gene synthesis (www.wormbuilder.org/transgenebuilder).

Description

Biological parts (BioParts), such as promoters, 3' UTRs, and protein-coding sequences, refer to DNA sequences that encode well-defined biological functions. Compatible, well-characterized BioParts can benefit the engineering of biological systems by enhancing reliability and reproducibility across laboratories (Shetty et al., 2008). In *E. coli* and *S. cerevisiae*, substantial effort has been put into developing standardized biological parts and toolkits (Malcı et al., 2022; Moore et al., 2016). In *C.*

elegans, the Fire lab vector kits are an early example of the benefits derived from the wide adoption of standardized reagents, and latter examples include comprehensive genome-wide BioPart collections, such as the promoterome (Dupuy et al., 2004), the ORFeome (Lamesch et al., 2004), the 3' UTRome (Steber et al., 2019), and the bacterial RNAi library (Kamath et al., 2003), which collectively serve as standard reagent platforms and toolkits for the community.

A collection of BioParts encoding promoters that are specific to individual classes of neurons is potentially useful for understanding and controlling the relatively simple *C. elegans* neuronal network consisting of 302 neurons belonging to 118 distinct neuronal classes (White et al., 1986). For example, neuron-specific expression enables optogenetic recording and manipulation of cellular activity (Suzuki et al., 2003) and genetic screens to identify factors required for neuronal specification (Hobert, 2021). Many studies have identified individual neuron-specific promoters (e.g., Inada et al., 2006; Sengupta et al., 1996; Yu et al., 1997), and there have been several large-scale efforts to comprehensively identify promoters expressed in single neurons (e.g., Dupuy et al., 2004; Hunt-Newbury et al., 2007). Furthermore, single-cell RNA expression data makes it increasingly feasible to test the expression of promoters from putative cell-specific transcripts (Lorenzo et al., 2020; Nava et al., 2023; Taylor et al., 2021). We have recently begun an effort to characterize and generate standardized reagents for neuron-specific expression, with a particular emphasis on identifying short promoters that are amenable to gene synthesis (AlHarbi & Frøkjær-Jensen, 2023). The *C. elegans* genome is relatively compact (Cutter et al., 2007); Froehlich & Rajewsky, 2023; Stefanakis et al., 2015; Wenick & Hobert, 2004), suggesting that it may be possible to identify a comprehensive collection of compact promoters. Here, we describe a short promoter for specific expression in BAG neurons, which are involved in sensing oxygen (Zimmer et al., 2009) and carbon dioxide (Bretscher et al., 2011).

To characterize a 300 bp shortened BAG-specific promoter (Brandt et al., 2012), P<u>flp-17</u>, we generated a fluorescent reporter construct (P<u>flp-17</u>::*mScarlet*::<u>tbb-2</u> 3'UTR) by gene synthesis. As part of the standardization, we removed *ApaI*, *SmaI*, and *BsaI* restriction sites to generate standardized reagents, added a standard start sequence ("aaaaATG"), and removed homopolymer runs, which interfere with gene synthesis. To determine expression, we generated transgenic animals with extrachromosomal arrays that carry many copies of the *mScarlet* reporter construct. We also included a piRNA interference (piRNAi) fragment targeting <u>him-5</u> (Priyadarshini et al., 2022) to induce males in transgenic animals. We imaged hermaphrodite and male transgenic worms and observed fluorescence starting at the gastrula stage with highly specific red fluorescence in two bilaterally symmetric neurons at the L4 stage that are consistent with the position and morphology of BAG neurons (**Figure 1A**). Expression was also readily visible from a single-copy transgene inserted by modular safe-harbor transgene insertion (MosTI) (El Mouridi et al., 2022) (**Figure 1B**), demonstrating that the short P*flp*-17 is relatively strong. We have developed a semi-quantitative visual score for expression strength (AlHarbi & Frøkjær-Jensen, 2023) (see methods), and the multicopy arrays scored 5 out of 6, and the single-copy insertion also scored 5 out of 6. The characteristics of the <u>flp-17</u> promoter are summarized in **Figure 1C**.

To enable easy use of the Pflp-17 BioPart we have generated standardized expression vectors containing red (*mScarlet*, pJRM3) and green (*gfp*, pJRM4) fluorophores in a plasmid backbone that is compatible with single-copy insertion and multicopy extra-chromosomal arrays (**Figure 1D**). The plasmids include restriction sites for N-terminal (*ApaI*) and C-terminal (*SmaI*) transgene fusions, sites for fluorophore exchange (*BsaI*), and a 5' decoy sequence to diminish misexpression (A. Fire, Fire lab kit). The plasmid design is identical to vectors for PVQ-specific expression (AlHarbi & Frøkjær-Jensen, 2023), and the plasmids are available from Addgene. Finally, the short Pflp-17 promoter is amenable to gene synthesis, and we have included the promoter in an online transgene design app (www.wormbuilder.org/transgenebuilder) (**Figure 1E**).

In conclusion, our ongoing efforts aim to develop a comprehensive collection of promoter BioParts for neuron-specific transgene expression. These design principles and reagents represent the initial steps towards establishing such a system. We aim to generate and characterize a complete collection of neuron-specific BioParts to advance our understanding of the nervous system and potentially encode novel circuits.

Methods

Molecular biology

We generated non-clonal synthetic transgenes by gene synthesis (Twist Bioscience, CA, USA) and clonal plasmids by Golden-Gate cloning (Engler et al., 2008) using *Esp*3I (New England Biolabs). We validated clonal plasmids by restriction digest. We shortened the *flp-17* promoter tested by Lorenzo et al., (2020) to 300 base pairs and removed *BsaI*, *Esp*3I, *ApaI*, *SmaI*, and *Eco*RV, as well as homopolymer runs. We changed the last four basepairs of the promoter to the consensus start site (aaaaATG).

Promoter strength quantification and expression onset determination

We quantified the fluorescence intensity of *mScarlet* reporter constructs using a scheme first described by Alharbi and Frøkjær-Jensen (2023). Transgenic animals with stable extra-chromosomal arrays under a dissection stereo microscope (ZEISS, Olympus SZX2-FOF) or an upright compound microscope (LEICA DM2500 LED), objectives 1x and 10x using an mTomato filter set and an LED light (X-Cite XYLIS, XT720L) with 40x oil immersion objective, a Rhodamine filter set (LEICA 11504205), and a mercury metal halide bulb (LEICA EL6000). We quantified fluorescence intensity visually in ten L4 animals at different magnifications. We performed the quantification by eye and scored on a scale from 1 to 6, with 1 being the dimmest and 6 being the brightest. We used the following scoring criteria. **Dissection microscope**. Score = 6: fluorescence visible with 1x objective and zoom = 1 (lowest). Score = 5: fluorescence visible with 1x objective and zoom = 8 (highest). Score = 4: fluorescence visible with 10x objective and zoom = 1. Score = 3: fluorescence visible with 10x objective and zoom = 8 (highest) and zoom = 8. **Compound microscope**. Score = 0: fluorescence not visible at 40x oil immersion objective. We screened 15 embryos from transgenic animals at various stages (gastrula, comma, 1.5-fold, 2-fold, and 3-fold) on NGM plates for the first visible expression.

Microscopy

We anesthetized transgenic animals on 2% agarose pads with a 50 mM sodium-azide M9 solution. Extra-chromosomal array animals were imaged with a Leica THUNDER Imaging System, equipped with a 20× oil immersion objective, and captured four image stacks—two for hermaphrodites and two for males. The images show a maximum intensity projection generated with the Leica LAS X software. Animals with single-copy insertions were imaged on a LEICA DM2500 LED using a 42x oil immersion objective.

Transgenic animals

To generate extrachromosomal arrays animals, we injected the injection mix composed of 10 $ng \cdot \mu L^{-1}$ non-clonal Pf[p-<u>17</u>::mScarlet::<u>tbb-2</u> 3' UTR dsDNA fragment, 10 $ng \cdot \mu L^{-1}$ pCFJ108 (<u>unc-119</u> rescue, linearized by *ApaLI*), 10 $ng \cdot \mu L^{-1}$ pCFJ782 (HygroR, linearized by *Eco*RV), 10 $ng \cdot \mu L^{-1}$ pMNK54 (piRNAi <u>him-5</u>, linearized by *ApaLI*), and 60 $ng \cdot \mu L^{-1}$ GeneRuler 1 kb plus DNA ladder (ThermoFisher SM1331) for a final concentration of 100 $ng \cdot \mu L^{-1}$, into <u>CFJ42</u> animals (MosTI II, <u>unc-119(ed3</u>) III). We kept injected animals at 25 °C and added hygromycin to the plates on day 3. We identified transgenic animals with extrachromosomal arrays based on <u>unc-119</u> rescue and antibiotic resistance.

Single-copy transgenes with pJRM3 (P<u>flp-17</u>::*mScarlet*::*tbb-2* 3' UTR) were inserted by MosTI into a landing site on Chr II (<u>CFJ42</u>, *ttTi5605* site) following standard protocols (El Mouridi et al., 2022). The injection mix was comprised of 20 ng·µL⁻¹ pJRM3 (P<u>flp-17</u>::*mScarlet*) MosTI targeting vector, 15 ng·µL⁻¹ pSEM318 (MosTI sgRNA, linearized by *NdeI*), 10 ng·µL⁻¹ pSEM231 (P<u>mlc-1</u>::*gfp*, co-injection marker, linearized by *NdeI*) (El Mouridi et al., 2020), 15 ng·µL⁻¹ pCFJ782 (hygromycin resistance, linearized by *Eco*RI) (Radman et al., 2013), 25 ng·µL⁻¹ pMDJ231 (heat-shock Cas9, linearized by *ApaLI*), and 15 ng·µL⁻¹ GeneRuler 1 kb plus DNA ladder (ThermoFisher SM1331) for a final concentration of 100 ng·µL⁻¹. We kept injected animals at 25 °C and added hygromycin to the plates on day 3. When the bacterial lawn was nearly exhausted, we heat-shocked the plates with transgenic animals in a 30 °C air incubator for 20 hours and kept animals at 20 °C for four days. We identified transgenic animals with a single-copy insertion four days after the heat shock based on <u>unc-119</u> rescue and the absence of green co-injection markers.

Software

We performed *in silico* design with "A plasmid Editor" (ApE) (Davis & Jorgensen, 2022), and generated the figure with Adobe Illustrator (v28.2).

A frozen version of the transgenebuilder software was described in Alharbi & Frøkjær-Jensen (2023) and is available at https://doi.org/10.22002/qs7eh-g0669.

Reagents

Plasmids

pJRM3 - P<u>flp-17</u>::*mScarlet*::<u>tbb-2</u> 3' UTR (MosTI compatible) (Addgene #204616) pJRM4 - P<u>flp-17</u>::*gfp*::<u>tbb-2</u> 3' UTR (MosTI compatible) (Addgene #204617) pMNK54 - <u>him-5</u> piRNAi fragment (Addgene #159818) (Priyadarshini et al., 2022) pCFJ108 - *cbr-unc-119* (Addgene #200367)



pSEM231 - <u>mlc-1</u>p::gfp::cbr-<u>tbb-2</u> (Addgene # 159897) (El Mouridi et al., 2020)

pSEM318 - *<u>rpr-1</u>*p::sgRNA targeting the *<u>ttTi5605</u> location Chr. II (Addgene # 159822)*

pMDJ231 – <u>hsp-16.41</u>p::cas9::<u>gpd-2</u>::tagRFP-t::smu-1 (Addgene # 191382)

pCFJ782 - P<u>rps-0</u>::hygromycinR::<u>rps-27</u> 3' UTR (Addgene #190933) (Radman et al., 2013).

Strains

<u>N2</u> – (wildtype)

<u>CFJ42</u> – *kstSi42*[<u>unc-119(p1, spc2)(-)] II; <u>unc-119(ed3</u>) III (available from CGC)</u>

Promoter sequence

WormBase ID	Gene Name	Promoter Sequence (modifications in upper case, consensus start site in bold)
WBGene00001460	f <u>lp-17</u>	tttctcagaagcctctaagtttcccgtgctttcttgtcaaattgattaggaaatgattcggattttctgaatttttgaaagatttttcggttttttcagtgatttcaatcggaaattcggagccggaaatgcattaggaaggggctaatagctataaaattgacacgggaaattcagattttcattca

Acknowledgements:

We thank WormBase (supported by grant P41-HG02223 from the US National Human Genome Research Institute and the British Medical Research Council). Some strains were provided by the CGC, which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440). We thank all Frøkjær-Jensen lab members for discussions.

References

AlHarbi S, Frøkjær-Jensen C. 2023. Characterizing a standardized BioPart for PVQ-specific expression in C. elegans. MicroPubl Biol 2023. PubMed ID: <u>37426742</u>

Brandt JP, Aziz-Zaman S, Juozaityte V, Martinez-Velazquez LA, Petersen JG, Pocock R, Ringstad N. 2012. A single gene target of an ETS-family transcription factor determines neuronal CO2-chemosensitivity. PLoS One 7(3): e34014. PubMed ID: 22479504

Bretscher AJ, Kodama-Namba E, Busch KE, Murphy RJ, Soltesz Z, Laurent P, de Bono M. 2011. Temperature, oxygen, and salt-sensing neurons in C. elegans are carbon dioxide sensors that control avoidance behavior. Neuron 69(6): 1099-113. PubMed ID: <u>21435556</u>

Cutter AD, Dey A, Murray RL. 2009. Evolution of the Caenorhabditis elegans genome. Mol Biol Evol 26(6): 1199-234. PubMed ID: <u>19289596</u>

Davis MW, Jorgensen EM. 2022. ApE, A Plasmid Editor: A Freely Available DNA Manipulation and Visualization Program. Front Bioinform 2: 818619. PubMed ID: <u>36304290</u>

Dupuy D, Li QR, Deplancke B, Boxem M, Hao T, Lamesch P, et al., Vidal M. 2004. A first version of the Caenorhabditis elegans Promoterome. Genome Res 14(10B): 2169-75. PubMed ID: <u>15489340</u>

El Mouridi S, Alkhaldi F, Frøkjær-Jensen C. 2022. Modular safe-harbor transgene insertion for targeted single-copy and extrachromosomal array integration in Caenorhabditis elegans. G3 (Bethesda) 12(9). PubMed ID: <u>35900171</u>

El Mouridi S, Peng Y, Frøkjær-Jensen C. 2020. Characterizing a strong pan-muscular promoter (Pmlc-1) as a fluorescent coinjection marker to select for single-copy insertions. MicroPubl Biol 2020. PubMed ID: <u>32908967</u>

Engler C, Kandzia R, Marillonnet S. 2008. A one pot, one step, precision cloning method with high throughput capability. PLoS One 3(11): e3647. PubMed ID: <u>18985154</u>

Etchberger JF, Lorch A, Sleumer MC, Zapf R, Jones SJ, Marra MA, et al., Hobert O. 2007. The molecular signature and cisregulatory architecture of a C. elegans gustatory neuron. Genes Dev 21(13): 1653-74. PubMed ID: <u>17606643</u>

Froehlich JJ, Rajewsky N. 2023. C. elegans Gene Regulatory Alleles and Reporter Bashing Studies. MicroPubl Biol 2023. PubMed ID: <u>36748042</u>

Hobert O. 2021. Homeobox genes and the specification of neuronal identity. Nat Rev Neurosci 22(10): 627-636. PubMed ID: <u>34446866</u>

Hunt-Newbury R, Viveiros R, Johnsen R, Mah A, Anastas D, Fang L, et al., Moerman DG. 2007. High-throughput in vivo analysis of gene expression in Caenorhabditis elegans. PLoS Biol 5(9): e237. PubMed ID: <u>17850180</u>

Inada H, Ito H, Satterlee J, Sengupta P, Matsumoto K, Mori I. 2006. Identification of guanylyl cyclases that function in thermosensory neurons of Caenorhabditis elegans. Genetics 172(4): 2239-52. PubMed ID: <u>16415369</u>

Kamath RS, Fraser AG, Dong Y, Poulin G, Durbin R, Gotta M, et al., Ahringer J. 2003. Systematic functional analysis of the Caenorhabditis elegans genome using RNAi. Nature 421(6920): 231-7. PubMed ID: <u>12529635</u>

Lamesch P, Milstein S, Hao T, Rosenberg J, Li N, Sequerra R, et al., Vidal M. 2004. C. elegans ORFeome version 3.1: increasing the coverage of ORFeome resources with improved gene predictions. Genome Res 14(10B): 2064-9. PubMed ID: <u>15489327</u>

Lorenzo R, Onizuka M, Defrance M, Laurent P. 2020. Combining single-cell RNA-sequencing with a molecular atlas unveils new markers for Caenorhabditis elegans neuron classes. Nucleic Acids Res 48(13): 7119-7134. PubMed ID: <u>32542321</u>

Malcı K, Watts E, Roberts TM, Auxillos JY, Nowrouzi B, Boll HO, et al., Rios-Solis L. 2022. Standardization of Synthetic Biology Tools and Assembly Methods for Saccharomyces cerevisiae and Emerging Yeast Species. ACS Synth Biol 11(8): 2527-2547. PubMed ID: <u>35939789</u>

Moore SJ, Lai HE, Kelwick RJ, Chee SM, Bell DJ, Polizzi KM, Freemont PS. 2016. EcoFlex: A Multifunctional MoClo Kit for E. coli Synthetic Biology. ACS Synth Biol 5(10): 1059-1069. PubMed ID: <u>27096716</u>

Nava S, Palma W, Wan X, Oh JY, Gharib S, Wang H, et al., Sternberg PW. 2023. A cGAL-UAS bipartite expression toolkit for Caenorhabditis elegans sensory neurons. Proc Natl Acad Sci U S A 120(51): e2221680120. PubMed ID: <u>38096407</u>

Priyadarshini M, Ni JZ, Vargas-Velazquez AM, Gu SG, Frøkjær-Jensen C. 2022. Reprogramming the piRNA pathway for multiplexed and transgenerational gene silencing in C. elegans. Nat Methods 19(2): 187-194. PubMed ID: <u>35115715</u>

Radman I, Greiss S, Chin JW. 2013. Efficient and rapid C. elegans transgenesis by bombardment and hygromycin B selection. PLoS One 8(10): e76019. PubMed ID: <u>24130756</u>

Sengupta P, Chou JH, Bargmann CI. 1996. odr-10 encodes a seven transmembrane domain olfactory receptor required for responses to the odorant diacetyl. Cell 84(6): 899-909. PubMed ID: <u>8601313</u>

Shetty RP, Endy D, Knight TF Jr. 2008. Engineering BioBrick vectors from BioBrick parts. J Biol Eng 2: 5. PubMed ID: <u>18410688</u>

Steber HS, Gallante C, O'Brien S, Chiu PL, Mangone M. 2019. The C. elegans 3' UTRome v2 resource for studying mRNA cleavage and polyadenylation, 3'-UTR biology, and miRNA targeting. Genome Res 29(12): 2104-2116. PubMed ID: <u>31744903</u>

Stefanakis N, Carrera I, Hobert O. 2015. Regulatory Logic of Pan-Neuronal Gene Expression in C. elegans. Neuron 87(4): 733-50. PubMed ID: <u>26291158</u>

Suzuki H, Kerr R, Bianchi L, Frøkjaer-Jensen C, Slone D, Xue J, et al., Schafer WR. 2003. In vivo imaging of C. elegans mechanosensory neurons demonstrates a specific role for the MEC-4 channel in the process of gentle touch sensation. Neuron 39(6): 1005-17. PubMed ID: <u>12971899</u>

Taylor SR, Santpere G, Weinreb A, Barrett A, Reilly MB, Xu C, et al., Miller DM 3rd. 2021. Molecular topography of an entire nervous system. Cell 184(16): 4329-4347.e23. PubMed ID: <u>34237253</u>

Wenick AS, Hobert O. 2004. Genomic cis-regulatory architecture and trans-acting regulators of a single interneuron-specific gene battery in C. elegans. Dev Cell 6(6): 757-70. PubMed ID: <u>15177025</u>

White JG, Southgate E, Thomson JN, Brenner S. 1986. The structure of the nervous system of the nematode Caenorhabditis elegans. Philos Trans R Soc Lond B Biol Sci 314(1165): 1-340. PubMed ID: <u>22462104</u>

Yu S, Avery L, Baude E, Garbers DL. 1997. Guanylyl cyclase expression in specific sensory neurons: a new family of chemosensory receptors. Proc Natl Acad Sci U S A 94(7): 3384-7. PubMed ID: <u>9096403</u>



Zimmer M, Gray JM, Pokala N, Chang AJ, Karow DS, Marletta MA, et al., Bargmann CI. 2009. Neurons detect increases and decreases in oxygen levels using distinct guanylate cyclases. Neuron 61(6): 865-79. PubMed ID: <u>19323996</u>

Funding:

This work was funded by a competitive research grant from the KAUST Office of Research Administration (CRG9 URF/1/4388-01-01).

Author Contributions: Xinglin Qian: investigation, writing - original draft, writing - review editing. Sarah AlHarbi: investigation, writing - review editing. Christian Frøkjær-Jensen: funding acquisition, conceptualization, supervision, writing - review editing.

Reviewed By: Anonymous

Nomenclature Validated By: Daniela Raciti

WormBase Paper ID: WBPaper00066554

History: Received February 7, 2024 Revision Received February 25, 2024 Accepted March 4, 2024 Published Online March 12, 2024 Indexed March 26, 2024

Copyright: © 2024 by the authors. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International (CC BY 4.0) License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Citation: Qian, X; AlHarbi, S; Frøkjær-Jensen, C (2024). Characterizing a standardized BioPart for BAG-specific expression in *C. elegans*. microPublication Biology. <u>10.17912/micropub.biology.001150</u>