

Deterministic Genetic Barcoding for Multiplexed Behavioral and Single-Cell Transcriptomic Studies

Reviewed Preprint

v2 • January 3, 2025

Revised by authors

Reviewed Preprint

v1 • August 1, 2023

Jorge Blanco Mendana, Margaret Donovan, Lindsey G O'Brien, Benjamin Auch, John Garbe, Daryl M Gohl 

University of Minnesota Genomics Center, Minneapolis, USA • Department of Genetics, Cell Biology, and Development, University of Minnesota, Minneapolis, USA

 https://en.wikipedia.org/wiki/Open_access

 Copyright information

eLife Assessment

This **useful** study presents a genetically encoded barcoding system that could advance transcriptomic studies and that has the potential for further applications, such as in high-throughput population-scale behavioral measurements. The evidence supporting the claims of the authors is **solid** and highlights both the usefulness and the limitations of the approach.

<https://doi.org/10.7554/eLife.88334.2.sa3>

Abstract

Advances in single-cell sequencing technologies have provided novel insights into the dynamics of gene expression throughout development, been used to characterize somatic variation and heterogeneity within tissues, and are currently enabling the construction of transcriptomic cell atlases. However, despite these remarkable advances, linking anatomical information to transcriptomic data and positively identifying the cell types that correspond to gene expression clusters in single-cell sequencing data sets remains a challenge. We describe a straightforward genetic barcoding approach that takes advantage of the powerful genetic tools available in *Drosophila* to allow *in vivo* tagging of defined cell populations. This method, called Targeted Genetically-Encoded Multiplexing (TaG-EM), involves inserting a DNA barcode just upstream of the polyadenylation site in a Gal4-inducible *UAS-GFP* construct so that the barcode sequence can be read out during single-cell sequencing, labeling a cell population of interest. By creating many such independently barcoded fly strains, TaG-EM will enable a number of potential applications that will improve the quality and information content of single-cell transcriptomic data including positive identification of cell types in cell atlas projects, identification of multiplet droplets, and barcoding of experimental timepoints, conditions, and replicates. Furthermore, we demonstrate that the barcodes from TaG-EM fly lines can be read out using next-generation sequencing to facilitate population-scale behavioral measurements. Thus, TaG-EM has the potential to enable large-scale behavioral screens in addition to improving the ability to reliably annotate cell atlas data, expanding the scope, and improving the robustness of single-cell transcriptomic experiments.

Introduction

Spatially and temporally regulated gene expression patterns are a hallmark of multicellular life and function to orchestrate patterning, growth, and differentiation throughout development (Ingham, 1988 [↗](#); Reeves et al., 2006 [↗](#)). In mature organisms, spatial expression patterns both in tissues and within cells define functionally distinct compartments and determine many aspects of cellular and organismal physiology (Martin and Ephrussi, 2009 [↗](#)). In addition, such expression patterns differentiate healthy and diseased tissue and impact disease etiology (Marusyk et al., 2012 [↗](#)). Spatial and temporal expression patterns, which can be used to distinguish between cell types and provide insight into cellular function, also provide a means to understand the organization and physiology of complex tissues such as the brain (Thompson et al., 2014 [↗](#)). Thus, robust and scalable tools for measuring spatial and temporal gene expression patterns at a genome-wide scale and at high resolution would be transformative research tools across many biological disciplines.

Single-cell sequencing technologies have provided insights into the dynamics of gene expression throughout development, been used to characterize somatic variation and heterogeneity within tissues, and are currently enabling the construction of transcriptomic cell atlases (Klein et al., 2015 [↗](#); Macosko et al., 2015 [↗](#); Zheng et al., 2017 [↗](#)). However, linking anatomical information to transcriptomic data and positively identifying the cell types that correspond to gene expression clusters in single-cell sequencing data sets remains a challenge. The cellular identities of gene expression clusters identified in cell atlas data sets are typically inferred from the expression of distinctive gene sets (Hung et al., 2020 [↗](#); Li et al., 2022 [↗](#); Ma et al., 2021 [↗](#)), and the lack of positive identification of gene expression clusters introduces an element of uncertainty in this analysis. Moreover, this process of manual annotation is labor-intensive and often requires additional experiments to determine or confirm the expression patterns of marker genes. Emerging spatial genomics technologies hold promise in linking anatomical and transcriptomic information (Lee et al., 2015 [↗](#); Lein et al., 2017 [↗](#)). Several of the emerging commercial spatial genomics technologies rely on *in situ* sequencing of marker genes allowing droplet-based single-cell transcriptomic data to be mapped onto a tissue. However, these technologies currently suffer from constraints related to cost, content, or applicability to specific model systems.

In addition to descriptive cell atlas projects, studies involving multiple experimental timepoints throughout development and aging, or studies assessing the effects of experimental exposures or genetic manipulations would benefit from increased ability to multiplex samples. Given the fixed costs of droplet-based single-cell sequencing, generating data for single-cell transcriptomic time courses or experimental manipulations can be costly. Outside of descriptive studies, these costs are also a barrier to including replicates to assess biological variability; consequently, a lack of biological replicates derived from independent samples is a common shortcoming of single-cell sequencing experiments. Antibody-based cell hashing or feature barcoding approaches have been developed to allow multiplexing of samples in droplet-based single-cell sequencing reactions (Stoeckius et al., 2018 [↗](#), 2017 [↗](#)). In addition, other multiplexing strategies for single-cell sequencing based on alternative methods for tagging cells (Cheng et al., 2021 [↗](#)) or making use of natural genetic variation have been used (Kurmangaliyev et al., 2020 [↗](#)). While such approaches can reduce per-sample costs, typically samples are barcoded at a population level and thus do not enable labeling of cell subpopulations within a sample.

We have developed a straightforward genetic barcoding approach that takes advantage of the powerful genetic tools available in *Drosophila* to allow deterministic *in vivo* tagging of defined cell populations. This method, called Targeted Genetically Encoded Multiplexing (TaG-EM), involves

inserting a DNA barcode just upstream of the poly-adenylation site in a Gal4-inducible *UAS-GFP* construct so that the barcode sequence can be read out during droplet-based single-cell sequencing, labeling a cell population of interest.

Genetic barcoding approaches have been employed in many unicellular systems, cell culture, and viral transfection to facilitate high throughput screening using sequencing-based readouts (Bhang et al., 2015 [↗](#); Smith et al., 2009 [↗](#); van Opijnen et al., 2009 [↗](#)). In multicellular animals, techniques such as GESTALT have enabled lineage tracing by using CRISPR to create unique barcodes in differentiating tissue (McKenna et al., 2016 [↗](#)), and barcode sequencing has also been employed to map connectivity in the brain (Chen et al., 2019 [↗](#)).

Genetically barcoded fly lines can also be used to enable highly multiplexed behavioral assays which can be read out using high throughput sequencing. Flies carrying TaG-EM barcodes can be exposed to different experimental perturbations and then tested in assays where flies, larvae, or embryos are fractionated based on behavioral outcomes or other phenotypes. Thus, TaG-EM has the potential to enable large-scale next-generation sequencing (NGS)-based behavioral or other fractionation screens analogous to BAR-Seq or Tn-Seq approaches employed in microbial organisms.

Results

TaG-EM: A novel genetic barcoding strategy for multiplexed behavioral and single-cell transcriptomics

We cloned a fragment containing a PCR handle sequence and a diverse 14 bp barcode sequence into the SV40 3' untranslated region (UTR) sequence just upstream of the polyadenylation sites in the 10xUAS-myr::GFP (pJFRC12, (Pfeiffer et al., 2010 [↗](#))) backbone (**Figure 1A** [↗](#)). A pool containing 29 unique barcode-containing plasmids was injected into *Drosophila* embryos for PhiC31-mediated integration into the attP2 landing site (Groth et al., 2004 [↗](#)) and transgenic lines were isolated and confirmed by Sanger sequencing (**Figure 1B** [↗](#), **Supplemental Figure 1** [↗](#)). We recovered 20 distinctly barcoded *Drosophila* lines, with some barcodes recovered from multiple crosses (**Supplemental Figure 1** [↗](#)). Such barcoded fly lines have the potential to enable population behavioral measurements, where different exposures, experimental timepoints, and genetic or neural perturbations can be multiplexed and analyzed by measuring barcode abundance in sequencing data (**Figure 1C** [↗](#)). In addition, the barcodes, which reside on a Gal4-inducible *UAS-GFP* construct, can be expressed tissue-specifically and read out during droplet-based single-cell sequencing, labeling a cell population and/or an experimental condition of interest (**Figure 1D** [↗](#)).

Testing the accuracy and reproducibility of TaG-EM behavioral measurements using structured pools

We conducted initial experiments to optimize amplification of the genetic barcodes using primers targeting the PCR handle inserted just upstream of the 14 bp barcode sequence and PCR primers downstream of the TaG-EM barcode in the SV40 3' UTR sequence (**Supplemental Figure 2** [↗](#)). To test the accuracy and reproducibility of sequencing-based measurements of TaG-EM barcodes, we constructed structured pools containing defined numbers of flies pooled either evenly with each of the 20 barcode constructs comprising 5% of the pool, or in a staggered manner with sets of barcodes differing in abundance in 2-fold increments (**Figure 2A** [↗](#)). To examine the impact of technical steps such as DNA extraction and PCR amplification on TaG-EM barcode measurements, even pools were made and extracted in triplicate and amplicon sequencing libraries were made in triplicate for each independently extracted DNA sample for both the even and staggered pools. The

resulting data indicated that TaG-EM measurements are highly accurate and reproducible. Technical replicates (indicated by error bars in **Figure 2B-E**) showed minimal variability. Likewise, the three independently extracted replicates of the even pools produced consistent data with all 20 barcodes detected at levels close to the expected 5% abundance (**Figure 2B-C**). Barcode abundance values for the staggered structured pools was generally consistent with the input values and in most cases, the 2-fold differences between the different groups of barcodes could be distinguished (**Figure 2D-E**). The coefficients of variation were largely consistent for groups of TaG-EM barcodes pooled evenly or at different levels within the staggered pools (**Supplemental Figure 3**). For the staggered pools, abundances correlated well with the expected values, particularly when multiple barcodes for an input level were averaged, in which case R^2 values were >0.99 (**Figure 2D-E**, inset plots). This indicates that a high level of quantitative accuracy can be attained using sequencing-based analysis of TaG-EM barcode abundance, particularly when averaging data for three to four independent barcodes for an experimental condition.

TaG-EM measurement of phototaxis behavior correlate well with video-based measurements

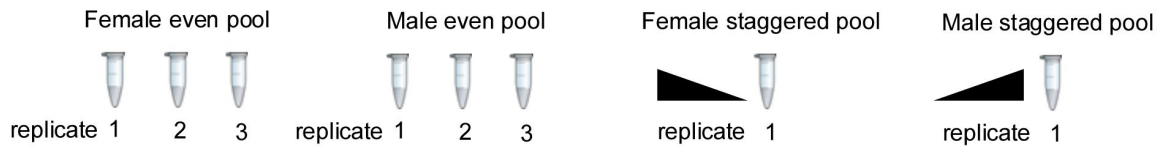
Next, we tested whether TaG-EM could be used to measure a phototaxis behavior. A mixture of barcoded wild type or blind *norpA* mutant flies were run together through a phototaxis assay. At the end of a period of light exposure, test tubes facing toward or away from the light were capped, DNA was isolated, and barcodes were amplified and sequenced for each tube. Raw read counts were scaled in proportion to the number of flies per tube and a preference index was calculated for each barcode (**Figure 3A**). In parallel, individual preference indices were calculated based on manual scoring of videos recorded for each line (**Figure 3B**). Preference indices calculated for the pooled, NGS-based TaG-EM measurements were nearly identical to conventional behavioral measurements for both wild type and *norpA* mutants (**Figure 3A-B**).

TaG-EM measurement of oviposition behavior and age-dependent fecundity

We next tested whether NGS-based pooled measurements of egg laying could be made. Fertilized females from each of the 20 barcode lines were placed together in egg laying cups, embryos were collected, aged for 12 hours to enable cell numbers to stabilize in the developing eggs, and then DNA was extracted from both the pooled adult flies and the embryos. In general, TaG-EM measurements of oviposition correlated with fly numbers, with the exception of barcode 14 which had reduced barcode abundance across multiple trials (**Supplemental Figure 4**). This suggests that despite the fact that the genetic barcode constructs are inserted in a common landing site, differences with respect to specific behaviors may exist among the lines, and thus one should test to make sure given lines are appropriate to use in specific behavioral assays.

To determine whether TaG-EM could be used to measure age-dependent fecundity, we collected flies from twelve different TaG-EM barcode lines at four time points separated by one week (three barcode lines per timepoint). We collected eggs from these fly lines individually and scored the number of viable eggs per female. Next, we pooled the barcoded flies from all timepoints and collected eggs from the pooled flies. These eggs were aged, DNA was extracted, and the TaG-EM barcodes were amplified and sequenced. While measurements from individual barcode lines were noisy, both for manual counts and sequencing based measurements, there was a general trend toward declining fecundity over time (**Supplemental Figure 5**), consistent with published reports (David et al., 1975). Manually scored viable egg numbers and TaG-EM barcode abundances were well correlated across two independent experimental trials (R^2 values of 0.52-

A Structured pool tests



	Barcode																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Female even	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
Male even	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
Female staggered	64	64	64	32	32	32	16	16	16	16	8	8	8	8	4	4	4	2	2	2
Male staggered	2	2	2	4	4	4	8	8	8	8	16	16	16	16	32	32	32	64	64	64

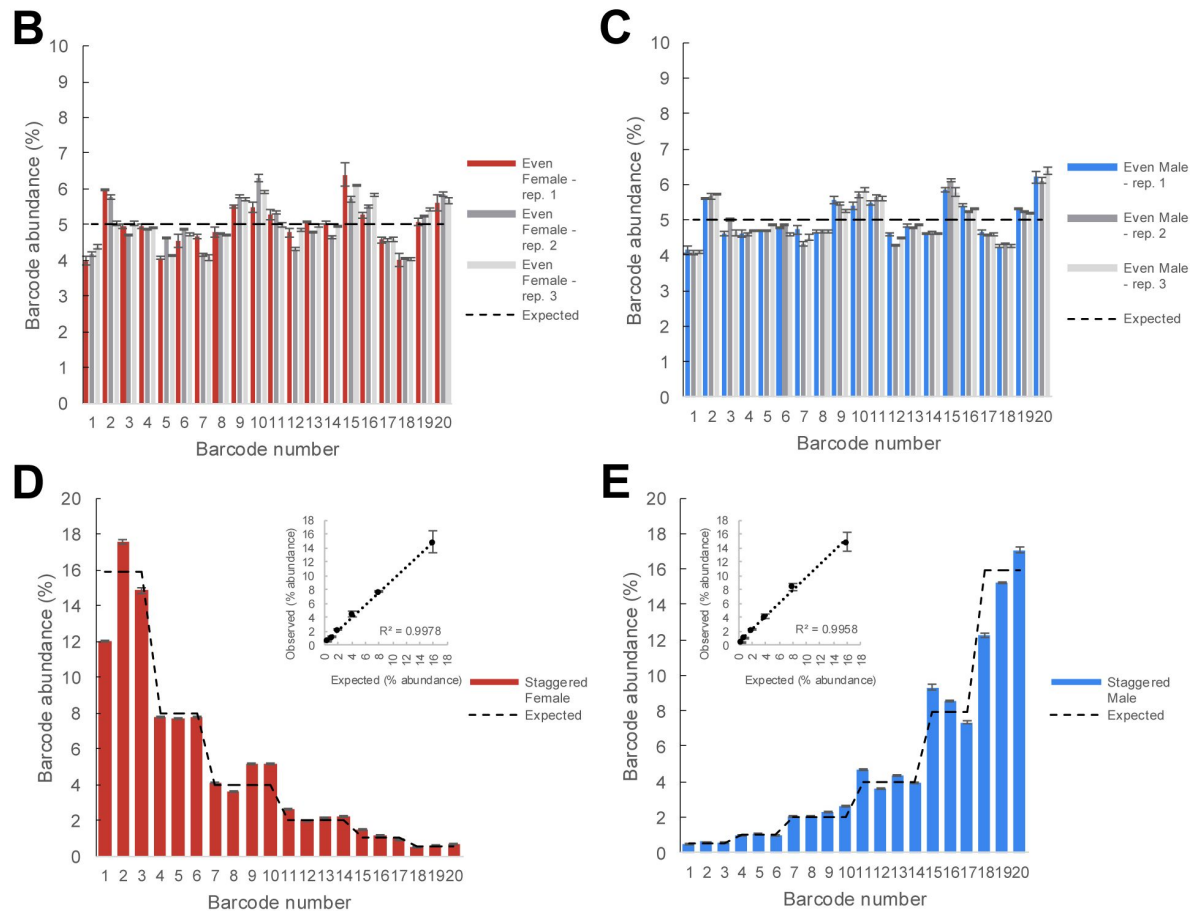


Figure 2.

Structured pool tests.

A) Overview of the construction of the structured pools for assessing the quantitative accuracy of TaG-EM barcode measurements. Male and female even pools were constructed and extracted in triplicate. The table shows the number of flies that were pooled for each experimental condition. B) Barcode abundance data for three independent replicates of the female even pool. C) Barcode abundance data for three independent replicates of the male even pool. D) Barcode abundance data for the female staggered pool. Inset plot shows the average observed barcode abundance among lines pooled at each level compared to the expected abundance. E) Barcode abundance data for the male staggered pool. Inset plot shows the average observed barcode abundance among lines pooled at each level compared to the expected abundance. For all plots, bars indicate the mean barcode abundance for three technical replicates of each pool, error bars are $\pm$ S.E.M.

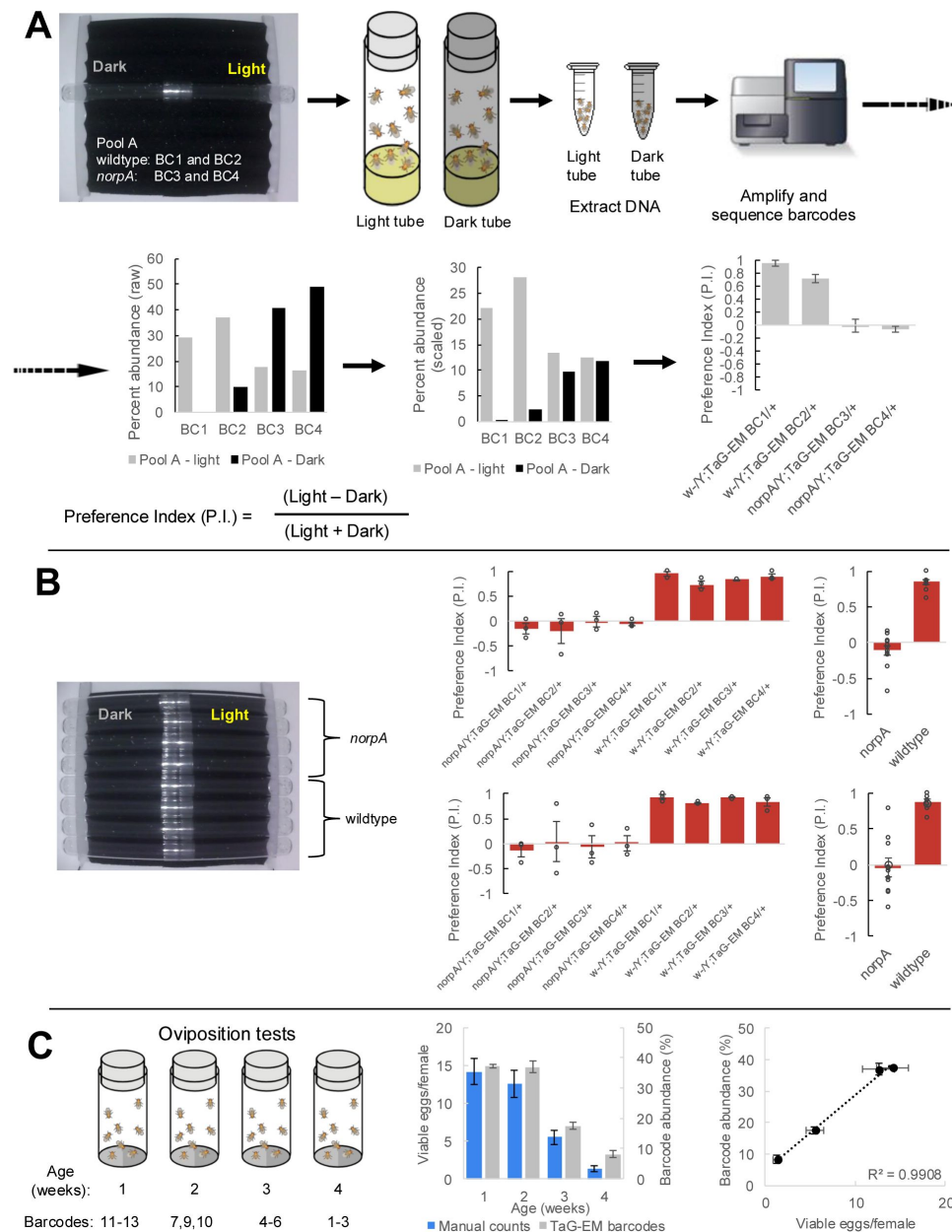


Figure 3.

TaG-EM barcode-based behavioral measurements.

A) TaG-EM barcode lines in either a wild-type or *norpA* background were pooled and tested in a phototaxis assay. After 30 seconds of light exposure, flies in tubes facing the light or dark side of the chamber were collected, DNA was extracted, and TaG-EM barcodes were amplified and sequenced. Barcode abundance values were scaled to the number of flies in each tube and used to calculate a preference index (P.I.). Average P.I. values for four different TaG-EM barcode lines in both the wild-type and *norpA* backgrounds are shown ($n=3$ biological replicates, error bars are $\pm$ S.E.M.). B) The same eight lines used for the sequencing-based TaG-EM barcode measurements were independently tested in the phototaxis assay and manually scored videos were used to calculate a P.I. for each genotype. Average P.I. values for each line are shown ($n=3$ biological replicates, error bars are $\pm$ S.E.M.) for TaG-EM-based quantification (top) and manual video-based quantification (bottom). C) Flies carrying different TaG-EM barcodes were collected and aged for one to four weeks and then eggs were collected and egg number and viability was manually scored for each line. In parallel the barcoded flies from each timepoint were pooled, and eggs were collected, aged, and DNA was extracted, followed by TaG-EM barcode amplification and sequencing. Average number of viable eggs per female (manual counts) and average barcode abundance are shown both as a bar plot and scatter plot ($n=3$ biological replicates for 3 barcodes per condition, error bars are $\pm$ S.E.M.).

0.61 for Trial 1 and 0.74-0.84 for Trial 2). When barcodes from each individual timepoint were averaged, R^2 values for the correlation between manual and sequencing-based measurements were 0.95 for Trial 1 and 0.99 for Trial 2 (**Figure 3C** [↗](#), **Supplemental Figure 6** [↗](#)).

Quantifying food transit time in the larval gut using TaG-EM

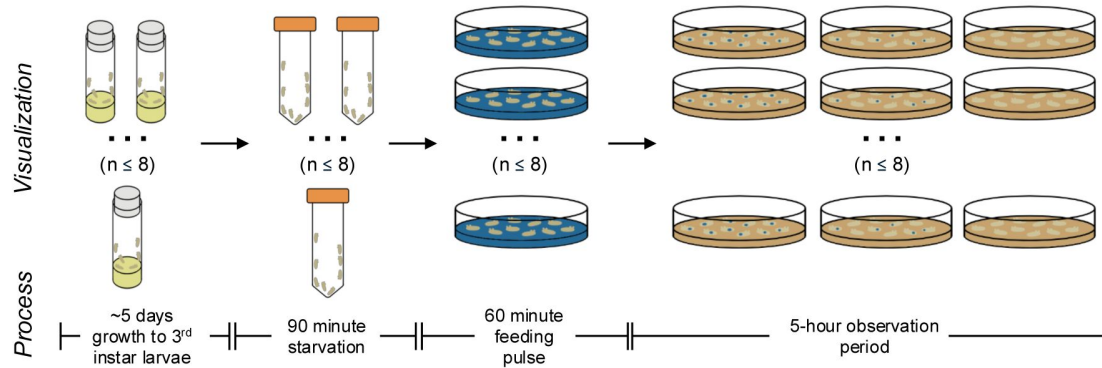
Gut motility defects underlie a number of functional gastrointestinal disorders in humans (Keller et al., 2018 [↗](#)). To study gut motility in *Drosophila*, we have developed an assay based on the time it takes a food bolus to transit the larval gut (**Figure 4A** [↗](#)), similar to approaches that have been employed for studying the role of the microbiome in human gut motility (Asnicar et al., 2021 [↗](#)). Third instar larvae were starved for 90 minutes and then fed food containing a blue dye. After 60 minutes, larvae in which a blue bolus of food was visible were transferred to plates containing non-dyed food, and food transit (indicated by loss of the blue food bolus) was scored every 30 minutes for five hours (**Supplemental Figure 7** [↗](#)).

Because this assay is highly labor-intensive and requires hands-on effort for the entire five-hour observation period, there is a limit on how many conditions or replicates can be scored in one session (~8 plates maximum). Thus, we decided to test whether food transit could be quantified in a more streamlined and scalable fashion by using TaG-EM (**Figure 4B** [↗](#)). Using the manual assay, we observed that while caffeine-containing food is aversive to larvae, the presence of caffeine reduces transit time through the gut (**Figure 4C** [↗](#), **Supplemental Figure 7** [↗](#)). This is consistent with previous observations in adult flies that bitter compounds (including caffeine) activate enteric neurons via serotonin-mediated signaling and promote gut motility (Yao and Scott, 2022 [↗](#)). We tested whether TaG-EM could be used to measure the effect of caffeine on food transit time in larvae. As with prior behavioral tests, the TaG-EM data recapitulated the results seen in the manual assay (**Figure 4D** [↗](#)). Conducting the transit assay via TaG-EM enables several labor-saving steps. First, rather than counting the number of larvae with and without a food bolus at each time point, one simply needs to transfer non-bolus-containing larvae to a collection tube. Second, because the TaG-EM lines are genetically barcoded, all the conditions can be tested at once on a single plate, removing the need to separately count each replicate of each experimental condition. This reduces the hands-on time for the assay to just a few minutes per hour. A summary of the anticipated cost and labor savings for the TaG-EM-based food transit assay is shown in **Supplemental Figure 8** [↗](#).

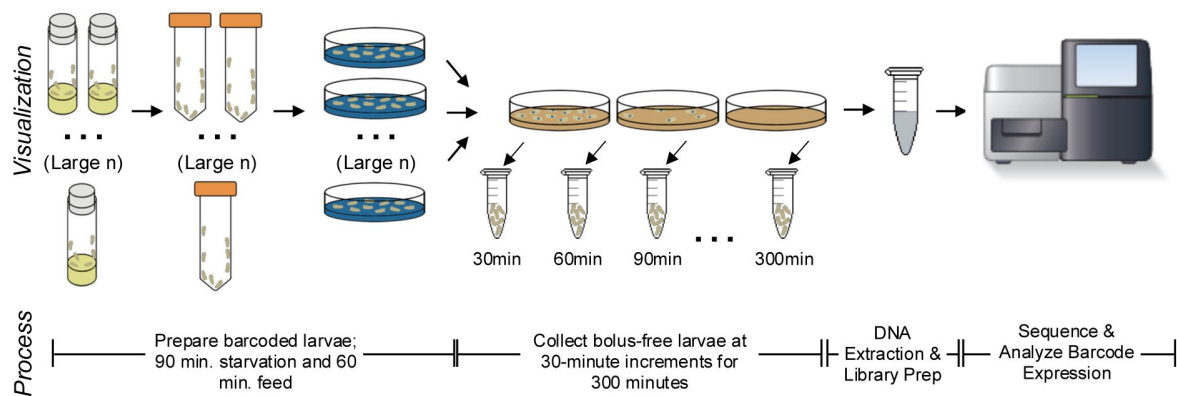
Tissue-specific expression of TaG-EM GFP constructs

To facilitate representation of the TaG-EM barcodes in single-cell sequencing data, genetic barcodes were placed just upstream of the polyadenylation signal sequences and polyA cleavage sites (**Figure 1A** [↗](#)). To verify that the inserted sequences did not interfere with Gal4-driven GFP expression, we crossed each of the barcoded TaG-EM lines to *decapentaplegic-Gal4* (*dpp-Gal4*). We observed GFP expression in the expected characteristic central stripe (Teleman and Cohen, 2000 [↗](#)) in the wing imaginal disc for 19/20 lines at similar expression levels to the base pJFRC12 *UAS-myr::GFP* construct inserted in the same landing site (**Figure 5A** [↗](#), **Supplemental Figure 9** [↗](#)). No GFP expression was visible for TaG-EM barcode number 8, which upon molecular characterization had an 853 bp deletion within the GFP coding region (data not shown). We generated and tested GFP expression of an additional 156 TaG-EM barcode lines (Alegria et al., 2024 [↗](#)), by crossing them to Mhc-Gal4 and observing expression in the adult thorax. All 156 additional TaG-EM lines had robust GFP expression (data not shown). Gal4-driven expression levels of TaG-EM barcoded GFP constructs were also similar to that of the pJFRC12 base construct for multiple driver lines (**Supplemental Figure 10** [↗](#)) indicating that the presence of the barcode does not generally impair expression of GFP.

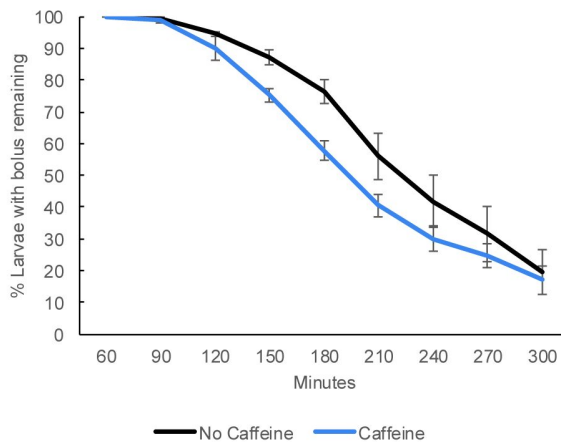
A Manual Gut Motility Assay



B TaG-EM Gut Motility Assay



C Manual Gut Motility Data



D TaG-EM Gut Motility Data

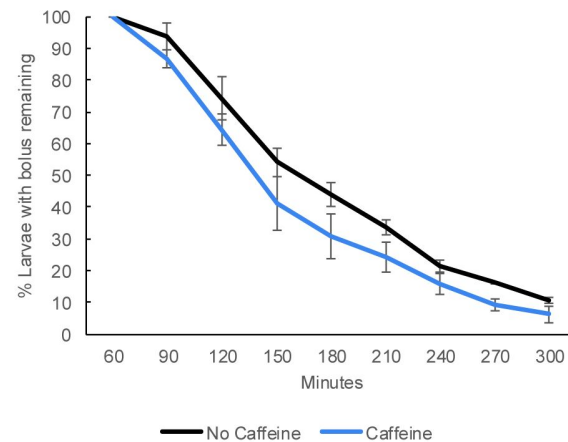


Figure 4.

TaG-EM barcode-based quantification of larval gut motility.

Schematics depicting A) manual and B) TaG-EM-based assays for quantifying food transit time in *Drosophila* larvae. C) Transit time of a food bolus in the presence and absence of caffeine measured using the manual assay ($p = 0.0340$). D) Transit time of a food bolus in the presence and absence of caffeine measured using the TaG-EM assay ($p = 0.0488$). A modified Chi-squared method was used for statistical testing (Hristova and Wimley, 2023 [DOI](#)).

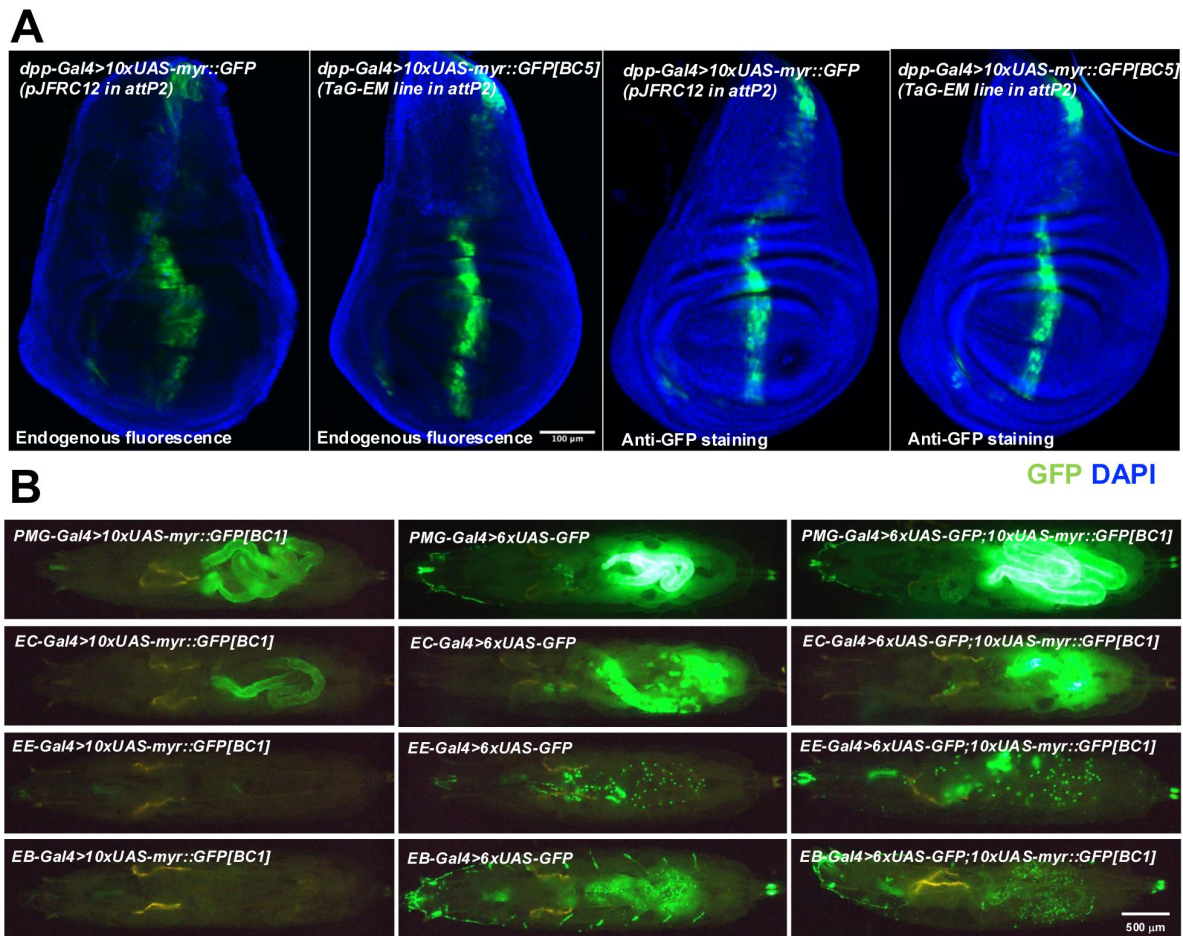


Figure 5.

Gal4 driven expression of GFP from TaG-EM lines.

A) Comparison of endogenous GFP expression and GFP antibody staining in the wing imaginal disc for the original pJFRC12 construct inserted in the attP2 landing site or for a TaG-EM barcode line driven by *dpp-Gal4*. Wing discs are counterstained with DAPI. B) Endogenous expression of GFP from either a TaG-EM barcode construct (left column), a hexameric GFP construct (middle column), or a line carrying both a TaG-EM barcode construct and a hexameric GFP construct (right column) driven by the indicated gut driver line (PMG-Gal4: Pan-midgut driver; EC-Gal4: Enterocyte driver; EE-Gal4: Enteroendocrine driver; EB-Gal4: Enteroblast driver).

Boosting the GFP signal of TaG-EM constructs to enable robust cell sorting

While with some driver lines, expression of the myr::GFP from the TaG-EM construct may be too weak to allow robust enrichment of the tagged cells, adding an additional hexameric GFP construct (Shearin et al., 2014 [\[link\]](#)) could boost expression of weak driver lines to levels that are sufficient for robust detection of labeled flies or larvae (**Figure 5B** [\[link\]](#)) and for labeling of dissociated cells for flow cytometry (**Supplemental Figure 11** [\[link\]](#)). Stocks with an additional UAS hexameric GFP construct recombined onto the same chromosome as the TaG-EM construct have been established for 20 TaG-EM barcode lines.

Correlation between expression of TaG-EM barcodes and intestinal cell marker genes in single-cell sequencing data

To test whether we could detect TaG-EM barcodes in single-cell sequencing data, we crossed three TaG-EM barcode lines to two different gut Gal4 driver lines (Ariyapala et al., 2020 [\[link\]](#)), one expressing in the enterocytes (EC-Gal4: TaG-EM barcodes 1, 2, and 3) and the other in intestinal precursor cells (PC-Gal4: TaG-EM barcodes 7, 8, and 9), which includes stem cells and enteroblasts (EBs). Due to weak GFP expression with the EC-Gal4 driver, we did not see visible GFP positive cells for this driver line. The PC-Gal4 driver line contained an additional UAS-Stinger (2xGFP) construct and expressed GFP at a level sufficient for flow sorting when crossed to the TaG-EM line (**Supplemental Figure 12** [\[link\]](#)). Larval guts were dissected, dissociated, stained with propidium iodide (PI) to label dead cells, and flow sorted to recover PI negative and GFP positive cells. Approximately 10,000 cells were loaded into a 10x Genomics droplet generator and a single-cell library was prepared and sequenced. Two clusters were observed in the resulting sequencing data, one of which had high read counts from mitochondrial genes suggesting that this cluster consisted of mitochondria, debris, or dead and dying cells. After filtering the cells with high mitochondrial reads, a single cluster remained (**Supplemental Figure 13** [\[link\]](#)). This cluster expressed known intestinal precursor cell markers such as *escargot* (*esg*), *klumpfuss* (*klu*), and *Notch* pathway genes like *E(spl)mbeta-HLH* (**Supplemental Figure 13** [\[link\]](#)). Expression of all three PC-Gal4 driven TaG-EM barcodes was observed in this cluster (**Supplemental Figure 13** [\[link\]](#)) indicating that TaG-EM barcodes can be detected in single-cell sequencing data. Interestingly, TaG-EM barcode 8, for which no GFP expression was observed, was represented in the single-cell sequencing data indicating that the lesion in the GFP coding region does not prevent mRNA expression for this line.

A previous study used droplet-based single-cell sequencing to characterize the cell types that make up the adult midgut (Hung et al., 2020 [\[link\]](#)). This study took advantage of two fluorescent protein markers, an *escargot* (*esg*)-GFP fusion protein and a *prospero* (*pros*)-Gal4 driven RFP to label the intestinal stem cells (ISCs) and enteroendocrine cells (EEs), respectively (Hung et al., 2020 [\[link\]](#)). The authors compared the resulting clusters to a list of known marker genes in the literature, including antibody staining, GFP, LacZ, and Gal4 reporter expression patterns to classify the cells in individual clusters, and also found that the *esg*-GFP expression was present in a broader subset of cells than anticipated. Thus, most of these cell classifications relied upon inference as opposed to direct positive labeling. Recently, a large collection of split-Gal4 lines were screened for expression in the adult and larval gut (Ariyapala et al., 2020 [\[link\]](#)). These include pan-midgut driver lines, split-Gal4 lines specific for the EBs, ECs, EEs, and ISC/EBs, as well as driver lines with regionalized gene expression. We crossed four different TaG-EM barcode lines with the pan-midgut driver (PMG-Gal4: TaG-EM barcodes 1, 2, 3, and 7), and one barcode line to each of the precursor cell (PC-Gal4: TaG-EM barcode 5), enterocyte (EC-Gal4: TaG-EM barcode 4), enteroblast (EB-Gal4: TaG-EM barcode 6) and enteroendocrine (EE-Gal4: TaG-EM barcode 9) drivers (Ariyapala et al., 2020 [\[link\]](#)). Larval guts were dissected from these lines and cells were dissociated, flow sorted

as described above to select live, GFP positive cells, and approximately 30,000 cells were loaded into a 10x Genomics droplet generator for single-cell sequencing (**Supplemental Figure 14** [↗](#)). Using the additional hexameric GFP construct to boost GFP expression resulted in visible fluorescent signal for all eight barcode Gal4 line combinations.

An advantage of cell barcoding both for cell hashing (Stoeckius et al., 2018 [↗](#)) and for TaG-EM *in vivo* barcoding is that such labeling facilitates the identification and removal of multiplets, which are an artifact of droplet-based single-cell sequencing approaches. After filtering and removing cells with a high percentage of mitochondrial or ribosomal reads, we used DoubletFinder (McGinnis et al., 2019 [↗](#)) to computationally identify multiplet droplets. In parallel, we searched for cells that co-expressed multiple TaG-EM barcodes. DoubletFinder identified 2,019 multiplet droplets, while TaG-EM barcodes identified 298 such droplets, 198 of which (66.4%) overlapped with those identified by DoubletFinder (**Supplemental Figure 15** [↗](#)). Thus, TaG-EM help identify an additional 100 doublets that would have otherwise been overlooked using computational doublet identification methods.

After doublet removal, the remaining cells were clustered (**Figure 6A** [↗](#), **Supplemental Figure 16** [↗](#)) and analyzed using Seurat (Satija et al., 2015 [↗](#)). Analysis of differentially expressed genes identified clusters expressing marker genes previously reported for adult gut cell types (Hung et al., 2020 [↗](#)). These included genes associated with precursor cells (*Notch* pathway genes), enterocytes (trypsins, serine proteases, amylase, mannosidases), and enteroendocrine cells (neuropeptides and neuropeptide receptors) (**Supplemental Figure 17** [↗](#), data not shown). TaG-EM barcodes derived from the eight multiplexed genotypes were observed in approximately one-quarter of the cells (**Supplemental Figure 17** [↗](#)).

In antibody-conjugated oligo cell hashing approaches, sparsity of barcode representation is overcome by spiking in an additional primer at the cDNA amplification step and amplifying the hashtag oligo by PCR. We employed a similar approach to attempt to enrich for TaG-EM barcodes in an additional library sequenced separately from the 10x Genomics gene expression library. Our initial attempts at barcode enrichment using spike-in and enrichment primers corresponding to the TaG-EM PCR handle were unsuccessful (**Supplemental Figure 18** [↗](#)). However, we subsequently optimized the TaG-EM barcode enrichment by 1) using a longer spike-in primer that more closely matches the annealing temperature used during the 10x Genomics cDNA creation step, and 2) using a nested PCR approach to amplify the cell-barcode and unique molecular identifier (UMI)-labeled TaG-EM barcodes (**Supplemental Figure 18** [↗](#)).

Using the enriched library, TaG-EM barcodes were detected in nearly 100% of the cells at high sequencing depths (**Supplemental Figure 19** [↗](#)). However, although we used a polymerase that has been engineered to have high processivity and that has been shown to reduce the formation of chimeric reads in other contexts (Gohl et al., 2016 [↗](#)), it is possible that PCR chimeras could lead to unreliable detection events for some cells. Indeed, many cells had a mixture of barcodes detected with low counts and single or low numbers of associated UMIs. To assess the reliability of detection, we analyzed the correlation between barcodes detected in the gene expression library and the enriched TaG-EM barcode library as a function of the purity of TaG-EM barcode detection for each cell (the percentage of the most abundant detected TaG-EM barcode, **Supplemental Figure 19** [↗](#)). For TaG-EM barcode detections where the most abundance barcode was a high percentage of the total barcode reads detected (~75%-99.99%), there was a high correlation between the barcode detected in the gene expression library and the enriched TaG-EM barcode library. Below this threshold, the correlation was substantially reduced.

In the enriched library, we identified 26.8% of cells with a TaG-EM barcode reliably detected, a very modest improvement over the gene expression library alone (23.96%), indicating that at least for this experiment, the main constraint is sufficient expression of the TaG-EM barcode and not detection. To identify TaG-EM barcodes in the combined data set, we counted a positive detection

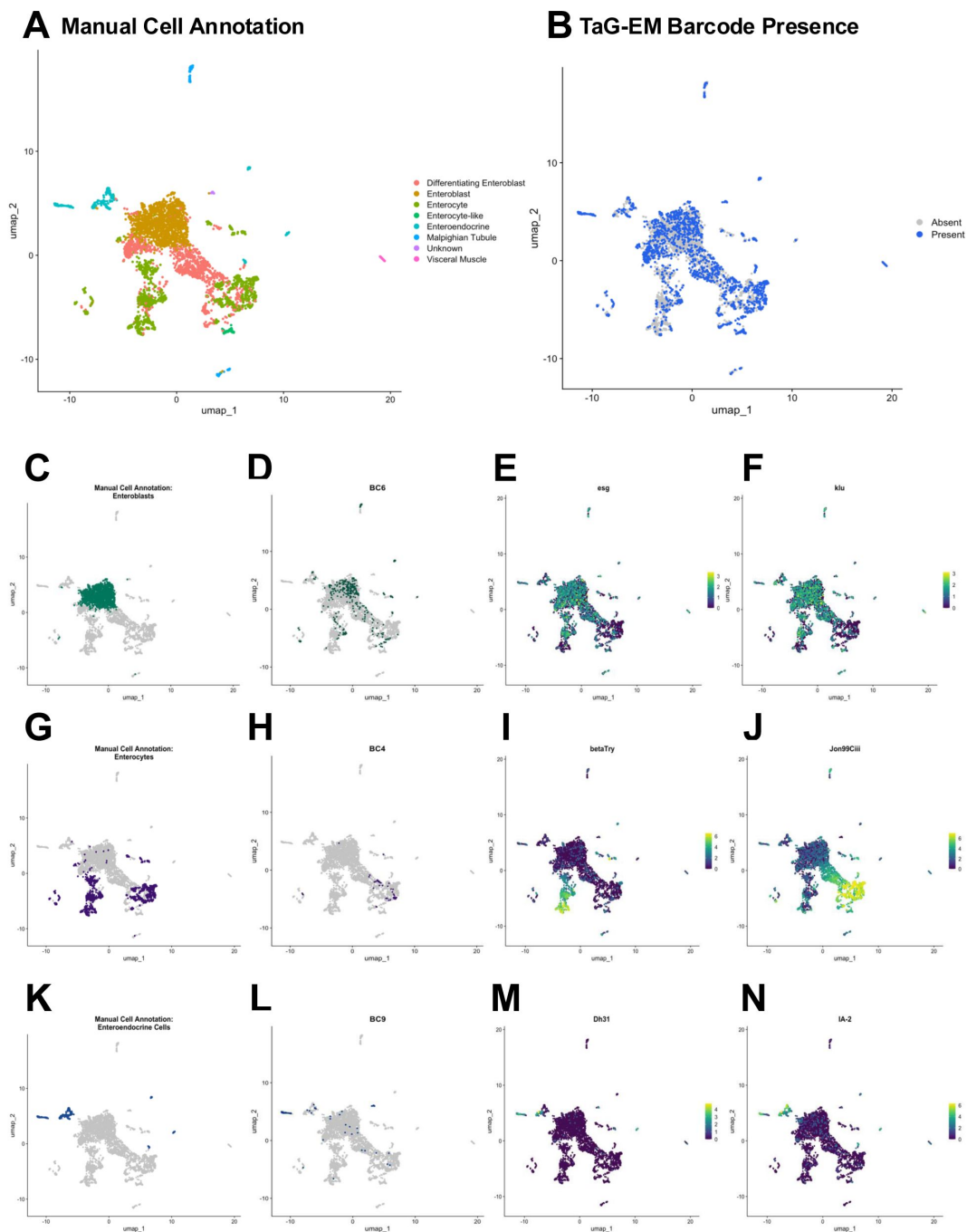


Figure 6.

Expression of TaG-EM genetic barcodes in larval intestinal cell types.

A) UMAP plot of *Drosophila* larval gut cell types. B) Annotation of cells associated with a TaG-EM barcode across all 8 multiplexed experimental conditions using data from the gene expression library and an enriched TaG-EM barcode library. C) Annotated enteroblast cells. D) Presence of TaG-EM barcode (BC6) driven by the EB-Gal4 line using data from the gene expression library and an enriched TaG-EM barcode library. Gene expression levels of enteroblast marker genes E) *esg*, F) *klu*. G) Annotated enterocyte cells. H) Presence of TaG-EM barcode (BC4) driven by the EC-Gal4 line using data from the gene expression library and an enriched TaG-EM barcode library. Gene expression levels of enterocyte marker genes I) *betaTry*, J) *Jon99Ciii*. K) Annotated enteroendocrine cells. L) Presence of TaG-EM barcode (BC9) driven by the EE-Gal4 line using data from the gene expression library and an enriched TaG-EM barcode library. Gene expression levels of enteroendocrine cell marker genes M) *Dh31*, N) *IA-2*.

as any barcode either identified in the gene expression library or any barcode identified in the enriched library with a purity of >75%. In the case of conflicting barcode calls, we assigned the barcode that was detected directly in the gene expression library. This increased the total fraction of cells where a barcode was identified to approximately 37% (**Figure 6B** [↗](#)).

As expected, the barcodes expressed by the pan-midgut driver were broadly distributed across the cell clusters (**Supplemental Figure 20** [↗](#)). However, the number of cells recovered varied significantly among the four pan-midgut driver associated barcodes. Expression of the cell type-specific barcodes showed more restricted patterns of expression among the cell clusters and were co-localized with known marker genes for these cell types (**Figure 6C-N** [↗](#)). For instance, TaG-EM barcode 6, driven by the EB-Gal4 line, was expressed primarily in cells that were annotated as enteroblasts (**Figure 6C-D** [↗](#)) and that expressed precursor cell markers such as *esg* (**Figure 6E** [↗](#)), *klu* (**Figure 6F** [↗](#)), and Notch pathway genes such as *e(spl)mbeta-HLH* and *e(spl)m3-HLH* (not shown).

TaG-EM barcode 4 expression, which was driven by the EC-Gal4 line, was seen primarily in a cluster that was annotated as enterocytes (**Figure 6G-H** [↗](#)) and that expressed enterocyte markers such as the serine protease *Jon99Ciii* and other enterocyte marker genes such as the amylase, *Amy-d* (not shown), but not the *beta-Trypsin* (*betaTry*) gene (**Figure 6I-J** [↗](#)). Detailed characterization of the EC-Gal4 line indicated that although this line labeled a high percentage of enterocytes, expression was restricted to an area at the anterior and middle of the midgut, with gaps between these regions and at the posterior (**Supplemental Figure 21** [↗](#)). This could explain the absence of subsets of enterocytes, such as those labeled by *betaTry*, which exhibits regional expression in R2 of the adult midgut (Buchon et al., 2013 [↗](#)).

Finally, expression of TaG-EM barcode 9, which was expressed using the EE-Gal4 driver line, was observed in clusters annotated as enteroendocrine cells (**Figure 6K-L** [↗](#)) and that also expressed enteroendocrine cell derived neuropeptide genes such as *Dh31* (**Figure 6M** [↗](#)) and other enteroendocrine markers such as *IA-2*, a tyrosine phosphatase involved in the secretion of insulin-like peptide (**Figure 6N** [↗](#)). Detailed characterization of the EE-Gal4 driver line indicated that ~80-85% of Prospero-positive enteroendocrine cells are labeled in the anterior and middle of the larval midgut, with a lower percentage (~65%) of Prospero-positive cells labeled in the posterior midgut (**Supplemental Figure 21** [↗](#)). As with the enterocyte labeling, and consistent with the Gal4 driver expression pattern, the EE-Gal4 expressed TaG-EM barcode 9 did not label all classes of enteroendocrine cells and other clusters of presumptive enteroendocrine cells expressing other neuropeptides such as *Orcokinin*, *AstA*, and *AstC*, or neuropeptide receptors such as *CCHa2* (not shown) were also observed. The EE-Gal4 driver uses *Dh31* regulatory elements, so it is not surprising that the TaG-EM barcodes specifically labeled *Dh31*-positive enteroendocrine cells and this result further highlights the ability to target specific genetically defined cell types using TaG-EM based on *in vivo* cell labeling. Taken together, these results demonstrate that TaG-EM can be used to label specific cell populations that correlate with Gal4-driven expression patterns *in vivo* for subsequent identification in single-cell sequencing data.

Discussion

Advances in next-generation sequencing, as well as single-cell and spatial genomics are enabling new types of detailed analyses to study important biological processes such as development and nervous system function. Here we describe TaG-EM, a genetic barcoding strategy that enables novel capabilities in several different experimental contexts (**Figure 1** [↗](#)).

We demonstrate that the genetic barcodes can be quantified from mixtures of barcoded fly lines using next-generation sequencing. Analysis of structured pools of flies with defined inputs suggests that TaG-EM barcode measurements are highly accurate and reproducible, particularly in

cases where multiple barcodes are used to label an experimental condition and averaged (**Figure 2**). Sequencing-based TaG-EM measurements recapitulated more laborious, one-at-a-time measurements in a phototaxis assay, an age-dependent fecundity assay, and a gut motility assay, demonstrating that TaG-EM can be used to measure behavior or other phenotypes in multiplexed, pooled populations (**Figures 3-4**). We did note that one line (TaG-EM barcode 14) exhibited poor performance in oviposition assays, suggesting that barcode performance should be verified for a specific assay. We excluded this poor performing barcode line from the fecundity tests, however, backcrossing is often used to bring reagents into a consistent genetic background for behavioral experiments and could also potentially be used to address behavior-specific issues with specific TaG-EM lines. In addition, other strategies such as averaging across multiple barcode lines (3-4 per condition, which yielded R^2 values >0.99 in tests with structured pools) or permutation of barcode assignment across replicates could also mitigate such deficiencies. Currently, up to 176 conditions can be multiplexed in a single pooled experiment with existing TaG-EM lines, but because sequencing indices can be added after amplification in a separate indexing PCR step, many hundreds or even thousands of such experiments can be multiplexed in a single sequencing run. While the utility of TaG-EM barcode-based quantification will vary based on the number of conditions being analyzed and the ease of quantifying the behavior or phenotype by other means, we demonstrate that TaG-EM can be employed to cost-effectively streamline labor-intensive assays and to quantify phenotypes with small effect sizes (**Figure 4**, **Supplemental Figure 8**). An additional benefit of multiplexed TaG-EM behavioral measurements is that the experimental conditions are effectively blinded as the multiplexed conditions are intermingled in a single assay.

In addition, we show that TaG-EM barcodes can be expressed by tissue-specific Gal4 drivers and used to tag specific cell populations upstream of single-cell sequencing (**Figures 5-6**). This capability will allow for positive identification of cell clusters in cell atlas projects and will facilitate multiplexing of single-cell sequencing experiments. Recently, a conceptually similar approach called RABID-Seq (Clark et al., 2021) has been described, which allows trans-synaptic labeling of neural circuits using barcoded viral transcripts. However, one distinction between the two approaches is that RABID-Seq relies on stochastic viral infection of mammalian cells while TaG-EM allows reproducible targeting of defined cell populations allowing unambiguous cell identification and potentially allowing the same cell populations to be assessed at different timepoints or in the context of different experimental manipulations. One current limitation is that TaG-EM barcodes are not observed in all cells in single-cell gene expression data. It is likely that the strength of the Gal4 driver contributes to the labeling density. However, we also observed variable recovery of TaG-EM barcodes that were all driven by the same pan-midgut Gal4 driver (**Supplemental Figure 20**). For single cell RNA-Seq experiments, the cost savings of multiplexing is roughly the cost of a run divided by the number of independent lines multiplexed, plus labor savings by also being able to multiplex upstream flow cytometry, minus loss of unbarcoded cells. Our experiments indicated that for the specific drivers we tested TaG-EM barcodes are detected in around one quarter of the cells if relying on endogenous expression in the gene expression library, though this fraction was higher ($\sim 37\%$) if sequencing an enriched TaG-EM barcode library in parallel (**Figure 6**, **Supplemental Figures 18-19**).

In the future, generation of additional TaG-EM lines will enable higher levels of multiplexing. In addition, while the original TaG-EM lines were made using a membrane-localized myr::GFP construct, variants that express GFP in other cell compartments such as the cytoplasm or nucleus could be constructed to enable increased expression levels or purification of nuclei. Nuclear labeling could also be achieved by co-expressing a nuclear GFP construct with existing TaG-EM lines in analogy to the use of hexameric GFP described above.

In summary, combined with the large collections of Gal4 and split-Gal4 lines that have been established in *Drosophila* that enable precise targeting of a high proportion of cell types (Ariyapala et al., 2020; Aso et al., 2014; Davis et al., 2020; Gohl et al., 2011; Kanca et al., 2022; Namiki et al., 2018; Pfeiffer et al., 2010, 2008; Venken et al., 2011; Zirin et al., 2024),

TaG-EM provides a means to target and label cells *in vivo* for subsequent detection in single-cell sequencing. Moreover, these genetic barcodes can be used to multiplex behavioral or other phenotypic measurements. Thus, TaG-EM provides a flexible system for barcoding cells and organisms.

Methods

Drosophila stocks and maintenance

Drosophila stocks were grown at 22°C on cornmeal agar unless otherwise indicated. The following stocks were used in this study:

Stock name	Source
IsoD1	Clandinin Lab, Stanford University
<i>w</i> ⁻ ; +; + (IsoD1)	Clandinin Lab, Stanford University
<i>attP2</i> line	BDSC #25710
<i>norpA</i>	BDSC# 9048
UAS-myr::GFP (pJFRC12)	BDSC #32197
Hexameric GFP lines	BDSC #91402, #91403
UAS-6XmCherry-HA	BDSC #52268
UAS-GFP.nls	BDSC #4776
esg-GFP.FPTB	BDSC #83386
dpp-Gal4 driver	BDSC #1553
Act-Gal4 driver	BDSC #4414
Tub-Gal4 driver	BDSC #5138
Mhc-Gal4 driver	BDSC #55132
PC-Gal4 driver	BDSC #73356 + #75528
PC-Gal4 driver (with UAS-Stinger)	BDSC #91400 + #91401
EC-Gal4 driver	BDSC #91406
EB-Gal4 driver	BDSC #91398 + #91404
EE-Gal4 driver	BDSC #91402 + #91403
PMG-Gal4 driver	BDSC #91398 + #91399
TaG-EM lines (x176)	This study, (Alegria et al., 2024)
TaG-EM lines + 6xGFP (x20)	This study

Design and cloning of TaG-EM constructs

A gBlock with the following sequence containing a part of the SV40 3' UTR with a PCR handle (uppercase, below) and a 14 bp randomer sequence just upstream of the SV40 polyadenylation site (bold and underlined, below) was synthesized (Integrated DNA Technologies, IDT):

```
caaagaaaaagctgcactgctatacaagaaaattatggaaaaatattgatgtatagtccttgactagagatcataatcagccata
ccacattgtagaggttttactgtcttaaaaaacccccacacccccctgaacctgaacataaaatgaatgcaattgtgtgttaact
tggtttatgcagcttataaCTCCAACAACCGGAAGTGANNNNNNNNNNNNNNNtggttacaataaagcaatag
catcacaatttcacaataaagcatgtttttcactgcattctagtgtgtgtgttgcctcaactcatcaatgtatcttatcatgtctgcatctc
ggccggccgtttaaacgaattctgaagacgaaaggcctcgtgatacgctattttatagtgtaatgcatgataataatg
```

The gBlock was resuspended in 20 µl EB, incubated at 50°C for 20 minutes and then cut with *PsiI* and *EcoRI* (New England Biolabs, NEB) using the following reaction conditions: 4 µl gBlock DNA (35 ng), 2 µl 10x CustSmart buffer (NEB), 1 µl *EcoRI* enzyme (NEB), 1 µl *PsiI* enzyme (NEB), and 12 µl nuclease-free water were mixed and incubated at 37°C for 1 hour followed by 65°C for 20 minutes to heat inactivate the restriction enzymes. pJFRC12-10XUAS-IVS-myr::GFP plasmid (Addgene, Plasmid #26222) (Pfeiffer et al., 2010 [DOI](#)) was digested with the following reaction conditions: 5 µl pJFRC12-10XUAS-IVS-myr::GFP plasmid DNA (~3 µg), 5 µl 10x CutSmart buffer (NEB), 1 µl *PsiI* enzyme (NEB), 1 µl *EcoRI* enzyme (NEB), and 38 µl nuclease-free water, were mixed and incubated at 37°C for 1 hour, followed by addition of 1 µl of CIP and incubation for an additional 30 minutes. The digested vector backbone was gel purified using the QiaQuick Gel Purification Kit (Qiagen). The digested gBlock was ligated into the digested pJFRC12-10XUAS-IVS-myr::GFP backbone using the following reactions conditions: 4 µl T4 ligase buffer (10x) (NEB), 20 µl plasmid backbone DNA (0.005 pmol), 5 µl gBlock digest DNA (0.03 pmol), 2 µl of T4 DNA ligase (NEB), and 9 µl nuclease-free water were mixed and incubated at 22°C for 2 hours. 2 µl of the ligation reaction was transformed into 50 µl of TOP10 competent cells (Invitrogen), and the cells were incubated on ice for 30 minutes, then heat shocked at 42°C for 30 seconds, and incubated on ice for 5 minutes. 250 µl SOC was added and the cells were plated on LB+ampicillin plates and incubated overnight at 37°C. DNA was isolated from 36 pJFRC12-gBlock colonies using a QIAprep Spin MiniPrep kit (Qiagen). Expected construct size was verified by diagnostic digest with *EcoRI* and *ApaLI*. DNA concentration was determined using a Quant-iT PicoGreen dsDNA assay (Thermo Fisher Scientific) and the randomer barcode for each of the constructs was determined by Sanger sequencing using the following primers: SV40_post_R: GCCAGATCGATCCAGACATGA SV40_5F: CTCCCCCTGAACCTGAAACA

Generation of TaG-EM transgenic lines

29 sequence verified constructs were normalized, pooled evenly, and injected as a pool into embryos (Rubin and Spradling, 1982 [DOI](#)) expressing PhiC31 integrase and the carrying the *attP2* landing site (BDSC #25710). Injected flies were outcrossed to w-flies, and up to three *white+* progeny per cross were identified, and the transgenic lines were homozygosed. DNA was extracted (GeneJET genomic DNA purification Kit, Thermo Scientific) and the region containing the DNA barcode was amplified with the following PCR reaction: 2.5 µl 1:10 diluted template DNA, 2 µl 10x Reaction Buffer (Qiagen), 0.2 µl dNTP mix (10 µM), 1 µl 10 µM SV40_5F primer (10 µM), 1 µl SV40_post_R primer (10 µM), 0.8 µl MgCl₂ (3 mM), 0.1 µl Taq polymerase (Qiagen), 12.4 µl nuclease-free water. Reactions were amplified using the following cycling conditions: 95°C for 5 minutes, followed by 30 cycles of 94°C for 30 seconds, 55°C for 30 seconds, 72°C for 30 seconds, followed by 72°C for 5 minutes. PCR products were treated with Exo-CIP using the following reaction conditions: 5 µl PCR product, 1 µl Exo-CIP Tube A (NEB), 1 µl Exo-CIP Tube B (NEB) were mixed and incubated at 37°C for 4 minutes, followed by 80°C for 1 minute. The barcode sequence for each of the independent transgenic lines was determined by Sanger sequencing using the SV40_5F and

SV40_PostR primers. Transgenic lines containing 20 distinct DNA barcodes were recovered (**Supplemental Figure 1**). An additional 156 TaG-EM barcode lines were isolated and sequence verified as described in a separate publication (Alegria et al., 2024).

Optimizing amplification of TaG-EM barcodes for next-generation sequencing

The following primers were evaluated to amplify the TaG-EM barcodes upstream of NGS: Forward primer pool; four primers with frameshifting bases to increase library sequence diversity in initial sequencing cycles were normalized to 10 μ M and pooled evenly to make a B2_3'F1_Nextera_0-6 primer pool:

B2_3'F1_Nextera: TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCTTCCAACAACCGGAAG*TGA

B2_3'F1_Nextera_2 TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGAGCTTCCAACAACCGGAAG*TGA

B2_3'F1_Nextera_4
TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTCGACTTCCAACAACCGGAAG*TGA

B2_3'F1_Nextera_6
TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGAAGAGCTTCCAACAACCGGAAG*TGA

The following two reverse primers were tested:

SV40_pre_R_Nextera:
GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGATTTGTGAAATTTGTGATGCTATTGC*T T

SV40_post_R_Nextera:
GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGCCAGATCGATCCAGACA*TGA

SV40_pre_R_nextera is designed to produce a shorter amplicon (200 bp with Illumina adapters and indices added) and SV40-post_R_Nextera is designed to produce a longer amplicon (290 bp with Illumina adapters and indices added).

An initial test was performed with three different polymerases (Qiagen Q5, KAPA HiFi, and Qiagen Taq) at two different annealing temperatures and with both the B2_3'F1_Nextera/ SV40_pre_R_Nextera and B2_3'F1_Nextera/ SV40_post_R_Nextera primer sets to determine whether the primers amplify as expected (**Supplemental Figure 1**). Two different samples were tested:

1. Pool of 8 putative transformant samples (pooled 5 μ l each of 1:10 diluted sample)
2. OreR (wild type - diluted 1:10) Set up the following PCR reactions:

Q5 polymerase: 2.5 μ l template DNA, 1 μ l 10 μ M Forward primer (10 μ M), 1 μ l Reverse primer (10 μ M), 10 μ l 2x Q5 Master Mix (Qiagen), 5.5 μ l nuclease-free water. Reactions were amplified using the following cycling conditions: 98°C for 30 seconds, followed by 30 cycles of 98°C for 20 seconds, 55°C or 60°C for 15 seconds, 72°C for 30 seconds, followed by 72°C for 5 minutes.

KAPA HiFi polymerase: 2.5 μ l template DNA, 1 μ l 10 μ M Forward primer (10 μ M), 1 μ l Reverse primer (10 μ M), 10 μ l 2x KAPA HiFi ReadyMix (Roche), 5.5 μ l nuclease-free water. Reactions were amplified using the following cycling conditions: 95°C for 5 minutes, followed by 30 cycles of 98°C for 20 seconds, 55°C or 60°C for 15 seconds, 72°C for 30 seconds, followed by 72°C for 5 minutes.

Taq polymerase: 2.5 µl template DNA, 2 µl 10x Reaction Buffer (Qiagen), 0.2 µl dNTP mix (10 µM), 1 µl 10 µM Forward primer (10 µM), 1 µl Reverse primer (10 µM), 0.8 µl MgCl₂ (3 mM), 0.1 µl Taq polymerase (Qiagen), 12.4 µl nuclease-free water. Reactions were amplified using the following cycling conditions: 95°C for 5 minutes, followed by 30 cycles of 94°C for 30 seconds, 55°C or 60°C for 30 seconds, 72°C for 30 seconds, followed by 72°C for 5 minutes.

Samples were run on a 2% agarose gel to verify amplification products (**Supplemental Figure 1**). Next, the TaG-EM barcode lines were pooled in either an even or staggered manner. To optimize reaction conditions for the barcode measurements, either 5 ng or 50 ng of DNA was amplified in triplicate for each pool for either 20, 25, or 30 cycles with either KAPA HiFi using the B2_Nextera_F 0-6 forward primer pool together with either the SV40_pre_R_Nextera or the SV40_post_R_Nextera reverse primer. Next, PCR reactions were diluted 1:100 in nuclease-free water and amplified in the following indexing reactions: 3 µl PCR 1 (1:100 dilution), 1 µl indexing primer 1 (5 µM), 1 µl indexing primer 2 (5 µM), and 5 µl 2x Q5 master mix. The following indexing primers were used (X indicates the positions of the 8 bp indices):

Forward indexing primer: AATGATACGGCGACCACCGAGATCTACACXXXXXXXXTCGTCGGCAGCGTC

Reverse indexing primer: CAAGCAGAAGACGGCATACGAGATXXXXXXXXGTCTCGTGGGCTCGG

Reactions were amplified using the following cycling conditions: 98°C for 30 seconds, followed by 10 cycles of 98°C for 20 seconds, 55°C for 15 seconds, 72°C for 1 minute, followed by 72°C for 5 minutes. Amplicons were then purified and normalized using a SequalPrep normalization plate (Thermo Fisher Scientific), followed by elution in 20 µl of elution buffer. An even volume of the normalized libraries was pooled and concentrated using 1.8x AmpureXP beads (Beckman Coulter). Pooled libraries were quantified using a Qubit dsDNA high sensitivity assay (Thermo Fisher Scientific) and libraries were normalized to 2 nM for sequencing on the Illumina MiSeq (see below).

Structured fly pool experiments

Male or female flies from TaG-EM barcode lines were pooled in either an even or staggered manner (**Figure 2A**). For the even pools, three independently pooled samples were constructed in order to assess sample-to-sample variability. DNA was extracted from these structured pools using a protocol adapted from Huang *et al.*, 2009 (Huang *et al.*, 2009), using homemade SPRI beads (DeAngelis *et al.*, 1995) in the last purification step and amplified in triplicate using 2.5 µl template DNA (50 ng), 1 µl 10 µM B2_Nextera_F 0-6 primers (10 µM), 1 µl SV40_pre_R_Nextera (10 µM), 10 µl 2x KAPA HiFi ReadyMix (Roche), 5.5 µl nuclease-free water. Reactions were amplified using the following cycling conditions: 98°C for 5 minutes, followed by 30 cycles of 98°C for 20 seconds, 60°C for 15 seconds, 72°C for 30 seconds, followed by 72°C for 5 minutes. Amplicons were indexed, normalized, quantified, and prepared for sequencing as described above.

Phototaxis experiments

Video-based measurements

A pair of white LED strip lights with Muzata LED Strip Light Diffusers (U1SW WW 1M, LU1) were mounted withing a light-tight box and controlled using an Vilros Uno Rev 3 microcontroller. Test tubes containing flies were held in place with Acoustic Foam Panels (1"x12"x12", ALPOWL). Videos and images were acquired using an Arducam 1080P Day & Night Vision USB Camera with an IR filter and using Photo Booth software (Apple). Wild type and *norpA* flies carrying one of 4 different TaG-EM barcodes were tested in three independent experimental replicates. 20 male flies of each genotype were transferred into 25 mm x 150 mm glass test tubes, incubated at 34°C for 10 minutes and then run in the phototaxis assay, where a light at one end of the chamber was turned on for 30

seconds. Videos of all tests were recorded through the end of the 30 second light pulse. Videos were independently scored by two observers to determine the number of flies in the light-facing or dark-facing tubes and the results were averaged. A preference index (P.I.) was calculated using the following formula: $[(\text{number of flies in light tube}) - (\text{number of flies in dark tube})]/(\text{total number of flies})$.

TaG-EM measurements

For TaG-EM barcode-based phototaxis measurements, the following genotypes were consolidated into a single test tube:

Pool A

norpA/Y;TaG-EM BC4/+

norpA/Y;TaG-EM BC3/+

w-/Y;TaG-EM BC2/+

w-/Y;TaG-EM BC1/+

Pool B

norpA/Y;TaG-EM BC2/+

norpA/Y;TaG-EM BC1/+

w-/Y;TaG-EM BC4/+

w-/Y;TaG-EM BC3/+

These pools were individually incubated at 34°C for 10 minutes and then run in the phototaxis assay. Videos of all tests were recorded and at the end of a 30 second light pulse the two test tubes were quickly separated and capped. Flies in each of these tubes were counted, then DNA was extracted from the flies from the light-facing or dark-facing tubes and amplified using 2.5 µl template DNA (50 ng), 1 µl 10 µM B2_Nextera_F 0-6 primers (10 µM), 1 µl SV40_pre_R_Nextera (10 µM), 10 µl 2x KAPA HiFi ReadyMix (Roche), 5.5 µl nuclease-free water. Reactions were amplified using the following cycling conditions: 95°C for 5 minutes, followed by 30 cycles of 98°C for 20 seconds, 60°C for 15 seconds, 72°C for 30 seconds, followed by 72°C for 5 minutes. Amplicons were indexed, normalized, quantified, and prepared for sequencing as described above.

Oviposition experiments

Newly hatched flies (males and females) from three barcode lines were collected at one-week intervals during 4 consecutive weeks (12 barcode lines in total). Fresh fly food was provided every 3-4 days. Ten days after the last collection, 10 females from each barcode line were taken and pooled together in a collection cage (10 females x 12 barcode lines = 120 females). The remaining females from each barcode line were separated from the males and put in individual collection cages. Two days later, the experiment started and run for 3 consecutive days. Each day a 1-1.5 hour pre-collection was followed by a 6 hour collection, both at 25°C. 100 embryos from each individual collection plate were transferred to new plates and incubated for 2 days at 18°C. The number of hatched larvae were counted and used to calculate the egg survival rate. The pooled collection plate was also incubated at 18°C and the next day the embryos were dechorionated and frozen. The 12 individual collection plates were kept at 4°C and the number of embryos counted in the following days. For the barcode measurements, DNA was extracted from the embryos, and

amplified using 2.5 µl template DNA (50 ng), 1 µl 10 µM B2_Nextera_F 0-6 primers (10 µM), 1 µl SV40_pre_R_Nextera (10 µM), 10 µl 2x KAPA HiFi ReadyMix (Roche), 5.5 µl nuclease-free water. Reactions were amplified using the following cycling conditions: 95°C for 5 minutes, followed by 30 cycles of 98°C for 20 seconds, 60°C for 15 seconds, 72°C for 30 seconds, followed by 72°C for 5 minutes. Amplicons were indexed, normalized, quantified, and prepared for sequencing as described above.

Larval gut motility experiments

Preparing Yeast Food Plates

Yeast agar plates were prepared by making a solution containing 20% Red Star Active Dry Yeast 32oz (Red Star Yeast) and 2.4% Agar Powder/Flakes (Fisher) and a separate solution containing 20% Glucose (Sigma-Aldrich). Both mixtures were autoclaved with a 45-minute liquid cycle and then transferred to a water bath at 55°C. After cooling to 55°C, the solutions were combined and mixed, and approximately 5 mL of the combined solution was transferred into 100 x 15 mm petri dishes (VWR) in a PCR hood or contamination-free area. For blue-dyed yeast food plates, 0.4% Blue Food Color (McCormick) was added to the yeast solution. For the caffeine assays, 300 µL of a solution of 100 mM 99% pure caffeine (Sigma-Aldrich) was pipetted onto the blue-dyed yeast plate and allowed to absorb into the food during the 90-minute starvation period.

Manual Gut Motility Assay

Third instar *Drosophila* larvae were transferred to empty conical tubes that had been misted with water to prevent the larvae from drying out. After a 90-minute starvation period the larvae were moved from the conical to a blue-dyed yeast plate with or without caffeine and allowed to feed for 60 minutes. Following the feeding period, the larvae were transferred to an undyed yeast plate. Larvae were scored for the presence or absence of a food bolus every 30 minutes over a 5-hour period. Up to 8 experimental replicates/conditions were scored simultaneously.

TaG-EM Gut Motility Assay

Third instar larvae were starved and fed blue dye-containing food with or without caffeine as described above. An equal number of larvae from each experimental condition/replicate were transferred to an undyed yeast plate. During the 5-hour observation period, larvae were examined every 30 minutes and larvae lacking a food bolus were transferred to a microcentrifuge tube labeled for the timepoint. Any larvae that died during the experiment were placed in a separate microcentrifuge tube and any larvae that failed to pass the food bolus were transferred to a microcentrifuge tube at the end of the experiment. DNA was extracted from the larvae in each tube and TaG-EM barcode libraries were prepared and sequenced as described above.

Dissection and immunostaining

Midguts from third instar larvae of driver lines crossed to UAS-GFP.nls or UAS-mCherry were dissected in 1xPBS and fixed with 4% paraformaldehyde (PFA) overnight at 4°C. Fixed samples were washed with 0.1% PBTx (1xPBS + 0.1% Triton X-100) three times for 10 minutes each and blocked in PBTxGS (0.1% PBTx + 3% Normal Goat Serum) for 2–4 hours at RT. After blocking, midguts were incubated in primary antibody solution overnight at 4°C. The next day samples were washed with 0.1% PBTx three times for 20 minutes each and were incubated in secondary antibody solution for 2–3 hours at RT (protected from light) followed by three washes with 0.1% PBTx for 20 minutes each. One µg/ml DAPI solution prepared in 0.1% PBTx was added to the sample and incubated for 10 minutes followed by washing with 0.1% PBTx three times for 10 minutes each. Finally, samples were mounted on a slide glass with 70% glycerol and imaged using a Nikon AX R confocal microscope. Confocal images were processed using Fiji software.

The primary antibodies used were rabbit anti-GFP (A6455, 1:1000 Invitrogen), mouse anti-mCherry (3A11, 1:20 DSHB), mouse anti-Prospero (MR1A, 1:50 DSHB) and mouse anti-Pdm1 (Nub 2D4, 1:30 DSHB). The secondary antibodies used were goat anti-mouse and goat anti-rabbit IgG conjugated to Alexa 647 and Alexa 488 (1:200) (Invitrogen), respectively. Five larval gut specimens per Gal4 line were dissected and examined.

Cell dissociation and isolation

Midguts from 3rd instar larvae were dissected in phosphate-buffered saline (PBS) and transferred to microcentrifuge tubes on ice containing PBS + 30% normal goat serum (NGS). After dissection, 150 μ L of 2.7 mg/mL elastase was added to each sample tube. The tubes were then incubated at 27°C for one hour. During incubation, samples were mixed by pipetting ~30 times every 15 minutes to improve elastase dissociation of the cells. Samples were then filtered through a 40 μ M FlowMi tip filter (Bel-Art) to reduce debris. Afterwards, the samples were quantified on the LUNA-FL Dual Fluorescence Cell Counter (Logos Biosystems) using 9 μ L of sample to 1 μ L AO/PI dye to ensure there were enough viable cells for flow sorting.

Once quantified, the samples were brought up to a volume of ~1.1 mL with the PBS + 30% NGS solution to facilitate flow sorting. The samples were then fluorescently sorted on a FACSaria II Cell Sorter (BD Biosciences) to isolate GFP+ cells. Following sorting, samples were centrifuged at 300x g for 10 minutes to concentrate the cells. The supernatant was aspirated off until 50 μ L cell concentrate remained in each sample. Then, the samples were carefully resuspended using wide bore pipette tips before being combined into one sample tube. This sample was quantified on the LUNA-FL Dual Fluorescence Cell Counter (Logos Biosystems) as described above. If necessary, cells were centrifuged, concentrated, and re-counted.

Preparation of single-cell sequencing libraries

The resulting pool was prepared for sequencing following the 10x Genomics Single Cell 3' protocol (version CG000315 Rev C). At step 2.2 of the protocol, cDNA amplification, 1 μ L of TaG-EM spike-in primer (10 μ M) was added to the reaction to amplify cDNA with the TaG-EM barcode. Gene expression cDNA and TaG-EM cDNA were separated using a double-sided SPRIselect (Beckman Coulter) bead clean up following 10x Genomics Single Cell 3' Feature Barcode protocol, step 2.3 (version CG000317 Rev E). The gene expression cDNA was created into a library following the CG000315 Rev C protocol starting at section 3. Custom nested primers were used for enrichment of TaG-EM barcodes after cDNA creation using PCR.

The following primers were tested (see **Supplemental Figure 18** [↗](#)):

UMGC_IL_TaGEM_SpikeIn_v1:

GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTTCCAACAACCGGAAGT*G*A

UMGC_IL_TaGEM_SpikeIn_v2:

GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGCAGCTTATAACTTCCAACAACCGGAAGT*G*A

UMGC_IL_TaGEM_SpikeIn_v3: TGTGCTCTTCCGATCTGCAGCTTATAACTTCCAACAACCGGAAGT*G*A

D701_TaGEM: CAAGCAGAAGACGGCATAACGAGATCGAGTAATGTGACTGGAGTTCAGACGTGTGCTCTTCGATCTGCAGC*T*T

SI PCR Primer: AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGC*T*C

UMGC_IL_DoubleNest:

GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGCAGCTTATAACTTCCAACAACCGG*A* A

P5: AATGATACGGCGACCACCGA

D701: GATCGGAAGAGCACACGTCTGAACTCCAGTCACATTACTCGATCTCGTATGCCGTCTTCTG CTTG

D702: GATCGGAAGAGCACACGTCTGAACTCCAGTCACTCCGAGAATCTCGTATGCCGTCTTCT GCTTG

After multiple optimization trials, the following steps yielded ~96% on-target reads for the TaG-EM library (**Supplemental Figure 18** [↗](#), note that for the enriched barcode data shown in **Figure 6** [↗](#) and **Supplemental Figure 19** [↗](#), a similar amplification protocol was used TaG-EM barcodes were amplified from the gene expression library cDNA and not the SPRI-selected barcode pool). TaG-EM cDNA was amplified with the following PCR reaction: 5 µl purified TaG-EM cDNA, 50 µl 2x KAPA HiFi ReadyMix (Roche), 2.5 µl UMGC_IL_DoubleNest primer (10 µM), 2.5 µl SI_PCR primer (10 µM), and 40 µl nuclease-free water. The reaction was amplified using the following cycling conditions: 98°C for 2 minutes, followed by 15 cycles of 98°C for 20 seconds, 63°C for 30 seconds, 72°C for 20 seconds, followed by 72°C for 5 minutes. After the first PCR, the amplified cDNA was purified with a 1.2x SPRIselect (Beckman Coulter) bead cleanup with 80% ethanol washes and eluted into 40 µl of nuclease-water. A second round of PCR was run with following reaction: 5 µl purified TaG-EM cDNA, 50 µl 2x KAPA HiFi ReadyMix (Roche), 2.5 µl D702 primer (10 µM), 2.5 µl p5 Primer (10 µM), and 40 µl nuclease-free water. The reaction was amplified using the following cycling conditions: 98°C for 2 minutes, followed by 10 cycles of 98°C for 20 seconds, 63°C for 30 seconds, 72°C for 20 seconds, followed by 72°C for 5 minutes. After the second PCR, the amplified cDNA was purified with a 1.2x SPRIselect (Beckman Coulter) bead cleanup with 80% ethanol washes and eluted into 40 µl of nuclease-water. The resulting 3' gene expression library and TaG-EM enrichment library were sequenced together following Scenario 1 of the BioLegend "Total-Seq-A Antibodies and Cell Hashing with 10x Single Cell 3' Reagents Kit v3 or v3.1" protocol. Additional sequencing of the enriched TaG-EM library also done following Scenario 2 from the same protocol.

Sequencing

Libraries for TaG-EM barcode analysis from structured pools or from phototaxis or oviposition experiments were denatured with NaOH and prepared for sequencing according to the protocols described in the Illumina MiSeq Denature and Dilute Libraries Guides. Single-cell libraries were sequenced on the Illumina NextSeq 2000 or Illumina NovaSeq 6000. One of the single cell enriched TaG-EM barcode libraries was sequenced on an Element Aviti sequencer following the manufacturers loading instructions.

Data analysis

Behavioral experiments

Demultiplexed fastq files were generated using bcl2fastq or bcl-convert. TaG-EM barcode data was analyzed using custom R and Python scripts and BioPython (Cock et al., 2009 [↗](#)). Leading primer sequences were trimmed using cutadapt (Martin, 2011 [↗](#)) and the first 14 bp of the remaining trimmed read were compared to a barcode reference file, with a maximum of 2 mismatches allowed, using a custom script (TaG-EM_barcode_analysis.py) which is available via Github: <https://github.com/darylgohl/TaG-EM> [↗](#).

Single-cell experiments

Datasets were first mapped and analyzed using the Cell Ranger analysis pipeline (10x Genomics). A custom Drosophila genome reference was made by combining the BDGP.28 reference genome assembly and Ensembl gene annotations. Custom gene definitions for each of the TaG-EM barcodes were added to the fasta genome file and .gtf gene annotation file. A Cell Ranger reference package was generated with the Cell Ranger *mkref* command. Subsequent single-cell data analysis

was performed using the R package Seurat (Satija et al., 2015 [↗](#)). Cells expressing less than 200 genes and genes expressed in fewer than three cells were filtered from the expression matrix. Next, percent mitochondrial reads, percent ribosomal reads cells counts, and cell features were graphed to determine optimal filtering parameters. DecontX (Yang et al., 2020 [↗](#)) was used to identify empty droplets, to evaluate ambient RNA contamination, and to remove empty cells and cells with high ambient RNA expression. DoubletFinder (McGinnis et al., 2019 [↗](#)) to identify droplet multiplets and remove cells classified as multiplets. Clustree (Zappia and Oshlack, 2018 [↗](#)) was used to visualize different clustering resolutions and to determine the optimal clustering resolution for downstream analysis. Finally, SingleR (Aran et al., 2019 [↗](#)) was used for automated cell annotation with a gut single-cell reference from the Fly Cell Atlas (Li et al., 2022 [↗](#)). The dataset was manually annotated using the expression patterns of marker genes known to be associated with cell types of interest. To correlate TaG-EM barcodes with cell IDs in the enriched TaG-EM barcode library, a custom Python script was used (TaG-EM_barcode_Cell_barcode_correlation.py), which is available via Github: <https://github.com/darylgohl/TaG-EM> [↗](#).

Declarations

Availability of data, code, and materials

Sequencing data for this project is available through the National Center for Biotechnology Information (NCBI) Sequence Read Archive BioProject PRJNA912199. Fly stocks containing 20 of the TaG-EM barcodes together with an additional UAS hexameric GFP expression construct will be available from the Bloomington Drosophila Stock Center. Additional TaG-EM barcode stocks are available upon request. Single cell analysis code and the TaG-EM barcode analysis script and barcode reference fasta files are available via Github: <https://github.com/darylgohl/TaG-EM> [↗](#).

Competing interests

None to declare.

Funding

This study was supported by a grant from the Winston and Maxine Wallin Neuroscience Discovery Fund.

Author contributions

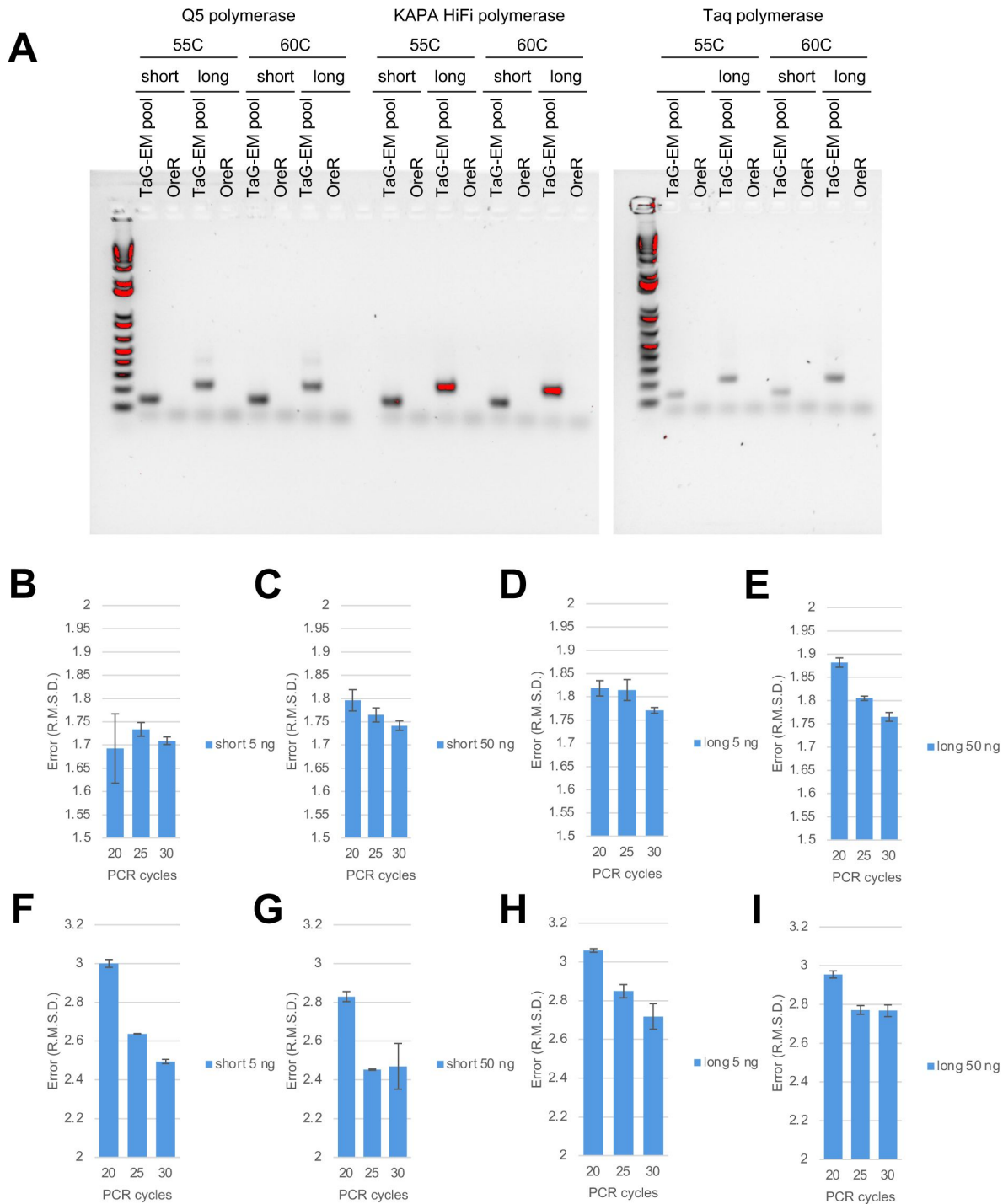
J.B.M. conceived and designed experiments, conducted experiments, analyzed data, revised manuscript, M.D. conducted experiments, analyzed data, revised manuscript, L.G. conducted experiments, revised manuscript, B.A. conducted experiments, revised manuscript, J.G. analyzed data, revised manuscript, D.M.G. conceived and designed experiments, conducted experiments, analyzed data, wrote and revised the manuscript.

Acknowledgements

We thank our colleagues in the University of Minnesota Genomics Center (RRID:SCR_012413), in particular Aaron Becker, Dylan Cole, and Logan Silber for help with DNA sequencing, Emma Stanley, Fernanda Rodriguez, and Patrick Grady for assistance and advice on single-cell sequencing, and Thomas Clandinin, Nam Chul Kim, Kenneth Beckman, Andrew Alegria, Troy Louwagie, and Aaron Barnes for helpful feedback and discussions. This work was supported by the resources and staff at the University of Minnesota University Imaging Centers (RRID:SCR_020997). The authors acknowledge the Minnesota Supercomputing Institute (MSI) at the

University of Minnesota for providing resources that contributed to the research results reported within this paper. URL: <http://www.msi.umn.edu> . Stocks obtained from the Bloomington Drosophila Stock Center (NIH P40OD018537) were used in this study. Antibodies were obtained from the Developmental Studies Hybridoma Bank (DSHB). Prospero (MR1A) was deposited to the DSHB by C.Q. Doe and Nub 2D4 was deposited to the DSHB by Michalis Averof.

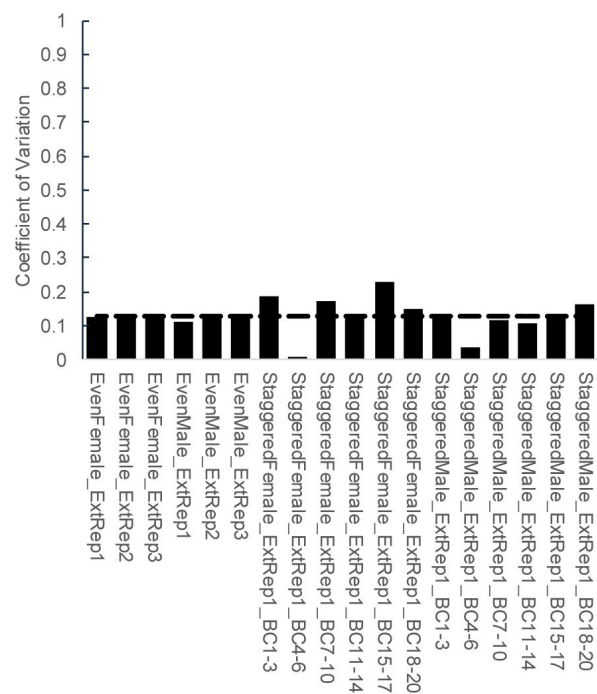
Supplemental Figures



Supplemental Figure 2.

Optimization of TaG-EM barcode amplification.

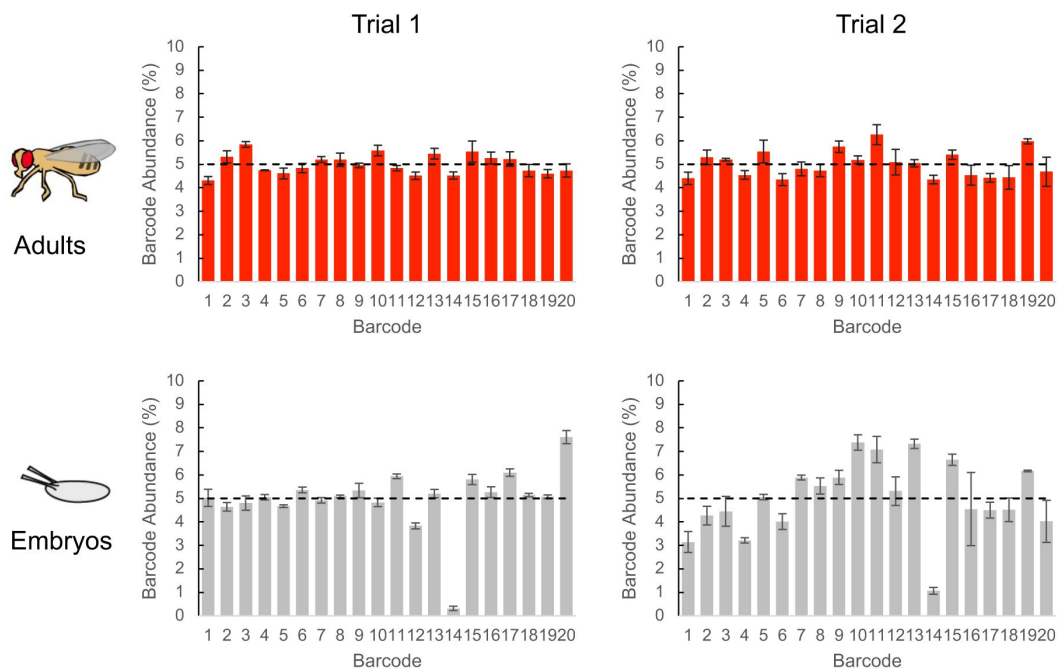
A) Gels showing bands produced when amplifying TaG-EM flies or a wild type control with the indicated polymerase, annealing temperature, and primer pair (short = B2_3'F1_Nextera/ SV40_pre_R_Nextera; long = B2_3'F1_Nextera/ SV40_post_R_Nextera). B-E) Mean error (R.M.S.D. root mean squared deviation from expected value) for even pool amplified with the indicated primer set, input amount, and cycle number using KAPA HiFi polymerase (n=3, error bars are +/- S.E.M.). F-G) Mean error (R.M.S.D. root mean squared deviation from expected value) for staggered pool amplified with the indicated primer set, input amount, and cycle number using KAPA HiFi polymerase (n=3, error bars are +/- S.E.M.).



Supplemental Figure 3.

Coefficient of variation for TaG-EM structured pools.

Plot showing coefficient of variation for different groups of TaG-EM barcodes in the structured pools. Dashed line indicates the mean coefficient of variation across all conditions.

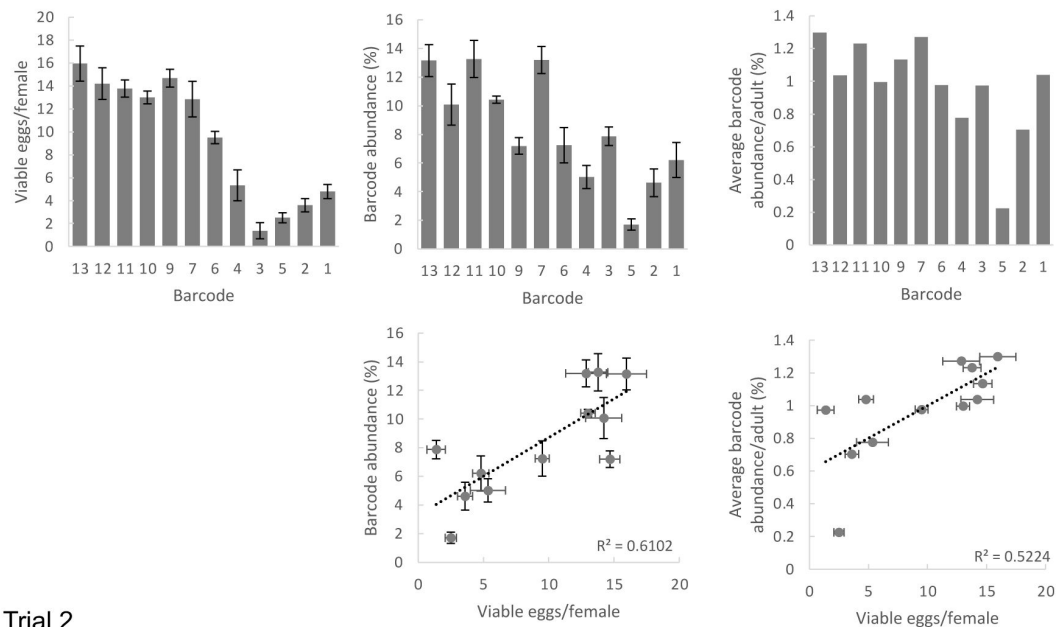


Supplemental Figure 4.

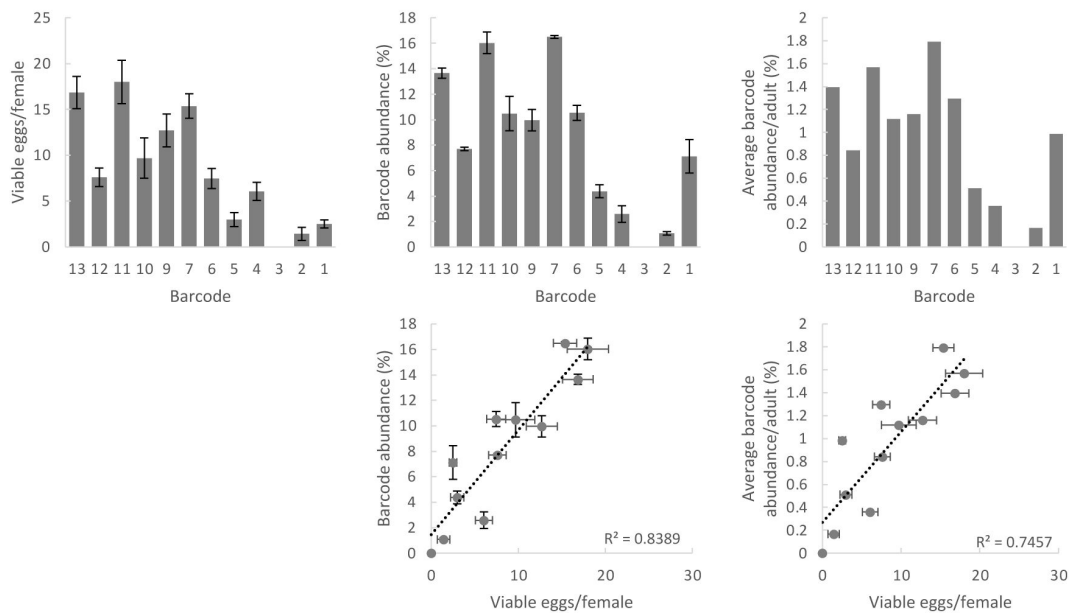
Oviposition tests with TaG-EM barcode lines.

Plots showing mean TaG-EM barcode abundance for adult females used in oviposition experiments (top) and eggs collected from these females (bottom). Data from two independent trials is shown ($n=3$ for each trial, error bars are $\pm$ S.E.M.). Dashed lines indicate the expected abundance values.

Trial 1



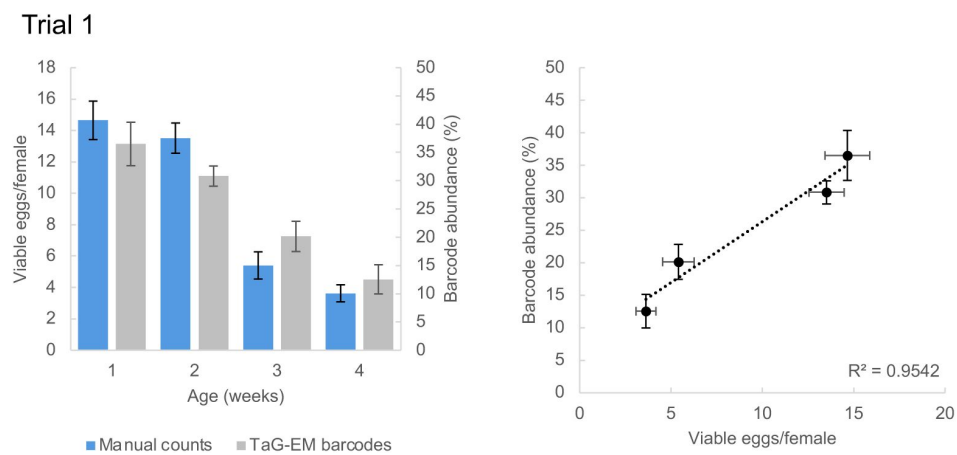
Trial 2



Supplemental Figure 5.

Fecundity data for individual TaG-EM lines.

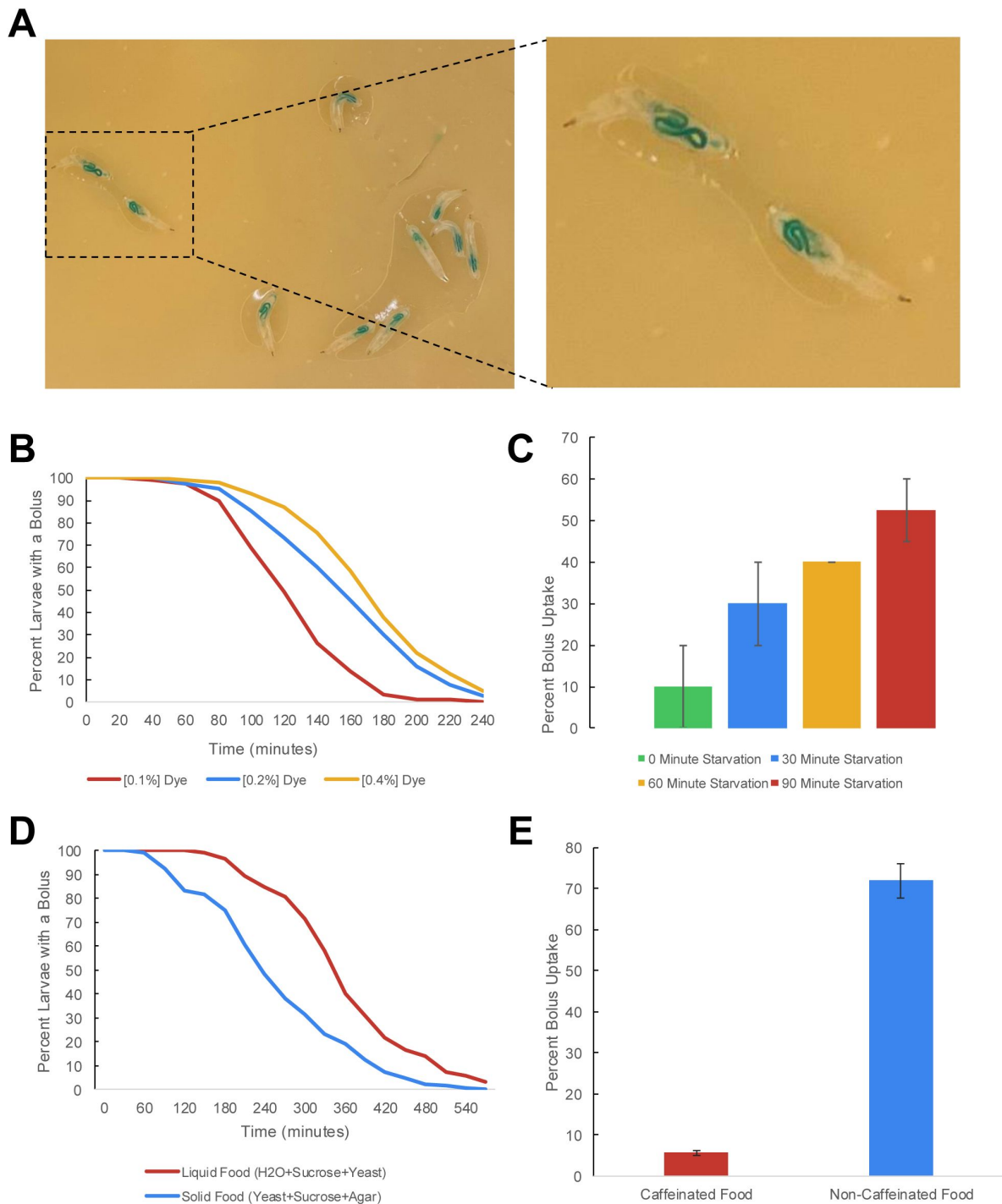
Manually collected data for mean number of viable eggs per female, barcode abundance data, and barcode abundance data normalized to adult fly barcode data for each of the TaG-EM barcode lines used in the age-dependent fecundity experiment. Scatterplots show correlations between manually collected data and barcode sequencing results. Data from two independent trials is shown (n=3 for each trial, error bars are +/- S.E.M.).



Supplemental Figure 6.

Average age-dependent fecundity data for Trial 1.

Average number of viable eggs per female (manual counts) and average barcode abundance are shown both as a bar plot and scatter plot ($n=3$ biological replicates for 3 barcodes per condition, error bars are $\pm$ S.E.M.). Data from Trial 2 is shown in [Figure 3C](#).

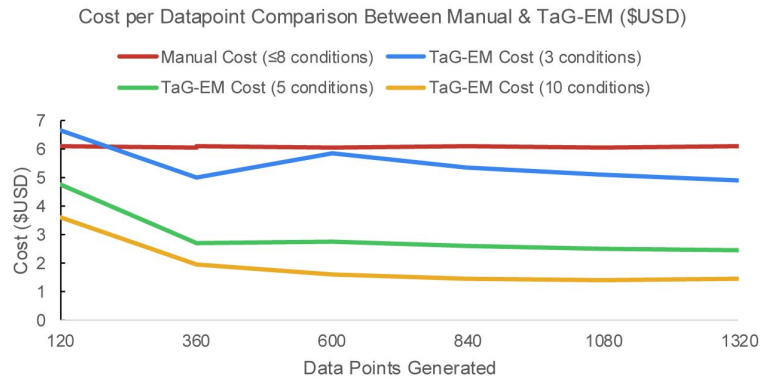


Supplemental Figure 7.

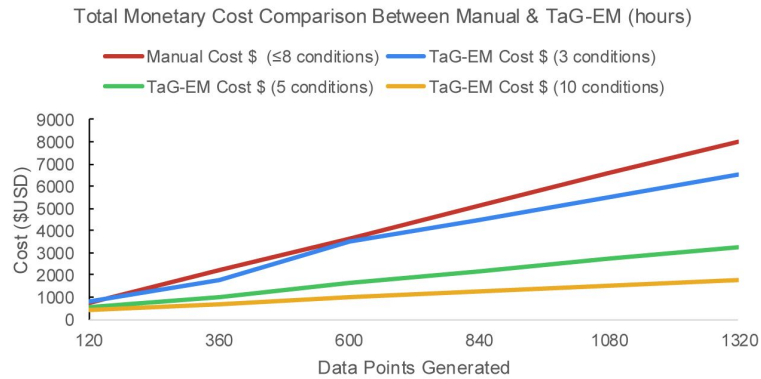
Larval gut motility assay parameters.

A) Images of larvae fed with blue-dyed yeast agar. B) Effect of dye concentration on food transit time. C) Effect of starvation time on feeding and uptake of the dyed food bolus. D) Effect of liquid versus solid diet on food transit time. E) Aversive effect of caffeine on food bolus uptake.

A Cost per Datapoint Analysis



B Monetary Cost Analysis (\$USD)



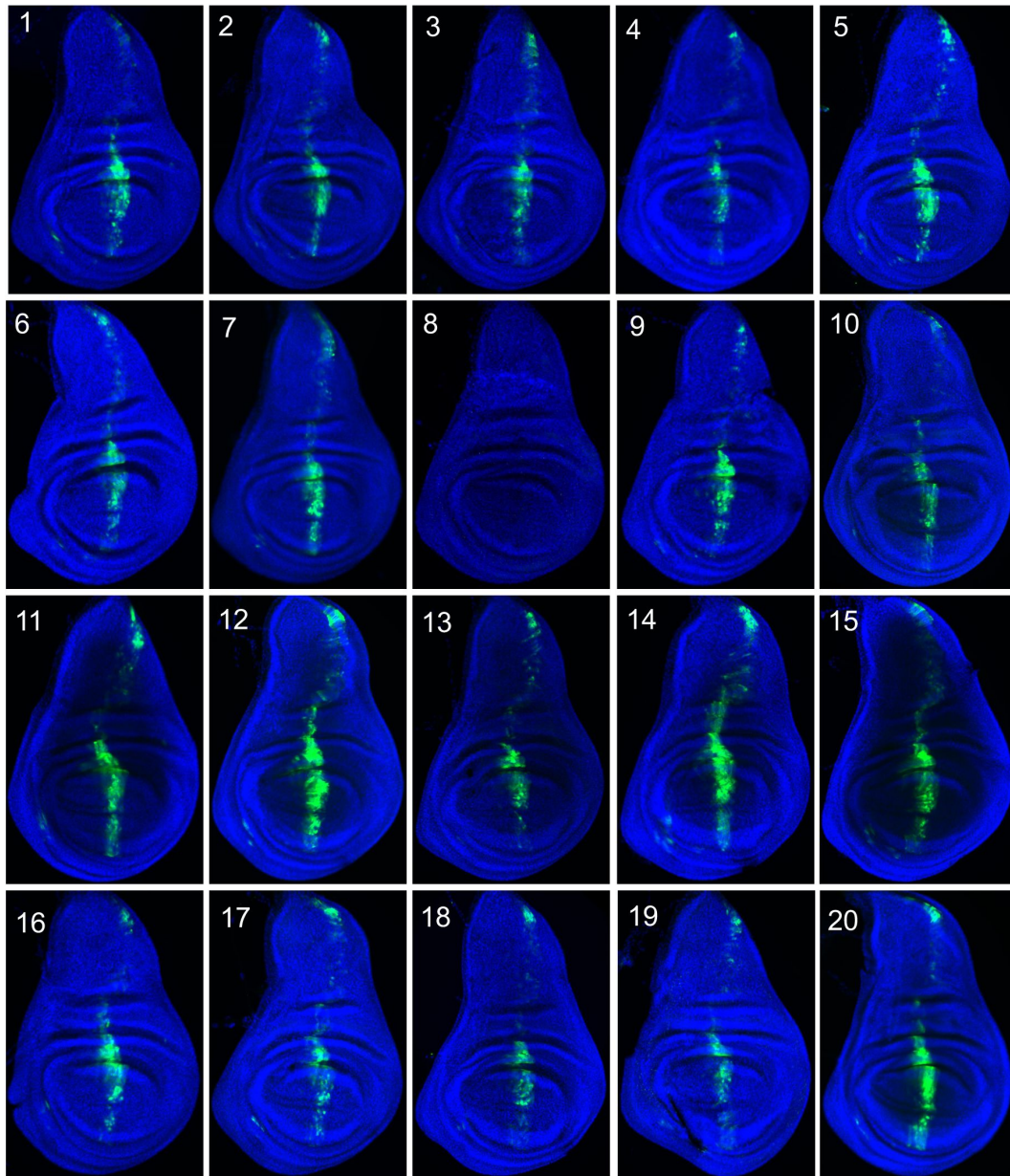
C Labor Time Cost Analysis (Hours)



Supplemental Figure 8.

Cost comparisons for manual and TaG-EM gut motility assays.

A) Cost per data point as a function of the number of data points generated and the number of experimental conditions screened. B) Overall experiment cost and C) labor effort as a function of the number of data points generated and the number of experimental conditions screened.

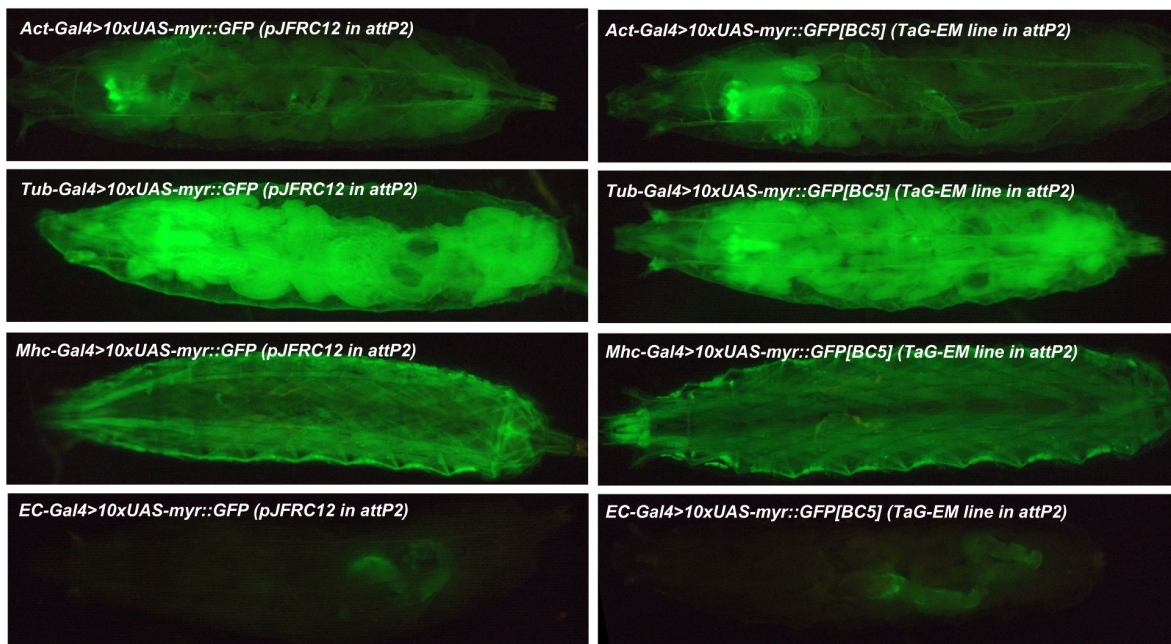


Supplemental Figure 9.

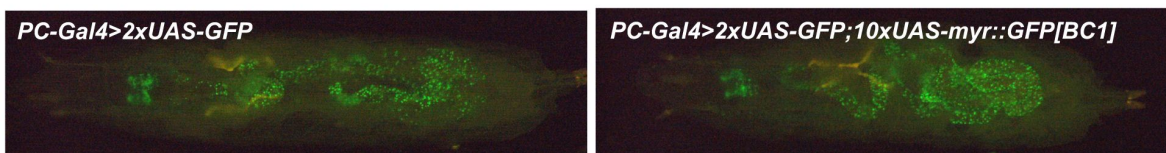
Expression driven by *dpp-Gal4* for 20 TaG-EM lines.

GFP antibody staining in the wing imaginal disc for the indicated TaG-EM barcode line driven by *dpp-Gal4*. Wing discs are counterstained with DAPI.

A



B



Supplemental Figure 10.

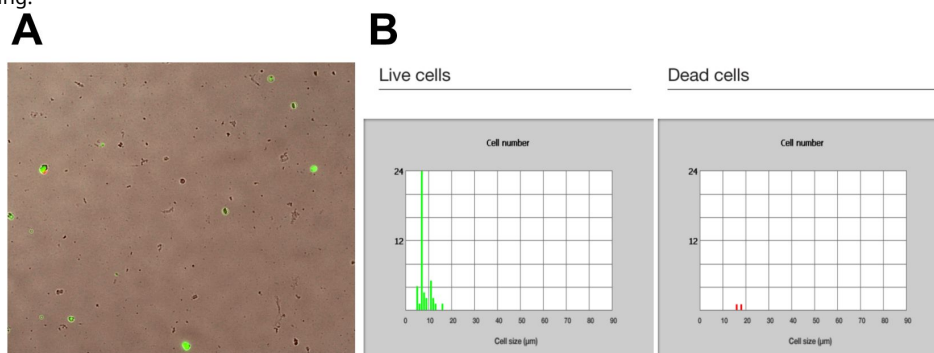
TaG-EM line GFP expression driven by different Gal4 drivers.

A) Comparison of endogenous GFP expression in larvae for the original pJFRC12 construct inserted in the attP2 landing site (left) or for a TaG-EM barcode line (right) expressed under the control of the indicated driver line. B) GFP expression of the PC-Gal (Precursor-Gal4) driver line together with either UAS-2xGFP or a combination of UAS-2xGFP and a TaG-EM barcode line.

Supplemental Figure 11.

Dissociated intestinal cell viability.

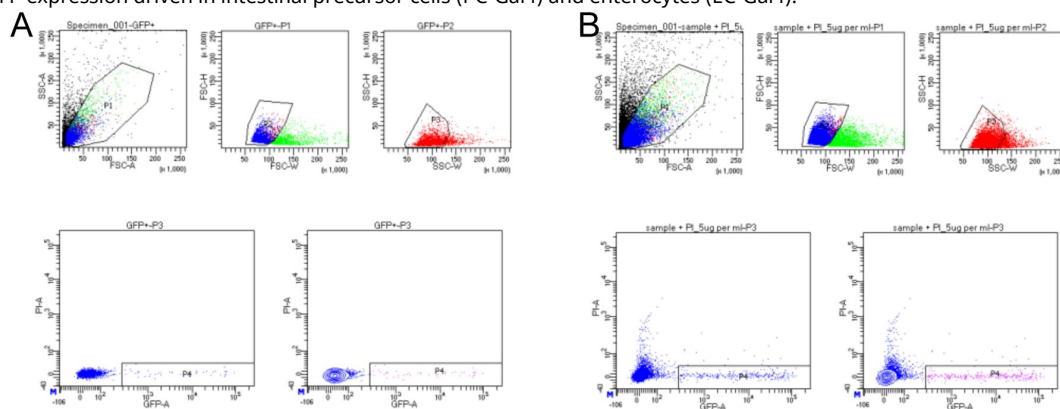
A) GFP expression visualized in dissociated cells from gut driver lines crossed to hexameric GFP and TaG-EM line. B) Proportion of live (left) and dead (right) cells post-isolation and flow sorting as assessed by GFP expression and propidium iodide staining.



Supplemental Figure 12.

BD FACSDiva 8.0.1 gating for sorted cells.

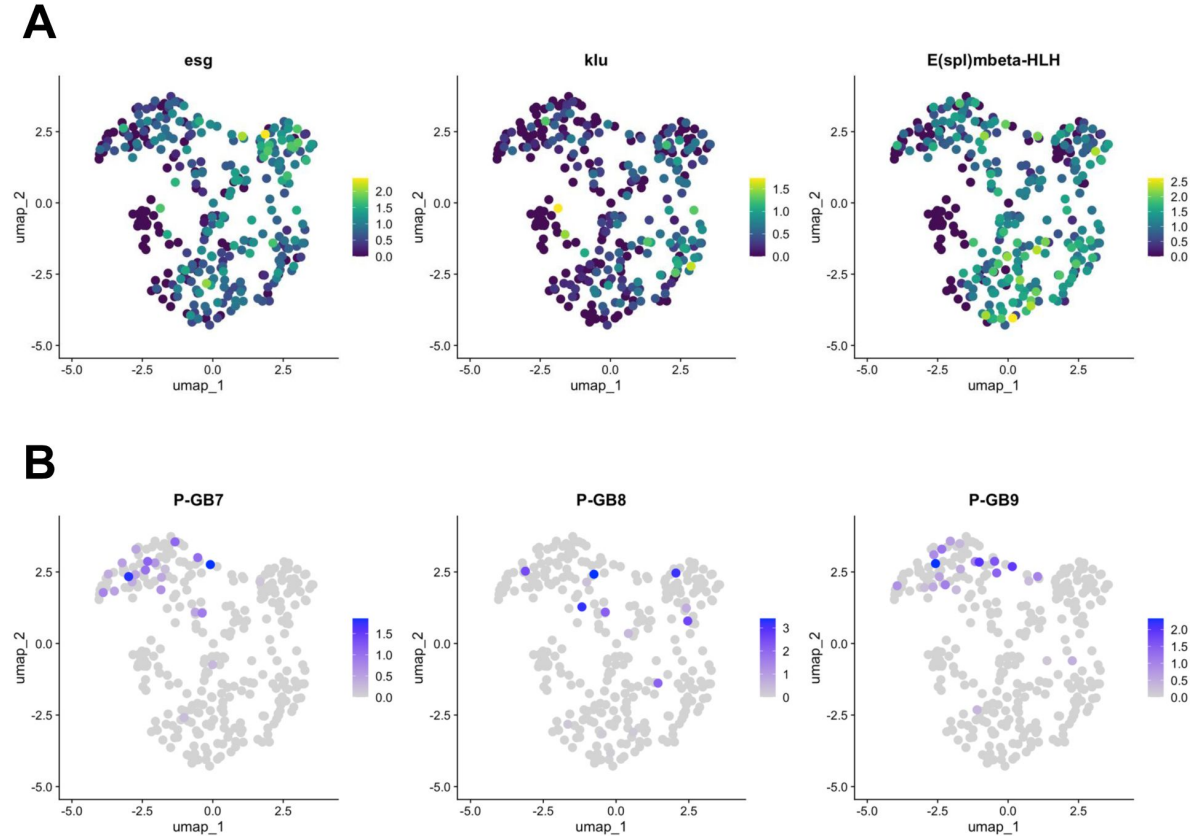
A) GFP gating created by analyzing a pool of GFP positive and negative cells. B) Flow gating for *Drosophila* gut cells with TaG-EM GFP expression driven in intestinal precursor cells (PC-Gal4) and enterocytes (EC-Gal4).



Supplemental Figure 13.

Expression of TaG-EM genetic barcodes in larval intestinal precursor cells.

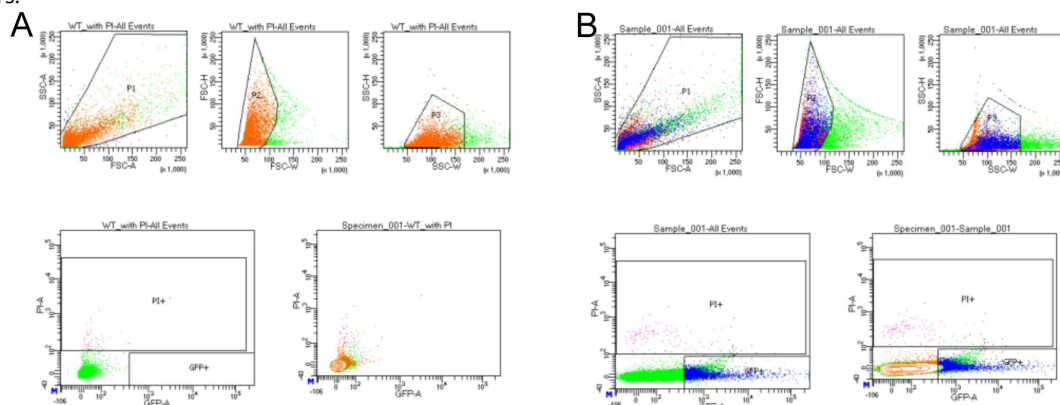
UMAP plots showing gene expression levels of A) enteroblast/ISC marker genes *esg*, *klu*, and *E(spl)mbeta-HLH*; and B) the TaG-EM barcodes 7, 8, and 9 driven by the PC-Gal4 line.

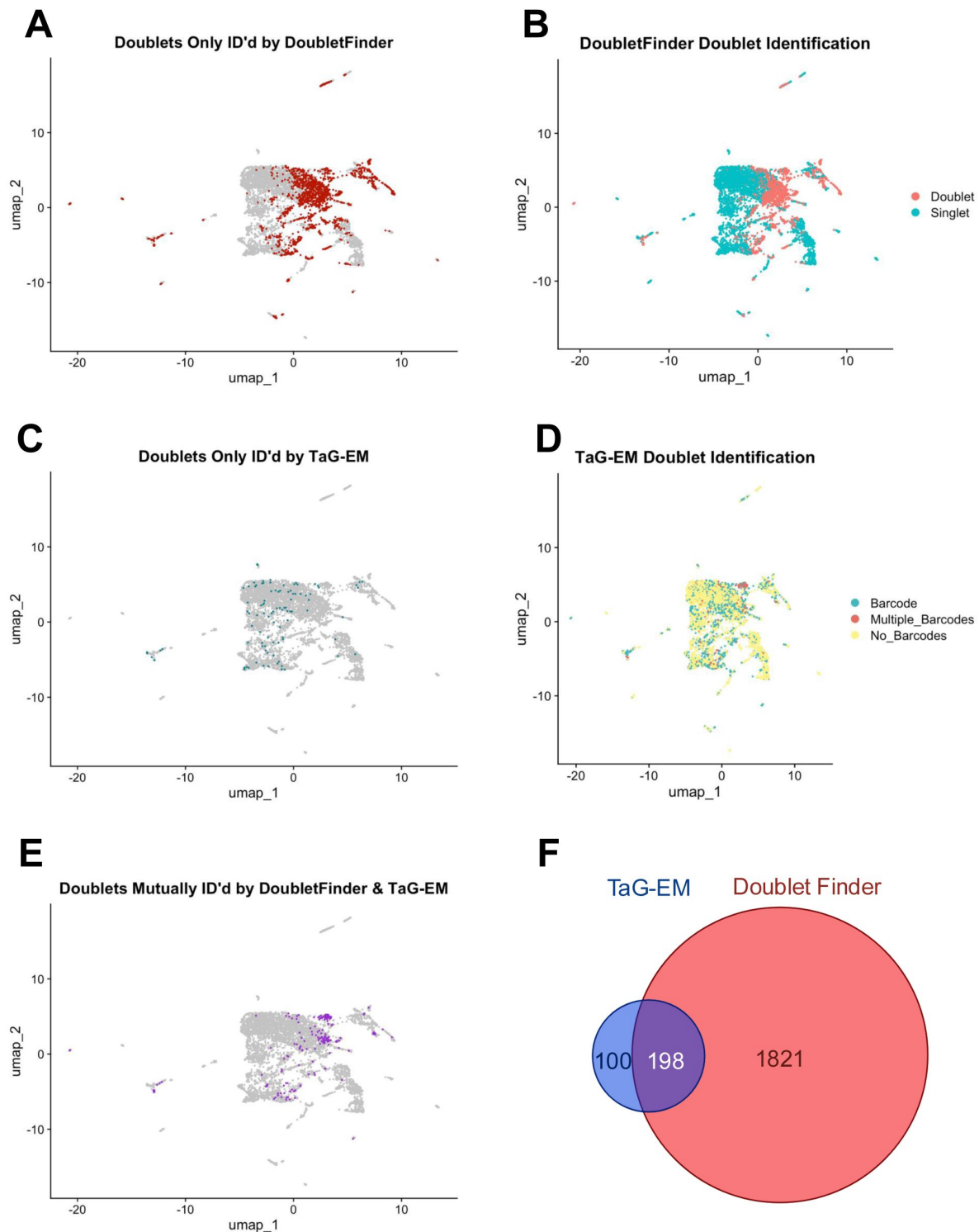


Supplemental Figure 14.

BD FACSDiva 8.0.1 gating for sorted cells.

A) Dead cell gating created by staining sample with propidium iodide (PI). B) Flow gating for *Drosophila* gut cells with TaG-EM and hexameric GFP expression driven by the pan-midgut, enteroblast, enterocyte, enteroendocrine, and precursor cell drivers.

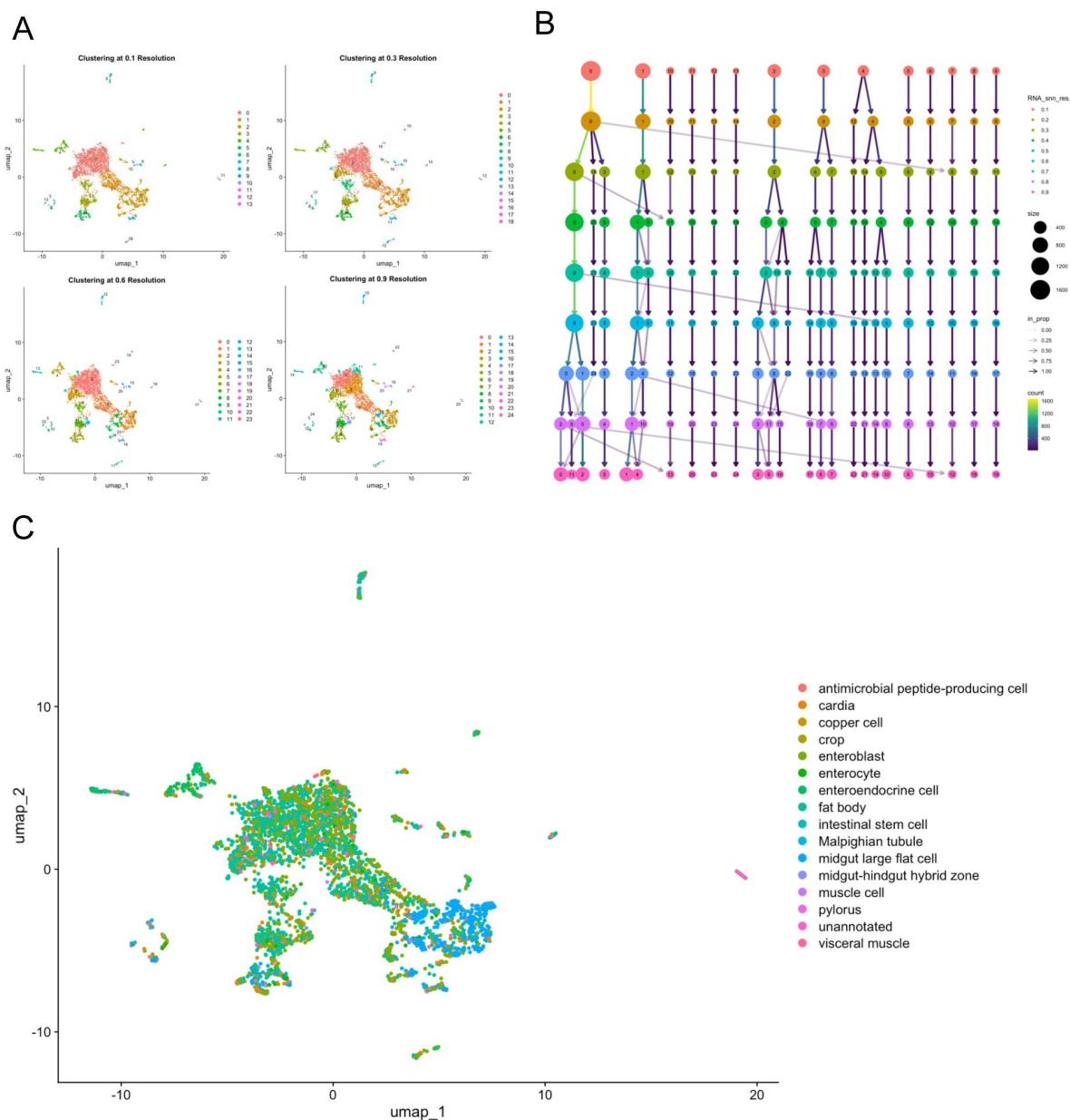




Supplemental Figure 15.

TaG-EM-based doublet identification.

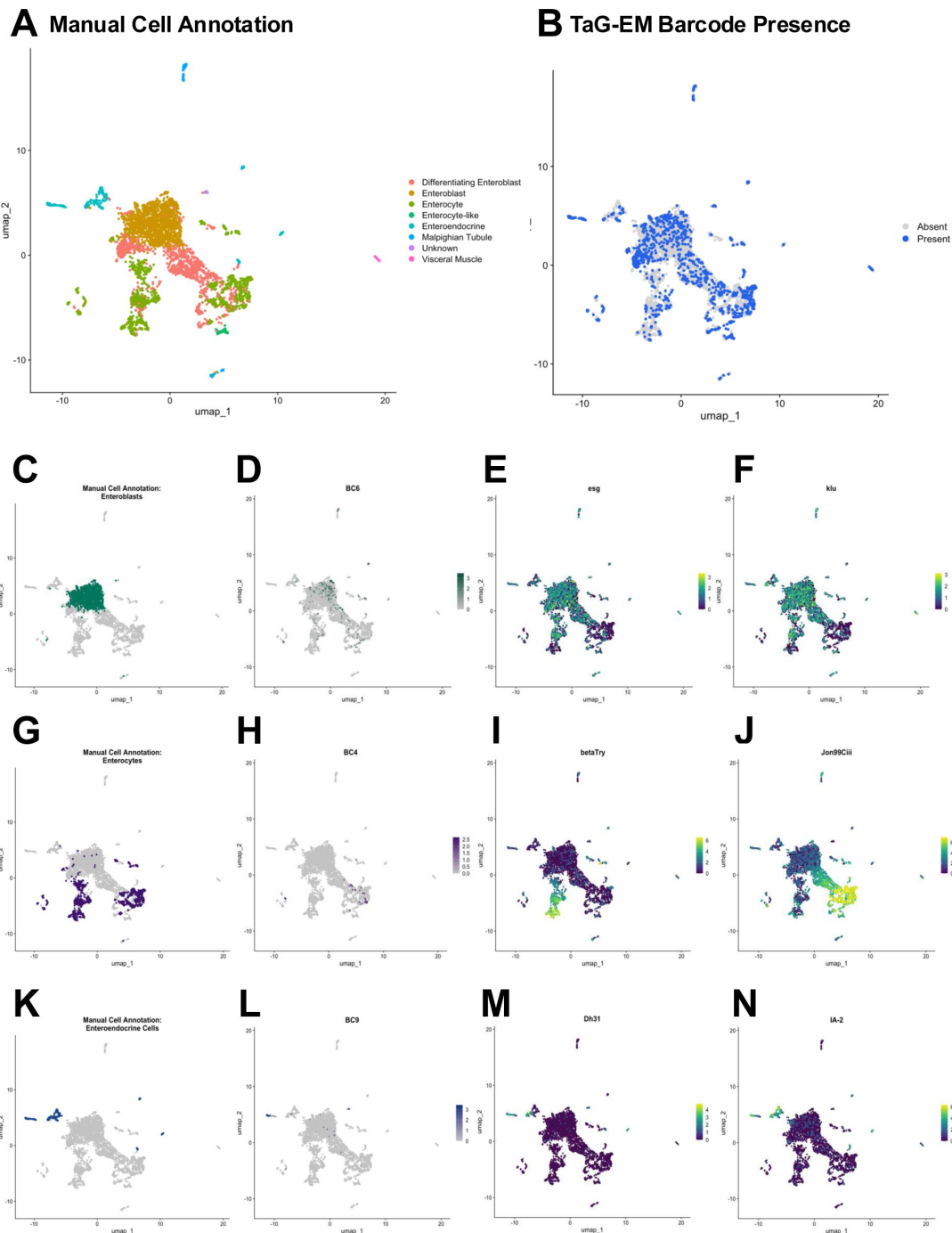
UMAP plots pre-doublet removal showing A) doublets uniquely identified by DoubletFinder, B) all doublets identified by DoubletFinder, C) doublets uniquely identified by TaG-EM barcodes, D) all doublets identified by TaG-EM barcodes, E) doublets mutually found by TaG-EM and DoubletFinder, F) Venn diagram of overlap between doublets identified by TaG-EM and DoubletFinder.



Supplemental Figure 16.

Clustering and automated annotation.

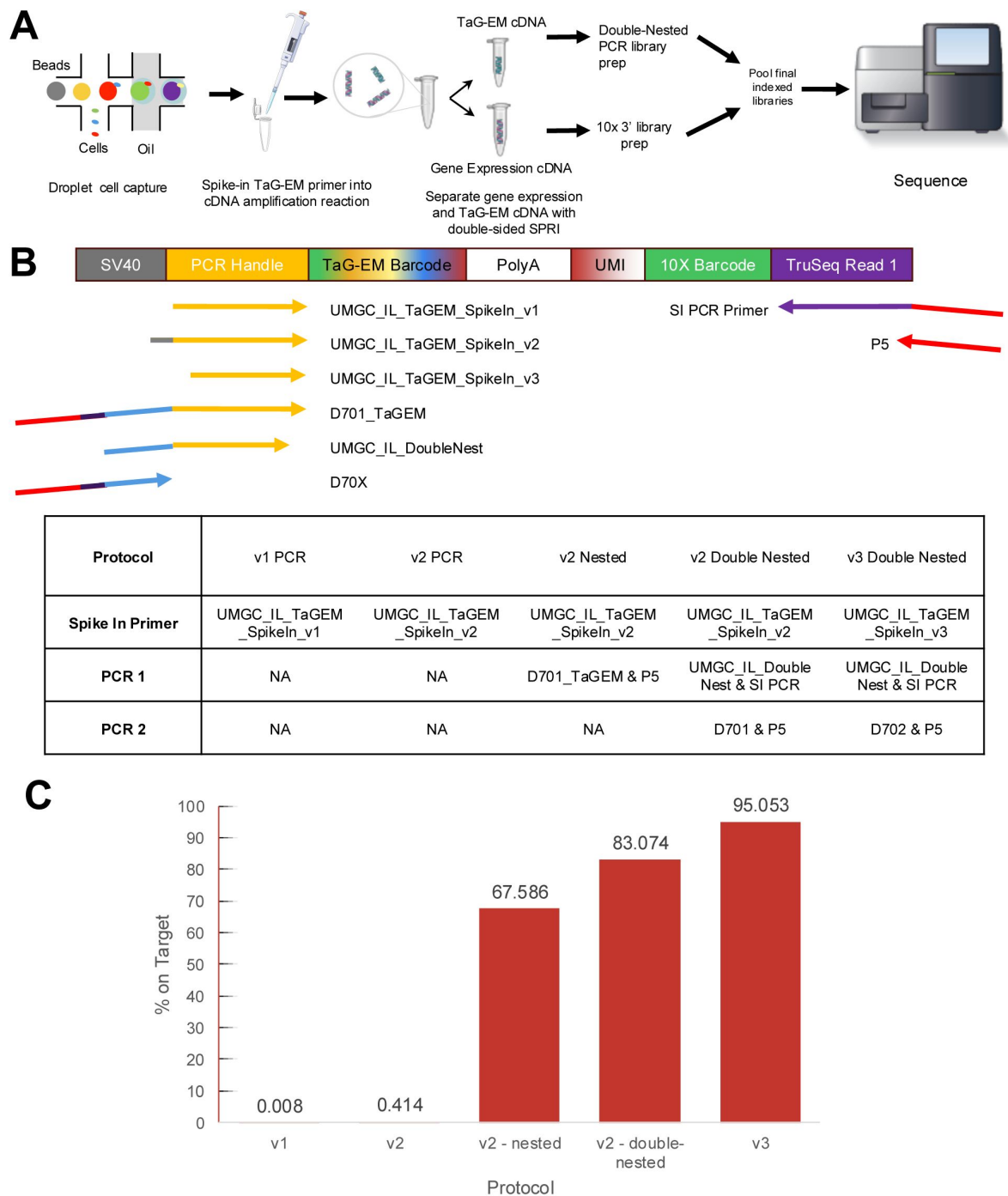
A) UMAP plots clustered at different resolutions. B) Clustree analysis of the effect of clustering resolution. C) Automated cell type annotation using data from the Fly Cell Atlas.



Supplemental Figure 17.

Expression of TaG-EM genetic barcodes in larval intestinal cell types.

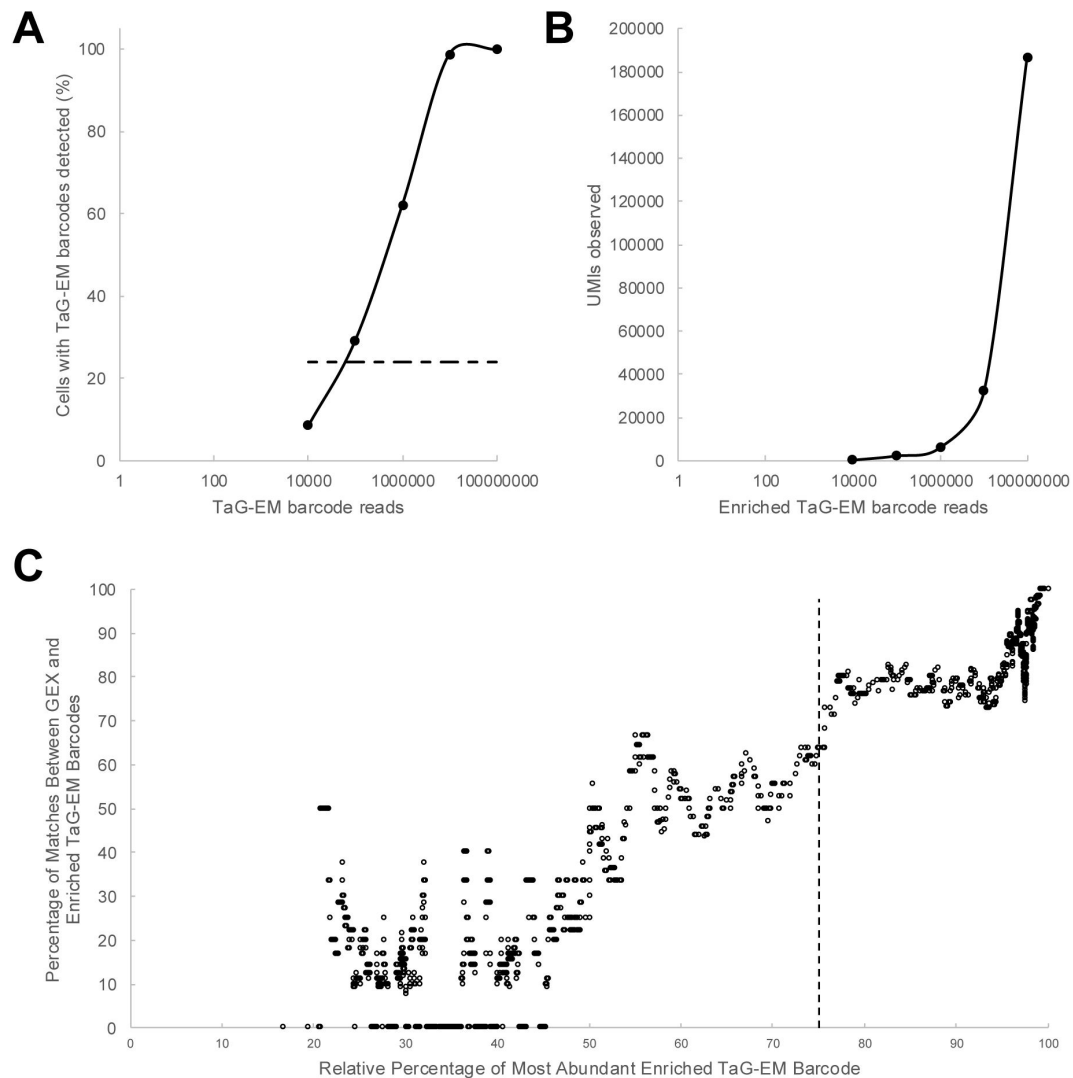
A) UMAP plot of *Drosophila* larval gut cell types. B) Annotation of cells associated with a TaG-EM barcode across all 8 multiplexed experimental conditions using data from the gene expression library only. C) Annotated enteroblast cells. D) Expression level of TaG-EM barcode (BC6) driven by the EB-Gal4 line using data from the gene expression library only. Gene expression levels of enteroblast marker genes E) *esg*, F) *klu*. G) Annotated enterocyte cells. H) Expression level of TaG-EM barcode (BC4) driven by the EC-Gal4 line using data from the gene expression library only. Gene expression levels of enterocyte marker genes I) *betaTry*, J) *Jon99Ciii*. K) Annotated enteroendocrine cells. L) Expression level of TaG-EM barcode (BC9) driven by the EE-Gal4 line using data from the gene expression library only. Gene expression levels of enteroendocrine cell marker genes M) *Dh31*, N) *IA-2*.



Supplemental Figure 18.

Optimizing amplification of the TaG-EM barcode library.

A) Workflow for single-cell capture; cDNA amplification with added spike-in primer for TaG-EM library followed by a SPRI size-selection clean-up, then PCR(s) to create library for sequencing. B) Spike-in primers and amplification primers used to enrich TaG-EM barcodes. Table summarizes different protocols tested to amplify the TaG-EM barcodes and create an enriched sequencing library. C) Percent of on-target reads for each enriched TaG-EM barcode library.

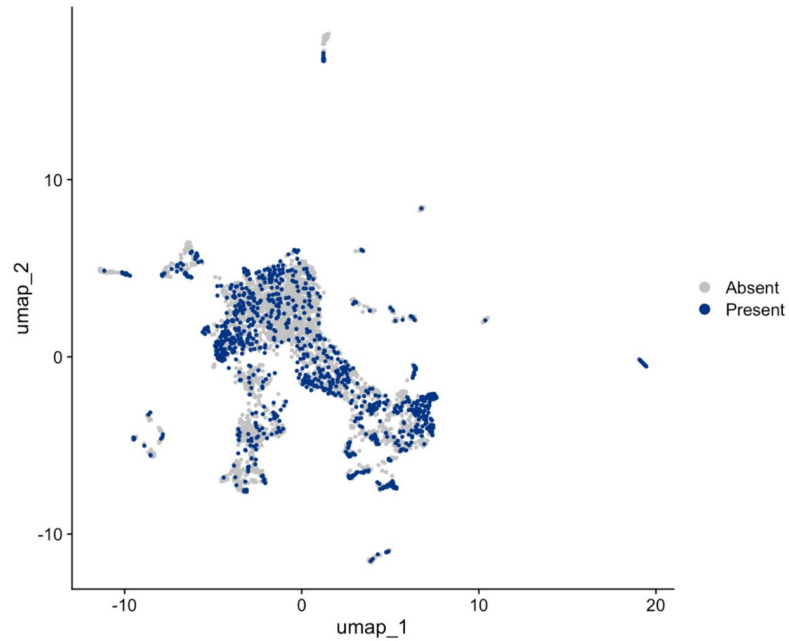


Supplemental Figure 19.

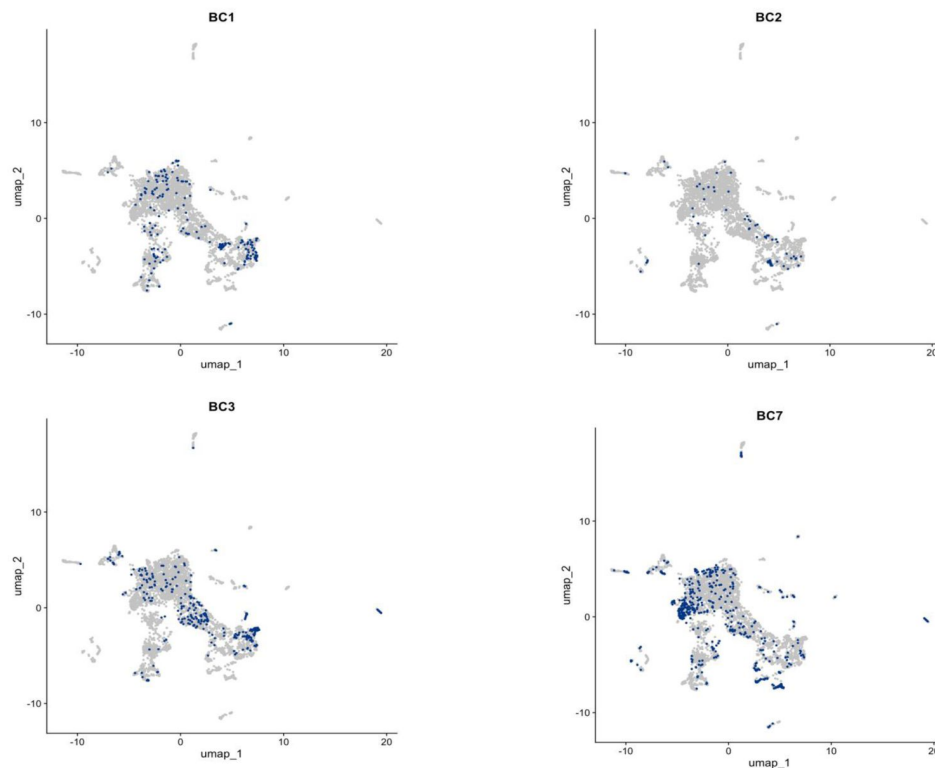
Performance of the enriched TaG-EM barcode library.

A) Proportion of cells with at least one barcode read assigned as a function of read depth for the enriched TaG-EM barcode library. Dashed line indicated percentage of cells with TaG-EM barcodes detected in the gene expression library B) Number of unique UMIs observed as a function of read depth. C) Correlation between barcodes detected in the gene expression (GEX) library and the enriched TaG-EM barcode library as a function of the purity of TaG-EM barcode assignment to the corresponding cell barcode. Dashed line indicates the threshold used for TaG-EM barcode calling in the enriched TaG-EM barcode library.

A Pan-midgut-associated TaG-EM barcode expression (BC1, BC2, BC3, BC7)



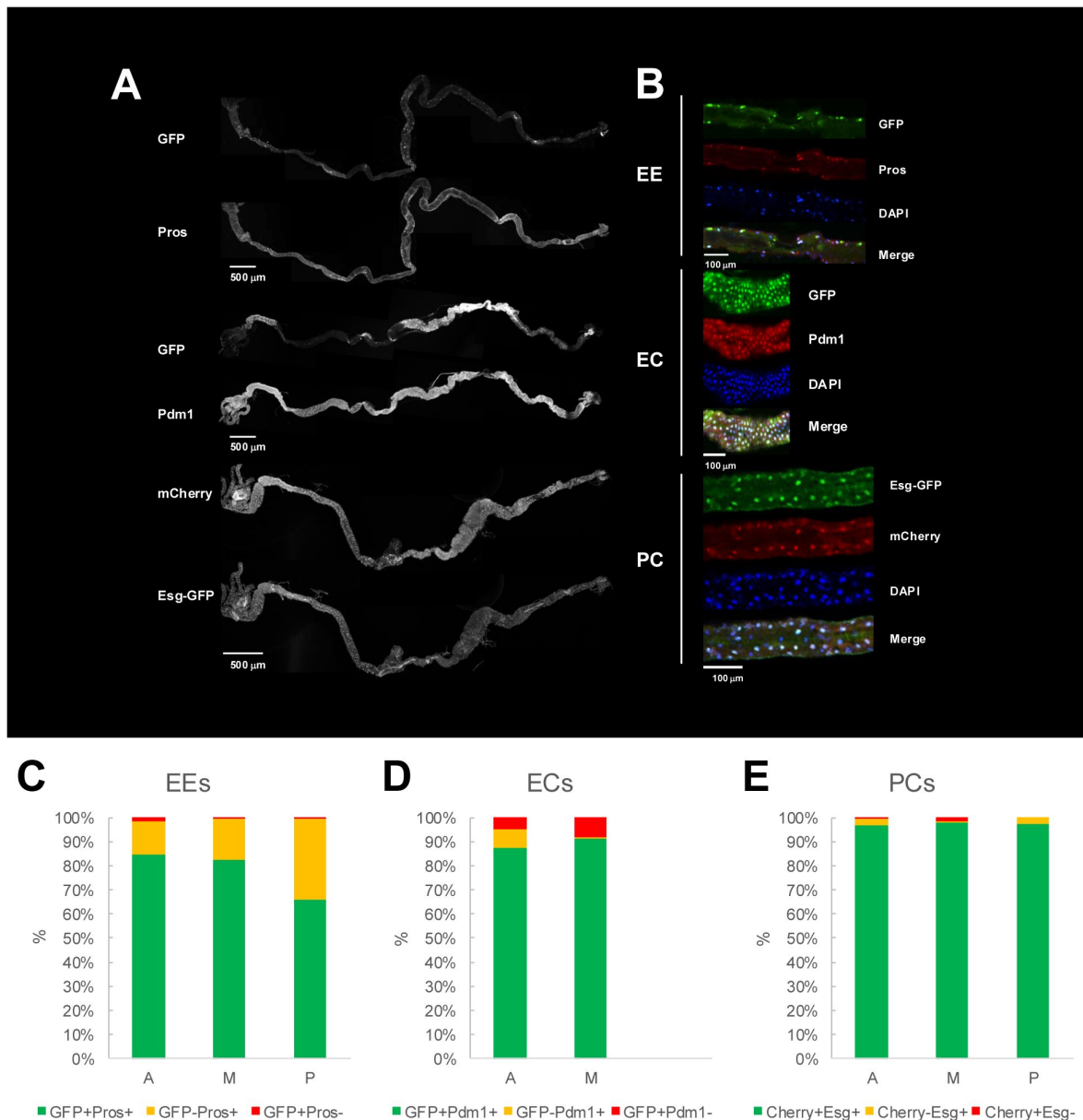
B



Supplemental Figure 20.

Expression of the PMG-Gal4 driven TaG-EM barcodes.

UMAP plots showing expression of the four PMG-Gal4 driven TaG-EM barcodes (BC1, BC2, BC3, and BC7) either A) in aggregate or B) individually.



Supplemental Figure 21.

Characterization of Gal4 line expression in the larval gut.

A) Confocal images of third instar midguts showing Gal4-driven fluorophore expression (GFP or mCherry) and comparison with immunostainings of the gut cell markers Prospero (enteroendocrine), Pdm1 (enterocyte) and Esg-GFP (progenitor cell). For each image, Z projections of the stacks recorded along the length of the midgut were manually stitched together. B) Representative single frames confocal images of a small region of the midgut showing immunostainings of the different gut cell markers and the Gal4-driven fluorophores. Quantification of overlapping and non-overlapping expression between the Gal4-driver fluorophore expression and the cell type marker in the anterior (A), middle (M), and posterior (P) regions for C) enteroendocrine cells (EC-Gal4), D) enterocytes (EC-Gal4), E) precursor cells (PC-Gal4). Five specimens for each Gal4 line were examined. In the case of the enterocyte-specific driver, only anterior and middle regions were analyzed since the driver is largely inactive in the posterior part of the midgut.

References

- Alegria AD, Joshi AS, Mendana JB, Khosla K, Smith KT, Auch B, Donovan M, Bischof J, Gohl DM, Kodandaramaiah SB (2024) **High-throughput genetic manipulation of multicellular organisms using a machine-vision guided embryonic microinjection robot** *Genetics* **226** <https://doi.org/10.1093/genetics/iyae025>
- Aran D *et al.* (2019) **Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage** *Nat Immunol* **20**:163–172 <https://doi.org/10.1038/s41590-018-0276-y>
- Ariyapala IS, Holsopple JM, Popodi EM, Hartwick DG, Kahsai L, Cook KR, Sokol NS (2020) **Identification of split-GAL4 drivers and enhancers that allow regional cell type manipulations of the drosophila melanogaster intestine** *Genetics* **216**:891–903 <https://doi.org/10.1534/genetics.120.303625>
- Asnicar F *et al.* (2021) **Blue poo: impact of gut transit time on the gut microbiome using a novel marker** *Gut* **70**:1665–1674 <https://doi.org/10.1136/gutjnl-2020-323877>
- Aso Y *et al.* (2014) **The neuronal architecture of the mushroom body provides a logic for associative learning** *eLife* **3** <https://doi.org/10.7554/ELIFE.04577>
- Bhang HC *et al.* (2015) **Studying clonal dynamics in response to cancer therapy using high-complexity barcoding** *Nature Medicine* **21**:440–448 <https://doi.org/10.1038/nm.3841>
- Buchon N, Osman D, David FPA, Yu Fang H, Boquete J-P, Deplancke B, Lemaitre B (2013) **Morphological and Molecular Characterization of Adult Midgut Compartmentalization in Drosophila** *Cell Reports* **3**:1725–1738 <https://doi.org/10.1016/j.celrep.2013.04.001>
- Chen X, Sun Y-C, Zhan H, Kebschull JM, Fischer S, Matho K, Huang ZJ, Gillis J, Zador AM (2019) **High-Throughput Mapping of Long-Range Neuronal Projection Using In Situ Sequencing** *Cell* **179**:772–786 <https://doi.org/10.1016/j.cell.2019.09.023>
- Cheng J, Liao J, Shao X, Lu X, Fan X (2021) **Multiplexing Methods for Simultaneous Large-Scale Transcriptomic Profiling of Samples at Single-Cell Resolution** *Advanced Science* **8** <https://doi.org/10.1002/advs.202101229>
- Clark IC *et al.* (2021) **Barcoded viral tracing of single-cell interactions in central nervous system inflammation** *Science* **372**
- Cock PJA *et al.* (2009) **Biopython: freely available Python tools for computational molecular biology and bioinformatics** *Bioinformatics (Oxford, England)* **25**:1422–3 <https://doi.org/10.1093/bioinformatics/btp163>
- David J, Cohet Y, Fouillet P (1975) **The variability between individuals as a measure of senescence: A study of the number of eggs laid and the percentage of hatched eggs in the case of Drosophila melanogaster** *Experimental Gerontology* **10**:17–25 [https://doi.org/10.1016/0531-5565\(75\)90011-X](https://doi.org/10.1016/0531-5565(75)90011-X)

Davis FP, Nern A, Picard S, Reiser MB, Rubin GM, Eddy SR, Henry GL (2020) **A genetic, genomic, and computational resource for exploring neural circuit function** *eLife* **9**

DeAngelis MM, Wang DG, Hawkins TL (1995) **Solid-phase reversible immobilization for the isolation of PCR products** *Nucleic Acids Res* **23**:4742–4743

Gohl DM, Silies MA, Gao XJ, Bhalerao S, Luongo FJ, Lin C-C, Potter CJ, Clandinin TR (2011) **A versatile in vivo system for directed dissection of gene expression patterns** *Nature Methods* **8**:231–237 <https://doi.org/10.1038/nmeth.1561>

Gohl DM *et al.* (2016) **Systematic improvement of amplicon marker gene methods for increased accuracy in microbiome studies** *Nature Biotechnology* **34**:942–949 <https://doi.org/10.1038/nbt.3601>

Groth AC, Fish M, Nusse R, Calos MP (2004) **Construction of transgenic Drosophila by using the site-specific integrase from phage phiC31** *Genetics* **166**:1775–82 <https://doi.org/10.1534/genetics.166.4.1775>

Hristova K, Wimley WC (2023) **Determining the statistical significance of the difference between arbitrary curves: A spreadsheet method** *PLoS One* **18** <https://doi.org/10.1371/journal.pone.0289619>

Huang AM, Rehm EJ, Rubin GM (2009) **Quick preparation of genomic DNA from Drosophila** *Cold Spring Harb Protoc* **2009** <https://doi.org/10.1101/pdb.prot5198>

Hung RJ *et al.* (2020) **A cell atlas of the adult Drosophila midgut** *Proceedings of the National Academy of Sciences of the United States of America* **117**:1514–1523 <https://doi.org/10.1073/pnas.1916820117>

Ingham PW (1988) **The molecular genetics of embryonic pattern formation in Drosophila** *Nature* **335**:25–34 <https://doi.org/10.1038/335025a0>

Kanca O *et al.* (2022) **An expanded toolkit for Drosophila gene tagging using synthesized homology donor constructs for CRISPR-mediated homologous recombination** *eLife* **11** <https://doi.org/10.7554/eLife.76077>

Keller J *et al.* (2018) **Expert consensus document: Advances in the diagnosis and classification of gastric and intestinal motility disorders** *Nat Rev Gastroenterol Hepatol* **15**:291–308 <https://doi.org/10.1038/nrgastro.2018.7>

Klein AM, Mazutis L, Akartuna I, Tallapragada N, Veres A, Li V, Peshkin L, Weitz DA, Kirschner MW (2015) **Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells** *Cell* **161**:1187–1201 <https://doi.org/10.1016/j.cell.2015.04.044>

Kurmangaliyev YZ, Yoo J, Valdes-Aleman J, Sanfilippo P, Zipursky SL (2020) **Transcriptional Programs of Circuit Assembly in the Drosophila Visual System** *Neuron* **108**:1045–1057 <https://doi.org/10.1016/j.NEURON.2020.10.006>

Lee JH *et al.* (2015) **Fluorescent in situ sequencing (FISSEQ) of RNA for gene expression profiling in intact cells and tissues** *Nature Protocols* **10**:442–458 <https://doi.org/10.1038/nprot.2014.191>

- Lein E, Borm LE, Linnarsson S (2017) **The promise of spatial transcriptomics for neuroscience in the era of molecular cell typing** *Science* **358**:64–69 <https://doi.org/10.1126/science.aan6827>
- Li H *et al.* (2022) **Fly Cell Atlas: A single-nucleus transcriptomic atlas of the adult fruit fly** *Science* **375** <https://doi.org/10.1126/science.abk2432>
- Ma D, Przybylski D, Abruzzi KC, Schlichting M, Li Q, Long X, Rosbash M (2021) **A transcriptomic taxonomy of drosophila circadian neurons around the clock** *eLife* **10**:1–19 <https://doi.org/10.7554/eLife.63056>
- Macosko EZ *et al.* (2015) **Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets** *Cell* **161**:1202–1214 <https://doi.org/10.1016/j.cell.2015.05.002>
- Martin KC, Ephrussi A (2009) **mRNA localization: gene expression in the spatial dimension** *Cell* **136**:719–30 <https://doi.org/10.1016/j.cell.2009.01.044>
- Martin M (2011) **Cutadapt removes adapter sequences from high-throughput sequencing reads** *EMBnet.journal* **17**:10–12
- Marusyk A, Almendro V, Polyak K (2012) **Intra-tumour heterogeneity: a looking glass for cancer?** *Nature Reviews Cancer* **12**:323–334 <https://doi.org/10.1038/nrc3261>
- McGinnis CS, Murrow LM, Gartner ZJ (2019) **DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors** *Cell Systems* **8**:329–337 <https://doi.org/10.1016/j.cels.2019.03.003>
- McKenna A, Findlay GM, Gagnon JA, Horwitz MS, Schier AF, Shendure J (2016) **Whole-organism lineage tracing by combinatorial and cumulative genome editing** *Science* **353** <https://doi.org/10.1126/science.aaf7907>
- Namiki S, Dickinson MH, Wong AM, Korff W, Card GM (2018) **The functional organization of descending sensory-motor pathways in drosophila** *eLife* **7**
- Pfeiffer BD *et al.* (2008) **Tools for neuroanatomy and neurogenetics in Drosophila** *Proceedings of the National Academy of Sciences of the United States of America* **105**:9715–20 <https://doi.org/10.1073/pnas.0803697105>
- Pfeiffer BD, Ngo T-TB, Hibbard KL, Murphy C, Jenett A, Truman JW, Rubin GM. (2010) **Refinement of tools for targeted gene expression in Drosophila** *Genetics* **186**:735–55 <https://doi.org/10.1534/genetics.110.119917>
- Reeves GT, Muratov CB, Schüpbach T, Shvartsman SY (2006) **Quantitative Models of Developmental Pattern Formation** *Developmental Cell* **11**:289–300 <https://doi.org/10.1016/j.DEVCEL.2006.08.006>
- Rubin GM, Spradling AC (1982) **Genetic transformation of Drosophila with transposable element vectors** *Science (New York, NY)* **218**:348–53 <https://doi.org/10.1126/science.6289436>
- Satija R, Farrell JA, Gennert D, Schier AF, Regev A (2015) **Spatial reconstruction of single-cell gene expression data** *Nature Biotechnology* **33**:495–502 <https://doi.org/10.1038/nbt.3192>

- Shearin HK, MacDonald IS, Spector LP, Steven Stowers R (2014) **Hexameric GFP and mCherry reporters for the Drosophila GAL4, Q, and LexA transcription systems** *Genetics* **196**:951–960 <https://doi.org/10.1534/GENETICS.113.161141/-/DC1/GENETICS.113.161141-5.PDF>
- Smith AM, Heisler LE, Mellor J, Kaper F, Thompson MJ, Chee M, Roth FP, Giaever G, Nislow C (2009) **Quantitative phenotyping via deep barcode sequencing** *Genome research* **19**:1836–42 <https://doi.org/10.1101/gr.093955.109>
- Stoeckius M, Hafemeister C, Stephenson W, Houck-Loomis B, Chattopadhyay PK, Swerdlow H, Satija R, Smibert P (2017) **Simultaneous epitope and transcriptome measurement in single cells** *Nature Methods* **14**:865–868 <https://doi.org/10.1038/nmeth.4380>
- Stoeckius M, Zheng S, Houck-Loomis B, Hao S, Yeung BZ, Mauck WM, Smibert P, Satija R. (2018) **Cell Hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics** *Genome Biology* **19**:1–12 <https://doi.org/10.1186/s13059-018-1603-1>
- Teleman AA, Cohen SM (2000) **Dpp gradient formation in the drosophila wing imaginal disc** *Cell* **103**:971–980 [https://doi.org/10.1016/S0092-8674\(00\)00199-9](https://doi.org/10.1016/S0092-8674(00)00199-9)
- Thompson CL *et al.* (2014) **A High-Resolution Spatiotemporal Atlas of Gene Expression of the Developing Mouse Brain** *Neuron* **83**:309–323 <https://doi.org/10.1016/j.NEURON.2014.05.033>
- van Opijnen T, Bodi KL, Camilli A. (2009) **Tn-seq: high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms** *Nature methods* **6**:767–72 <https://doi.org/10.1038/nmeth.1377>
- Venken KJT *et al.* (2011) **MiMIC: a highly versatile transposon insertion resource for engineering Drosophila melanogaster genes** *Nature methods* **8**:737–43
- Yang S, Corbett SE, Koga Y, Wang Z, Johnson WE, Yajima M, Campbell JD (2020) **Decontamination of ambient RNA in single-cell RNA-seq with DecontX** *Genome Biology* **21** <https://doi.org/10.1186/s13059-020-1950-6>
- Yao Z, Scott K (2022) **Serotonergic neurons translate taste detection into internal nutrient regulation** *Neuron* **110**:1036–1050 <https://doi.org/10.1016/j.neuron.2021.12.028>
- Zappia L, Oshlack A (2018) **Clustering trees: a visualization for evaluating clusterings at multiple resolutions** *GigaScience* **7** <https://doi.org/10.1093/gigascience/giy083>
- Zheng GXY *et al.* (2017) **Massively parallel digital transcriptional profiling of single cells** *Nature Communications* **8**:1–12 <https://doi.org/10.1038/ncomms14049>
- Zirin J, Jusiak B, Lopes R, Ewen-Campen B, Bosch JA, Risbeck A, Forman C, Villalta C, Hu Y, Perrimon N (2024) **Expanding the Drosophila toolkit for dual control of gene expression** *eLife* **12** <https://doi.org/10.7554/eLife.94073>

Author information

Jorge Blanco Mendana

University of Minnesota Genomics Center, Minneapolis, USA

Margaret Donovan

University of Minnesota Genomics Center, Minneapolis, USA

Lindsey G O'Brien

University of Minnesota Genomics Center, Minneapolis, USA

Benjamin Auch[†]

University of Minnesota Genomics Center, Minneapolis, USA

ORCID iD: [0000-0003-3125-5797](https://orcid.org/0000-0003-3125-5797)

[†]Present address: Phase Genomics, Seattle, USA

John Garbe

University of Minnesota Genomics Center, Minneapolis, USA

Daryl M Gohl

University of Minnesota Genomics Center, Minneapolis, USA, Department of Genetics, Cell Biology, and Development, University of Minnesota, Minneapolis, USA

ORCID iD: [0000-0002-4434-2788](https://orcid.org/0000-0002-4434-2788)

For correspondence: dmgoehl@umn.edu

Editors

Reviewing Editor

Sonia Sen

Tata Institute for Genetics and Society, Bangalore, India

Senior Editor

Claude Desplan

New York University, New York, United States of America

Reviewer #1 (Public review):

The aim of this paper is to describe a novel method for genetic labelling of animals or cell populations, using a system of DNA/RNA barcodes.

Strengths:

- The author's attempt at providing a straightforward method for multiplexing *Drosophila* samples prior to scRNA-seq is commendable. The perspective of being able to load multiple samples on a 10X Chromium without antibody labelling is appealing.
- The authors are generally honest about potential issues in their method, and areas that would benefit from future improvement.
- The article reads well. Graphs and figures are clear and easy to understand.

Weaknesses:

- The usefulness of TaG-EM for phototaxis, egg laying or fecundity experiments is questionable. The behaviours presented here are all easily quantifiable, either manually or using automated image-based quantification, even when they include a relatively large number of groups and replicates. Despite their claims (e.g., L311-313), the authors do not

present any real evidence about the cost- or time-effectiveness of their method in comparison to existing quantification methods.

- Behavioural assays presented in this article have clear outcomes, with large effect sizes, and therefore do not really challenge the efficiency of TaG-EM. By showing a T-maze in Fig 1B, the authors suggest that their method could be used to quantify more complex behaviours. Not exploring this possibility in this manuscript seems like a missed opportunity.
- Experiments in Figs S3 and S6 suggest that some tags have a detrimental effect on certain behaviours or on GFP expression. Whereas the authors rightly acknowledge these issues, they do not investigate their causes. Unfortunately, this question the overall suitability of TaG-EM, as other barcodes may also affect certain aspects of the animal's physiology or behaviour. Revising barcode design will be crucial to make sure that sequences with potential regulatory function are excluded.
- For their single-cell experiments, the authors have used the 10X Genomics method, which relies on sequencing just a short segment of each transcript (usually 50-250bp - unknown for this study as read length information was not provided) to enable its identification, with the matching paired-end read providing cell barcode and UMI information (Macosko et al., 2015). With average fragment length after tagmentation usually ranging from 300-700bp, a large number of GFP reads will likely not include the 14bp TaG-EM barcode. When a given cell barcode is not associated with any TaG-EM barcode, then demultiplexing is impossible. This is a major problem, which is particularly visible in Figs 5 and S13. In 5F, BC4 is only detected in a couple of dozen cells, even though the *Jon99Ciii* marker of enterocytes is present in a much larger population (Fig 5C). Therefore, in this particular case, TaG-EM fails to detect most of the GFP-expressing cells. Similarly, in S13, most cells should express one of the four barcodes, however many of them (maybe up to half - this should be quantified) do not. Therefore, the claim (L277-278) that "the pan-midgut driver were broadly distributed across the cell clusters" is misleading. Moreover, the hypothesis that "low expressing driver lines may result in particularly sparse labelling" (L331-333) is at least partially wrong, as Fig S13 shows that the same Gal4 driver can lead to very different levels of barcode coverage.
- Comparisons between TaG-EM and other, simpler methods for labelling individual cell populations are missing. For example, how would TaG-EM compare with expression of different fluorescent reporters, or a strategy based on the brainbow/flybow principle?
- FACS data is missing throughout the paper. The authors should include data from their comparative flow cytometry experiment of TaG-EM cells with or without additional hexameric GFP, as well as FSC/SSC and fluorescence scatter plots for the FACS steps that they performed prior to scRNA-seq, at least in supplementary figures.
- The authors should show the whole data described in L229, including the cluster that they chose to delete. At least, they should provide more information about how many cells were removed. In any case, the fact that their data still contains a large number of debris and dead cells despite sorting out PI negative cells with FACS and filtering low abundance barcodes with Cellranger is concerning.

Overall, although a method for genetic tagging cell populations prior to multiplexing in single-cell experiments would be extremely useful, the method presented here is inadequate. However, despite all the weaknesses listed above, the idea of barcodes expressed specifically in cells of interest deserves more consideration. If the authors manage to improve their design to resolve the major issues and demonstrate the benefits of their method more clearly, then TaG-EM could become an interesting option for certain applications.

Comments on revisions:

The authors have addressed many important points, providing reassurances about the initial weaknesses of their work. Although the TaG-EM is unlikely to have a significant influence on the field due to its limited benefits, the results are now sound and provide the reader with an unbiased view of the possibilities and limitations of the method.

Reviewer #2 (Public review):

The authors developed the TaG-EM system to address challenges in multiplexing *Drosophila* samples for behavioral and transcriptomic studies. This system integrates DNA barcodes upstream of the polyadenylation site in a UAS-GFP construct, enabling pooled behavioral measurements and cell type tracking in scRNA-seq experiments. The revised manuscript expands on the utility of TaG-EM by demonstrating its application to complex assays, such as larval gut motility, and provides a refined analysis of its limitations and cost-effectiveness.

Strengths

- (1) Novelty and Scope: The study demonstrates the potential for TaG-EM to streamline multiplexing in both behavioral and transcriptomic contexts. The additional application to labor-intensive larval gut motility assays highlights its scalability and practical utility.
- (2) Data Quality and Clarity: Figures and supplemental data are mostly clear and significantly enhanced in the revised manuscript. The addition of Supplemental Figures 18-21 addresses initial concerns about scRNA-seq data and driver characterization.
- (3) Cost-Effectiveness Analysis: New analyses of labor and cost savings (e.g., Supplemental Figure 8) provide a practical perspective.
- (4) Improvements in Barcode Detection and Analysis: Enhanced enrichment protocols (Supplemental Figures 18-19) demonstrate progress in addressing limitations of barcode detection and increase the detection rate of labeled cells.

Weaknesses

- (1) Barcode Detection Efficiency: While improvements are noted, the low barcode detection rate (~37% in optimized conditions) limits the method's scalability in some applications, such as single-cell sequencing experiments with complex cell populations.
- (2) Sparse Labeling: Sparse labeling of cell populations, particularly in scRNA-seq assays, remains a concern. Variability in driver strength and regional expression introduces inconsistencies in labeling density.
- (3) Behavioral Applications: The utility of TaG-EM in quantifying more complex behaviors remains underexplored, limiting the generalizability of the method beyond simpler assays like phototaxis and oviposition.
- (4) Driver Line Characterization: While improvements in driver line characterization were made, variability in expression patterns and sparse labeling emphasize the need for further refinement of constructs and systematic backcrossing to standardize the genetic background.

<https://doi.org/10.7554/eLife.88334.2.sa1>

Author response:

The following is the authors' response to the original reviews.

Public Reviews:

Reviewer #1 (Public Review):

The aim of this paper is to describe a novel method for genetic labelling of animals or cell populations, using a system of DNA/RNA barcodes.

Strengths:

- *The author's attempt at providing a straightforward method for multiplexing Drosophila samples prior to scRNA-seq is commendable. The perspective of being able to load multiple samples on a 10X Chromium without antibody labelling is appealing.*
- *The authors are generally honest about potential issues in their method, and areas that would benefit from future improvement.*
- *The article reads well. Graphs and figures are clear and easy to understand.*

We thank the reviewer for these positive comments.

Weaknesses:

- *The usefulness of TaG-EM for phototaxis, egg laying or fecundity experiments is questionable. The behaviours presented here are all easily quantifiable, either manually or using automated image-based quantification, even when they include a relatively large number of groups and replicates. Despite their claims (e.g., L311-313), the authors do not present any real evidence about the cost- or time-effectiveness of their method in comparison to existing quantification methods.*

While the behaviors that were quantified in the original manuscript were indeed relatively easy to quantify through other methods, they nonetheless demonstrated that sequencing-based TaG-EM measurements faithfully recapitulated manual behavioral measurements. In response to the reviewer's comment, we have added additional experiments that demonstrate the utility of TaG-EM-based behavioral quantification in the context of a more labor-intensive phenotypic assay (measuring gut motility via food transit times in *Drosophila* larvae, Figure 4, Supplemental Figure 7). We found that food transit times in the presence and absence of caffeine are subtly different and that, as with larger effect size behaviors, TaG-EM data recapitulates the results of the manual assay. This experiment demonstrates both that TaG-EM can be used to streamline labor-intensive behavioral assays (we have included an estimate of the savings in hands-on labor for this assay by using a multiplexed sequencing approach, Supplemental Figure 8) and that TaG-EM can quantify small differences between experimental groups. We also note in the discussion that an additional benefit of TaGEM-based behavioral assays is that the observed is blinded as to the experimental conditions as they are intermingled in a single multiplexed assay. We have added the following text to the paper describing these experiments.

Results:

“Quantifying food transit time in the larval gut using TaG-EM

Gut motility defects underlie a number of functional gastrointestinal disorders in humans (Keller et al., 2018). To study gut motility in *Drosophila*, we have developed an assay based on the time it takes a food bolus to transit the larval gut (Figure 4A), similar to approaches that have been employed for studying the role of the microbiome in human gut motility (Asnicar et al., 2021). Third instar larvae were starved for 90 minutes and then fed food containing a blue dye. After 60 minutes, larvae in which a blue bolus of food was visible were transferred to plates containing non-dyed food, and food transit (indicated by loss of the blue food bolus) was scored every 30 minutes for five hours (Supplemental Figure 7).

Because this assay is highly labor-intensive and requires hands-on effort for the entire five-hour observation period, there is a limit on how many conditions or replicates can be scored in one session (~8 plates maximum). Thus, we decided to test whether food transit could be quantified in a more streamlined and scalable fashion by using TaG-EM (Figure 4B). Using the manual assay, we observed that while caffeine-containing food is aversive to larvae, the presence of caffeine reduces transit time through the gut (Figure 4C, Supplemental Figure 7). This is consistent with previous observations in adult flies that bitter compounds (including caffeine) activate enteric neurons via serotonin-mediated signaling and promote gut motility (Yao and Scott, 2022). We tested whether TaG-EM could be used to measure the effect of caffeine on food transit time in larvae. As with prior behavioral tests, the TaG-EM data recapitulated the results seen in the manual assay (Figure 4D). Conducting the transit assay via TaGEM enables several labor-saving steps. First, rather than counting the number of larvae with and without a food bolus at each time point, one simply needs to transfer nonbolus-containing larvae to a collection tube. Second, because the TaG-EM lines are genetically barcoded, all the conditions can be tested at once on a single plate, removing the need to separately count each replicate of each experimental condition. This reduces the hands-on time for the assay to just a few minutes per hour. A summary of the anticipated cost and labor savings for the TaG-EM-based food transit assay is shown in Supplemental Figure 8.”

Discussion:

“While the utility of TaG-EM barcode-based quantification will vary based on the number of conditions being analyzed and the ease of quantifying the behavior or phenotype by other means, we demonstrate that TaG-EM can be employed to cost-effectively streamline labor-intensive assays and to quantify phenotypes with small effect sizes (Figure 4, Supplemental Figure 8). An additional benefit of multiplexed TaG-EM behavioral measurements is that the experimental conditions are effectively blinded as the multiplexed conditions are intermingled in a single assay.”

Methods:

“Larval gut motility experiments

Preparing Yeast Food Plates

Yeast agar plates were prepared by making a solution containing 20% Red Star Active Dry Yeast 32oz (Red Star Yeast) and 2.4% Agar Powder/Flakes (Fisher) and a separate solution containing 20% Glucose (Sigma-Aldrich). Both mixtures were autoclaved with a 45-minute liquid cycle and then transferred to a water bath at 55°C. After cooling to 55°C, the solutions were combined and mixed, and approximately 5 mL of the combined solution was transferred into 100 x 15 mm petri dishes (VWR) in a PCR hood or contamination-free area. For blue-dyed yeast food plates, 0.4% Blue Food Color (McCormick) was added to the yeast solution. For the caffeine assays, 300 μ L of a solution of 100 mM 99% pure caffeine (Sigma-Aldrich) was pipetted onto the blue-dyed yeast plate and allowed to absorb into the food during the 90-minute starvation period.

Manual Gut Motility Assay

Third instar *Drosophila* larvae were transferred to empty conical tubes that had been misted with water to prevent the larvae from drying out. After a 90-minute starvation period the larvae were moved from the conical to a blue-dyed yeast plate with or without caffeine and allowed to feed for 60 minutes. Following the feeding period, the larvae were transferred to an undyed yeast plate. Larvae were scored for the presence or absence of a food bolus every 30 minutes over a 5-hour period. Up to 8 experimental replicates/conditions were scored simultaneously.

TaG-EM Gut Motility Assay

Third instar larvae were starved and fed blue dye-containing food with or without caffeine as described above. An equal number of larvae from each experimental condition/replicate were transferred to an undyed yeast plate. During the 5-hour observation period, larvae were examined every 30 minutes and larvae lacking a food bolus were transferred to a microcentrifuge tube labeled for the timepoint. Any larvae that died during the experiment were placed in a separate microcentrifuge tube and any larvae that failed to pass the food bolus were transferred to a microcentrifuge tube at the end of the experiment. DNA was extracted from the larvae in each tube and TaG-EM barcode libraries were prepared and sequenced as described above.”

• *Behavioural assays presented in this article have clear outcomes, with large effect sizes, and therefore do not really challenge the efficiency of TaG-EM. By showing a Tmaze in Fig 1B, the authors suggest that their method could be used to quantify more complex behaviours. Not exploring this possibility in this manuscript seems like a missed opportunity.*

See the response to the previous point.

• *Experiments in Figs S3 and S6 suggest that some tags have a detrimental effect on certain behaviours or on GFP expression. Whereas the authors rightly acknowledge these issues, they do not investigate their causes. Unfortunately, this question the overall suitability of TaG-EM, as other barcodes may also affect certain aspects of the animal's physiology or behaviour. Revising barcode design will be crucial to make sure that sequences with potential regulatory function are excluded.*

We have determined that the barcode (BC#8) that had no detectable Gal4induced gene expression in Figure S6 (now Supplemental Figure 9) has a deletion in the GFP coding region that ablates GFP function. Interestingly, the expressed TaG-EM barcode transcript is still detectable in single cell sequencing experiments, but obviously this line cannot be used for cell enrichment (at least based solely on GFP expression from the TaG-EM construct). While it is unclear how this line came to have a lesion in the GFP gene, we have subsequently generated >150 additional TaG-EM stocks and we have tested the GFP expression of these newly established stocks by crossing them to *Mhc-Gal4*. All of the additional stocks had GFP expression in the expected pattern, indicating that the BC#8 construct is an outlier with respect to inducibility of GFP. We have added the following text to the results section to address this point:

“No GFP expression was visible for TaG-EM barcode number 8, which upon molecular characterization had an 853 bp deletion within the GFP coding region (data not shown). We generated and tested GFP expression of an additional 156 TaG-EM barcode lines (Alegria et al., 2024), by crossing them to *Mhc-Gal4* and observing expression in the adult thorax. All 156 additional TaG-EM lines had robust GFP expression (data not shown).”

It is certainly the case that future improvements to the construct design may be necessary or desirable and that back-crossing could likely be used to alleviate line-to-line differences for specific phenotypes, we also address this point in the discussion with the following text:

“We excluded this poor performing barcode line from the fecundity tests, however, backcrossing is often used to bring reagents into a consistent genetic background for behavioral experiments and could also potentially be used to address behavior-specific issues with specific TaG-EM lines. In addition, other strategies such as averaging across multiple barcode lines or permutation of barcode assignment across replicates could also mitigate such deficiencies.”

• For their single-cell experiments, the authors have used the 10x Genomics method, which relies on sequencing just a short segment of each transcript (usually 50-250bp - unknown for this study as read length information was not provided) to enable its identification, with the matching paired-end read providing cell barcode and UMI information (Macosko et al., 2015). With average fragment length after tagmentation usually ranging from 300-700bp, a large number of GFP reads will likely not include the 14bp TaG-EM barcode.

The 10x Genomics 3' workflows that were used for sequencing TaG-EM samples reads the cell barcode and UMI in read one and the expressed RNA sequence in read two. We sequenced the samples shown in Figure 5 in the initial manuscript using a run configuration that generated 150 bp for read two. The TaG-EM barcodes are located just upstream of the polyadenylation sites (based on the sequencing data, we observe two different poly-A sites and the TaG-EM barcode is located 35 and 60 bp upstream of these sites). Based on the location of the TaG-EM barcodes, 150 bp reads is sufficient to see the barcode in any GFP-associated read (when using the 3' gene expression workflow). In addition to detecting the expression of the TaG-EM barcodes in the 10x Genomics gene expression library, it is possible to make a separate library that enriches the barcode sequence (similar to hashtag or CITE-Seq feature barcode libraries). We have added experimental data where we successfully performed an enrichment of the TaG-EM barcodes and sequenced this as a separate hashtag library (Supplemental Figure 18). We have added text to the results describing this work and also included a detailed information in the methods for performing TaG-EM barcode enrichment during 10x library prep.

Results:

“In antibody-conjugated oligo cell hashing approaches, sparsity of barcode representation is overcome by spiking in an additional primer at the cDNA amplification step and amplifying the hashtag oligo by PCR. We employed a similar approach to attempt to enrich for TaG-EM barcodes in an additional library sequenced separately from the 10x Genomics gene expression library. Our initial attempts at barcode enrichment using spike-in and enrichment primers corresponding to the TaG-EM PCR handle were unsuccessful (Supplemental Figure 18). However, we subsequently optimized the TaG-EM barcode enrichment by 1) using a longer spike-in primer that more closely matches the annealing temperature used during the 10x Genomics cDNA creation step, and 2) using a nested PCR approach to amplify the cell-barcode and unique molecular identifier (UMI)-labeled TaG-EM barcodes (Supplemental Figure 18). Using the enriched library, TaG-EM barcodes were detected in nearly 100% of the cells at high sequencing depths (Supplemental Figure 19). However, although we used a polymerase that has been engineered to have high processivity and that has been shown to reduce the formation of chimeric reads in other contexts (Gohl et al., 2016), it is possible that PCR chimeras could lead to unreliable detection events for some cells. Indeed, many cells had a mixture of barcodes detected with low counts and single or low numbers of associated UMIs. To assess the reliability of detection, we analyzed the correlation between barcodes detected in the gene expression library and the enriched TaG-EM barcode library as a function of the purity of TaG-EM barcode detection for each cell (the percentage of the most abundant detected TaG-EM barcode, Supplemental Figure 19). For TaG-EM barcode detections where the most abundance barcode was a high percentage of the total barcode reads detected (~75%-99.99%), there was a high correlation between the barcode detected in the gene expression library and the enriched TaG-EM barcode library. Below this threshold, the correlation was substantially reduced.

In the enriched library, we identified 26.8% of cells with a TaG-EM barcode reliably detected, a very modest improvement over the gene expression library alone (23.96%), indicating that at least for this experiment, the main constraint is sufficient expression of the TaG-EM

barcode and not detection. To identify TaG-EM barcodes in the combined data set, we counted a positive detection as any barcode either identified in the gene expression library or any barcode identified in the enriched library with a purity of >75%. In the case of conflicting barcode calls, we assigned the barcode that was detected directly in the gene expression library. This increased the total fraction of cells where a barcode was identified to approximately 37% (Figure 6B)."

Methods:

"The resulting pool was prepared for sequencing following the 10x Genomics Single Cell 3' protocol (version CG000315 Rev C), At step 2.2 of the protocol, cDNA amplification, 1 µl of TaG-EM spike-in primer (10 µM) was added to the reaction to amplify cDNA with the TaG-EM barcode. Gene expression cDNA and TaG-EM cDNA were separated using a double-sided SPRIselect (Beckman Coulter) bead clean up following 10x Genomics Single Cell 3' Feature Barcode protocol, step 2.3 (version CG000317 Rev E). The gene expression cDNA was created into a library following the CG000315 Rev C protocol starting at section 3. Custom nested primers were used for enrichment of TaG-EM barcodes after cDNA creation using PCR. The following primers were tested (see Supplemental Figure 18):

UMGC_IL_TaGEM_SpikeIn_v1:

GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCTTCCAACAACCGGAAGT*G*A

UMGC_IL_TaGEM_SpikeIn_v2:

GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGCAGCTTATAACTTCCAACAACCGGAAGT*G*A

UMGC_IL_TaGEM_SpikeIn_v3:

TGTGCTCTTCCGATCTGCAGCTTATAACTTCCAACAACCGGAAGT*G*A D701_TaGEM:

CAAGCAGAAGACGGCATACGAGATCGAGTAATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGCAGC*T*T

SI PCR Primer:

AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGC*T*C

UMGC_IL_DoubleNest:

GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGCAGCTTATAACTTCCAACAACCGG*A*A

P5: AATGATACGGCGACCACCGA

D701:

GATCGGAAGAGCACACGTCTGAACTCCAGTCACATTACTCGATCTCGTATGCCGTCTTCTGCTTG

D702:

GATCGGAAGAGCACACGTCTGAACTCCAGTCACTCCGAGAATCTCGTATGCCGTCTTCTGCTTG

After multiple optimization trials, the following steps yielded ~96% on-target reads for the TaG-EM library (Supplemental Figure 18, note that for the enriched barcode data shown in Figure 6 and Supplemental Figure 19, a similar amplification protocol was used TaG-EM barcodes were amplified from the gene expression library cDNA and not the SPRI-selected barcode pool). TaG-EM cDNA was amplified with the following PCR reaction: 5 µl purified TaG-EM cDNA, 50 µl 2x KAPA HiFi ReadyMix (Roche), 2.5 µl UMGc_IL_DoubleNest primer (10 µM), 2.5 µl SI_PCR primer (10 µM), and 40 µl nuclease-free water. The reaction was amplified using the following cycling conditions: 98°C for 2 minutes, followed by 15 cycles of 98°C for 20 seconds, 63°C for 30 seconds, 72°C for 20 seconds, followed by 72°C for 5 minutes. After the

first PCR, the amplified cDNA was purified with a 1.2x SPRIselect (Beckman Coulter) bead cleanup with 80% ethanol washes and eluted into 40 μ L of nuclease-water. A second round of PCR was run with following reaction: 5 μ L purified TaG-EM cDNA, 50 μ L 2x KAPA HiFi ReadyMix (Roche), 2.5 μ L D702 primer (10 μ M), 2.5 μ L p5 Primer (10 μ M), and 40 μ L nuclease-free water. The reaction was amplified using the following cycling conditions: 98°C for 2 minutes, followed by 10 cycles of 98°C for 20 seconds, 63°C for 30 seconds, 72°C for 20 seconds, followed by 72°C for 5 minutes. After the second PCR, the amplified cDNA was purified with a 1.2x SPRIselect (Beckman Coulter) bead cleanup with 80% ethanol washes and eluted into 40 μ L of nuclease-water. The resulting 3' gene expression library and TaG-EM enrichment library were sequenced together following Scenario 1 of the BioLegend "Total-Seq-A Antibodies and Cell Hashing with 10x Single Cell 3' Reagents Kit v3 or v3.1" protocol. Additional sequencing of the enriched TaG-EM library also done following Scenario 2 from the same protocol."

When a given cell barcode is not associated with any TaG-EM barcode, then demultiplexing is impossible. This is a major problem, which is particularly visible in Figs 5 and S13. In 5F, BC4 is only detected in a couple of dozen cells, even though the Jon99Ciii marker of enterocytes is present in a much larger population (Fig 5C). Therefore, in this particular case, TaG-EM fails to detect most of the GFP-expressing cells.

Figure 5 in the original manuscript represented data from an experiment in which there were eight different TaG-EM barcoded samples present, including four replicates of the pan-midgut driver (each of which included enterocyte populations). One would not expect the BC4 enterocyte driver expression to be observed in all of the Jon99Ciii cells, since the majority of the GFP+ cells shown in the UMAP plot were likely derived from and are labeled by the pan-midgut driver-associated barcodes. Thus, the design and presentation of this particular experiment (in particular, the presence of eight distinct samples in the data set) is making the detection of the TaG-EM barcodes look sparser than it actually is. We have added a panel in both Figure 6B and Supplemental Figure 17B that shows the overall detection of barcodes in the enriched barcode library and gene expression library or the gene expression library only, respectively, for this experiment.

However, the reviewer's overall point regarding barcode detection is still valid in that if we consider all eight barcodes, we only see TaG-EM barcode labeling associated with about a quarter of all the cells in this gene expression library, or about 37% of cells when we include the enriched TaG-EM barcode library. While improving barcode detection will improve the yield and is necessary for some applications (such as robust detection of multiplets), we would argue that even at the current level of success this approach has significant utility. First, if one's goal is to unambiguously label a cell cluster and trace it to a defined cell population *in vivo*, sparse labeling may be sufficient. Second, demultiplexing is still possible (as we demonstrate) but involves a trade off in yield (not every cell is recovered and there is some extra sequencing cost as some sequenced cells cannot be assigned to a barcode).

Similarly, in S13, most cells should express one of the four barcodes, however many of them (maybe up to half - this should be quantified) do not. Therefore, the claim (L277278) that "the pan-midgut driver were broadly distributed across the cell clusters" is misleading. Moreover, the hypothesis that "low expressing driver lines may result in particularly sparse labelling" (L331-333) is at least partially wrong, as Fig S13 shows that the same Gal4 driver can lead to very different levels of barcode coverage.

As described above, since this experiment included eight different TaG-EM barcodes expressed by five different drivers, the expectation is that only about half of the cells in Figure S13 (now Figure S20) should express a TaG-EM barcode. It is not clear why BC2 is underrepresented in terms of the number of cells labeled and BC7 is overrepresented. We

agree with the reviewer that this should be described more accurately in the paper and that it does impact our interpretation related to driver strength and barcode detection. We have revised this sentence in the discussion and also added additional text in the results describing the within driver variability seen in this experiment.

Results text:

“As expected, the barcodes expressed by the pan-midgut driver were broadly distributed across the cell clusters (Supplemental Figure 20). However, the number of cells recovered varied significantly among the four pan-midgut driver associated barcodes.”

Discussion text:

“It is likely that the strength of the Gal4 driver contributes to the labeling density. However, we also observed variable recovery of TaG-EM barcodes that were all driven by the same pan-midgut Gal4 driver (Supplemental Figure 20).”

- *Comparisons between TaG-EM and other, simpler methods for labelling individual cell populations are missing. For example, how would TaG-EM compare with expression of different fluorescent reporters, or a strategy based on the brainbow/flybow principle?*

The advantage of TaG-EM is that an arbitrarily large number of DNA barcodes can be used (contingent upon the availability of transgenic lines – we described 20 barcoded lines in our initial manuscript and we have now extended this collection to over 170 lines), while the number of distinguishable FPs is much lower. Brainbow/Flybow uses combinatorial expression of different FPs, but because this combinatorial expression is stochastic, tracing a single cell transcriptome to a defined cell population in vivo based on the FP signature of a Brainbow animal would likely not be possible (and would almost certainly be impossible at scale).

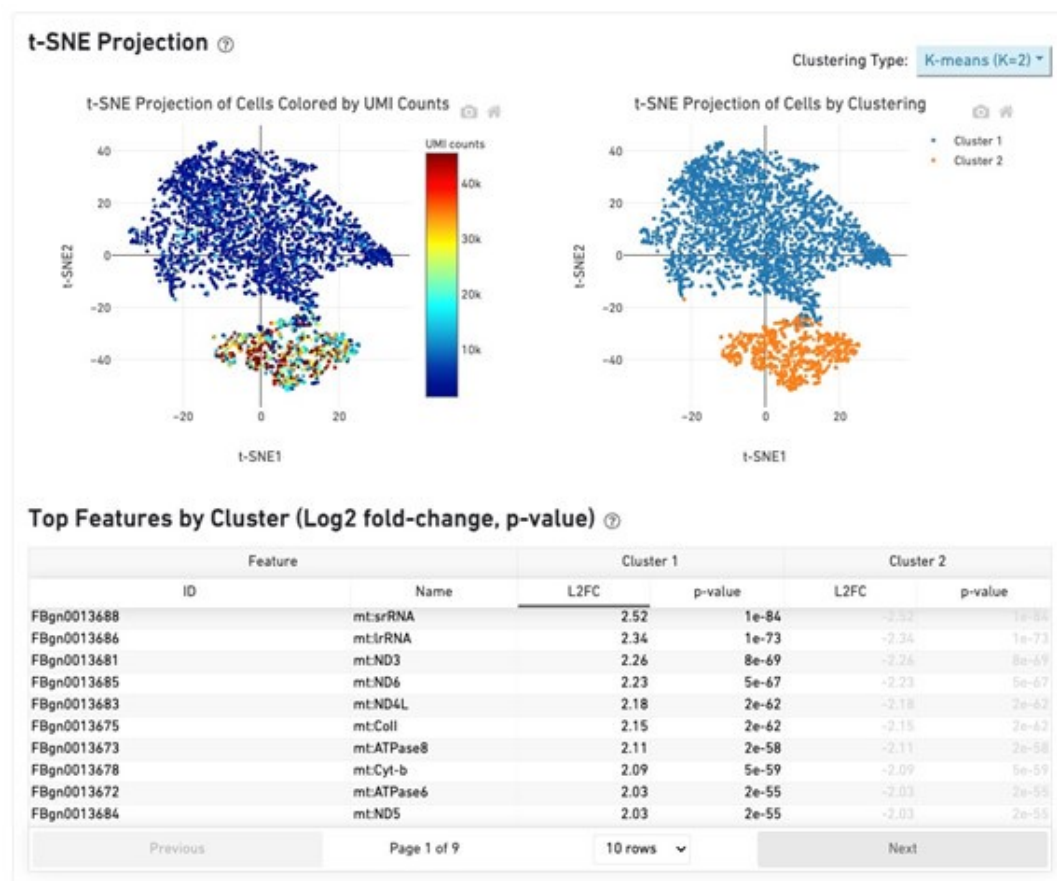
- *FACS data is missing throughout the paper. The authors should include data from their comparative flow cytometry experiment of TaG-EM cells with or without additional hexameric GFP, as well as FSC/SSC and fluorescence scatter plots for the FACS steps that they performed prior to scRNA-seq, at least in supplementary figures.*

We have added Supplemental Figures with the FACS data for all of the single cell sequencing data presented in the manuscript (Supplemental Figures 12 and 14).

- *The authors should show the whole data described in L229, including the cluster that they chose to delete. At least, they should provide more information about how many cells were removed. In any case, the fact that their data still contains a large number of debris and dead cells despite sorting out PI negative cells with FACS and filtering low abundance barcodes with Cellranger is concerning.*

This description was referring to the unprocessed Cellranger output (not filtered for low abundance barcodes). Prior to filtering for cell barcodes with high mitochondria or rRNA (or other processing in Seurat/Scanpy), we saw two clusters, one with low UMI counts and enrichment of mitochondrial genes (see Cellranger report below).

Author response image 1.



These cell barcodes were removed by downstream quality filtering and the remaining cells showed expression of expected intestinal stem cell and enteroblast marker genes.

Overall, although a method for genetic tagging cell populations prior to multiplexing in single-cell experiments would be extremely useful, the method presented here is inadequate. However, despite all the weaknesses listed above, the idea of barcodes expressed specifically in cells of interest deserves more consideration. If the authors manage to improve their design to resolve the major issues and demonstrate the benefits of their method more clearly, then TaG-EM could become an interesting option for certain applications.

We thank the reviewer for this comment and hope that the above responses and additional experiments and data that we have added have helped to alleviate the noted weaknesses.

Reviewer #2 (Public Review):

In this manuscript, Mendana et al developed a multiplexing method - Targeted Genetically-Encoded Multiplexing or TaG-EM - by inserting a DNA barcode upstream of the polyadenylation site in a Gal4-inducible UAS-GFP construct. This Multiplexing method can be used for population-scale behavioral measurements or can potentially be used in single-cell sequencing experiments to pool flies from different populations. The authors created 20 distinctly barcoded fly lines. First, TaG-EM was used to measure phototaxis and oviposition behaviors. Then, TaG-EM was applied to the fly gut cell types to

demonstrate its applications in single-cell RNA-seq for cell type annotation and cell origin retrieving.

This TaG-EM system can be useful for multiplexed behavioral studies from nextgeneration sequencing (NGS) of pooled samples and for Transcriptomic Studies. I don't have major concerns for the first application, but I think the scRNA-seq part has several major issues and needs to be further optimized.

Major concerns:

(1) It seems the barcode detection rate is low according to Fig S9 and Fig 5F, J and N. Could the authors evaluate the detection rate? If the detection rate is too low, it can cause problems when it is used to decode cell types.

See responses to Reviewer #1 on this topic above.

(2) Unsuccessful amplification of TaG-EM barcodes: The authors attempted to amplify the TaG-EM barcodes in parallel to the gene expression library preparation but encountered difficulties, as the resulting sequencing reads were predominantly offtarget. This unsuccessful amplification raises concerns about the reliability and feasibility of this amplification approach, which could affect the detection and analysis of the TaG-EM barcodes in future experiments.

As noted above, we have now established a successful amplification protocol for the TaG-EM barcodes. This data is shown in Figure 6, and Supplemental Figures 18-19 and we have included a detailed information in the methods for performing TaG-EM barcode enrichment during 10x library prep. We have also included code in the paper's Github repository for assigning TaG-EM barcodes from the enriched library to the associated 10x Genomics cell barcodes.

(3) For Fig 5, the single-cell clusters are not annotated. It is not clear what cell types are corresponding to which clusters. So, it is difficult to evaluate the accuracy of the assignment of barcodes.

We have added annotation information for the cell clusters based on expression of cell-type-specific marker genes (Figure 6A, Supplemental Figures 16-17).

(4) The scRNA-seq UMAP in Fig 5 is a bit strange to me. The fly gut epithelium contains only a few major cell types, including ISC, EB, EC, and EE. However, the authors showed 38 clusters in fig 5B. It is true that some cell types, like EE (Guo et al., 2019, Cell Reports), have sub-populations, but I don't expect they will form these many subtypes. There are many peripheral small clusters that are not shown in other gut scRNAseq studies (Hung et al., 2020; Li et al., 2022 Fly Cell Atlas; Lu et al., 2023 Aging Fly Cell Atlas). I suggest the authors try different data-processing methods to validate their clustering result.

For all of the single cell experiments, after doublet and ambient RNA removal (as suggested below), we have reclustered the datasets and evaluated different resolutions using Clustree. As the Reviewer points out, there are different EE subtypes, as well as regionalized expression differences in EC and other cell populations, so more than four clusters are expected (an analysis of the adult midgut identified 22 distinct cell types). With this revised analysis our results more closely match the cell populations observed in other studies (though it should be noted that the referenced studies largely focus on the adult and not the larval stage).

(5) Different gut drivers, *PMC*-, *PC*-, *EB*-, *EC*-, and *EE-GAL4*, were used. The authors should carefully characterize these *GAL4* expression in larval guts and validate sequencing data. For example, does the ratio of each cell type in Fig 5B reflect the *in vivo* cell type ratio? The authors used cell-type markers mostly based on the knowledge from adult guts, but there are significant morphological and cell ratio differences between larval and adult guts (e.g., Mathur...Ohlstein, 2010 Science).

We have characterized the *PC* driver which is highlighted in Supplemental Figure 13, and the *EC* and *EE* drivers which are highlighted in Figure 6G-N in detail in larval guts and have added this data to the paper (Supplemental Figure 21). The *EB* driver was not characterized histologically as *EB*-specific antibodies are not currently available. The *PMG-Gal4* line exhibits strong expression throughout the larval gut (Figure 5B and barcodes are recovered from essentially all of the larval gut cell clusters using this driver (Supplemental Figure 20). We don't necessarily expect the ratios of cells observed in the scRNA-Seq data to reflect the ratios typically observed in the gut as we performed pooled flow sorting on a multiplexed set of eight genotypes and driver expression levels, flow sorting, and possibly other processing steps could all influence the relative abundance of different cell types. However, detailed characterization of these driver lines did reveal spatial expression patterns that help explain aspects of the scRNA-Seq data. We have also added the following text to the paper to further describe the characterization of the drivers:

Results:

"Detailed characterization of the *EC-Gal4* line indicated that although this line labeled a high percentage of enterocytes, expression was restricted to an area at the anterior and middle of the midgut, with gaps between these regions and at the posterior (Supplemental Figure 21). This could explain the absence of subsets of enterocytes, such as those labeled by *betaTry*, which exhibits regional expression in R2 of the adult midgut (Buchon et al., 2013)."

"Detailed characterization of the *EE-Gal4* driver line indicated that ~80-85% of Prospero-positive enteroendocrine cells are labeled in the anterior and middle of the larval midgut, with a lower percentage (~65%) of Prospero-positive cells labeled in the posterior midgut (Supplemental Figure 21). As with the enterocyte labeling, and consistent with the *Gal4* driver expression pattern, the *EE-Gal4* expressed TaG-EM barcode 9 did not label all classes of enteroendocrine cells and other clusters of presumptive enteroendocrine cells expressing other neuropeptides such as *Orcokinin*, *AstA*, and *AstC*, or neuropeptide receptors such as *CCHa2* (not shown) were also observed."

Methods:

"Dissection and immunostaining

Midguts from third instar larvae of driver lines crossed to UAS-GFP.nls or UAS-mCherry were dissected in 1xPBS and fixed with 4% paraformaldehyde (PFA) overnight at 4°C. Fixed samples were washed with 0.1% PBTx (1xPBS + 0.1% Triton X-100) three times for 10 minutes each and blocked in PBTxGS (0.1% PBTx + 3% Normal Goat Serum) for 2–4 hours at RT. After blocking, midguts were incubated in primary antibody solution overnight at 4°C. The next day samples were washed with 0.1% PBTx three times for 20 minutes each and were incubated in secondary antibody solution for 2–3 hours at RT (protected from light) followed by three washes with 0.1% PBTx for 20 minutes each. One µg/ml DAPI solution prepared in 0.1% PBTx was added to the sample and incubated for 10 minutes followed by washing with 0.1% PBTx three times for 10 minutes each. Finally, samples were mounted on a slide glass with 70% glycerol and imaged using a Nikon AX R confocal microscope. Confocal images were processed using Fiji software.

The primary antibodies used were rabbit anti-GFP (A6455, 1:1000 Invitrogen), mouse anti-mCherry (3A11, 1:20 DSHB), mouse anti-Prospero (MR1A, 1:50 DSHB) and mouse anti-Pdm1 (Nub 2D4, 1:30 DSHB). The secondary antibodies used were goat antimouse and goat anti-rabbit IgG conjugated to Alexa 647 and Alexa 488 (1:200) (Invitrogen), respectively. Five larval gut specimens per Gal4 line were dissected and examined.”

(6) Doublets are removed based on the co-expression of two barcodes in Fig 5A. However, there are also other possible doublets, for example, from the same barcode cells or when one cell doesn't have detectable barcode. Did the authors try other computational approaches to remove doublets, like DoubleFinder (McGinnis et al., 2019) and Scrublet (Wolock et al., 2019)?

We have included DoubleFinder-based doublet removal in our data analysis pipeline. This is now described in the methods (see below).

(7) Did the authors remove ambient RNA which is a common issue for scRNA-seq experiments?

We have also used DecontX to remove ambient RNA. This is now described in the methods:

“Datasets were first mapped and analyzed using the Cell Ranger analysis pipeline (10x Genomics). A custom Drosophila genome reference was made by combining the BDGP.28 reference genome assembly and Ensembl gene annotations. Custom gene definitions for each of the TaG-EM barcodes were added to the fasta genome file and .gtf gene annotation file. A Cell Ranger reference package was generated with the Cell Ranger *mkref* command. Subsequent single-cell data analysis was performed using the R package Seurat (Satija et al., 2015). Cells expressing less than 200 genes and genes expressed in fewer than three cells were filtered from the expression matrix. Next, percent mitochondrial reads, percent ribosomal reads cells counts, and cell features were graphed to determine optimal filtering parameters. DecontX (Yang et al., 2020) was used to identify empty droplets, to evaluate ambient RNA contamination, and to remove empty cells and cells with high ambient RNA expression. DoubletFinder (McGinnis et al., 2019) to identify droplet multiplets and remove cells classified as multiplets. Clustree (Zappia and Oshlack, 2018) was used to visualize different clustering resolutions and to determine the optimal clustering resolution for downstream analysis. Finally, SingleR (Aran et al., 2019) was used for automated cell annotation with a gut single-cell reference from the Fly Cell Atlas (Li et al., 2022). The dataset was manually annotated using the expression patterns of marker genes known to be associated with cell types of interest. To correlate TaG-EM barcodes with cell IDs in the enriched TaG-EM barcode library, a custom Python script was used (TaGEM_barcode_Cell_barcode_correlation.py), which is available via Github: <https://github.com/darylgohl/TaG-EM>.”

(8) Why does TaG-EM barcode #4, driven by EC-GAL4, not label other classes of enterocyte cells such as betaTry+ positive ECs (Figures 5D-E)? similarly, why does TaG-EM barcode #9, driven by EE-GAL4, not label all EEs? Again, it is difficult to evaluate this part without proper data processing and accurate cell type annotation.

As noted in the response to a comment by Reviewer #1 above, part of this apparent sparsity of labeling is due to the way that this experiment was designed and visualized. We have added a new Figure panel in both Figure 6B and Supplemental Figure 17B that shows the overall detection of barcodes in the enriched barcode library and gene expression library or the gene expression library only, respectively, to better illustrate the efficacy of barcode detection. See also the response to point 5 above. Both the lack of labelling of betaTry+ ECs

and subsets of EEs is consistent with the expression patterns of the EC-Gal4 and EE-Gal4 drivers.

(9) For Figure 2, when the authors tested different combinations of groups with various numbers of barcodes. They found remarkable consistency for the even groups. Once the numbers start to increase to 64, barcode abundance becomes highly variable (range of 12-18% for both male and female). I think this would be problematic because the differences seen in two groups for example may be due to the barcode selection rather than an actual biologically meaningful difference.

While there is some barcode-to-barcode variability for different amplification conditions, the magnitude of this variation is relatively consistent across the conditions tested. We looked at the coefficient of variation for the evenly pooled barcodes or for the staggered barcodes pooled at different relative levels. While the absolute magnitude of the variation is higher for the highly abundant barcodes in the staggered conditions, the CVs for these conditions (0.186 for female flies and for 0.163 male flies) were only slightly above the mean CV (0.125) for all conditions (see Supplemental Figure 3):

We have added this analysis as Supplemental Figure 3 and added the following text to the paper:(

“The coefficients of variation were largely consistent for groups of TaG-EM barcodes pooled evenly or at different levels within the staggered pools (Supplemental Figure 3).”

(10) Barcode #14 cannot be reliably detected in oviposition experiment. This suggests that the BC 14 fly line might have additional mutations in the *atp2* chromosome arm that affects this behavior. Perhaps other barcode lines also have unknown mutations and would cause issues for other untested behaviors. One possible solution is to backcross all 20 lines with the same genetic background wild-type flies for >7 generations to make all these lines to have the same (or very similar) genetic background. This strategy is common for aging and behavior assays.

See response to Reviewer #1 above on this topic.

Reviewer #3 (Public Review):

The work addresses challenges in linking anatomical information to transcriptomic data in single-cell sequencing. It proposes a method called Targeted Genetically-Encoded Multiplexing (TaG-EM), which uses genetic barcoding in *Drosophila* to label specific cell populations in vivo. By inserting a DNA barcode near the polyadenylation site in a UASGFP construct, cells of interest can be identified during single-cell sequencing. TaG-EM enables various applications, including cell type identification, multiplet droplet detection, and barcoding experimental parameters. The study demonstrates that TaGEM barcodes can be decoded using next-generation sequencing for large-scale behavioral measurements. Overall, the results are solid in supporting the claims and will be useful for a broader fly community. I have only a few comments below:

We thank the reviewer for these positive comments.

Specific comments:

(1) The authors mentioned that the results of structure pool tests in Fig. 2 showed a high level of quantitative accuracy in detecting the TaG-EM barcode abundance. Although the data were generally consistent with the input values in most cases, there were some obvious exceptions such as barcode 1 (under-represented) and barcodes 15, 20

(overrepresented). It would be great if the authors could comment on these and provide a guideline for choosing the appropriate barcode lines when implementing this TaG-EM method.

See the response to point 9 from Reviewer 2. Although there seem to be some systematic differences in barcode amplification, the coefficient of variation was relatively consistent across all of the barcode combinations and relative input levels that we examined. Our recommendation (described in the text) is to average across 3-4 independent barcodes (which yielded a R2 values of >0.99 with expected abundance in the structured pooled tests).

(2) In Supplemental Figure 6, the authors showed GFP antibody staining data with 20 different TaG-EM barcode lines. The variability in GFP antibody staining results among these different TaG-EM barcode lines concerns the use of these TaG-EM barcode lines for sequencing followed by FACS sorting of native GFP. I expected the native GFP expression would be weaker and much more variable than the GFP antibody staining results shown in Supplemental Figure 6. If this is the case, variation of tissue-specific expression of TaG-EM barcode lines will likely be a confounding factor.

Aside from barcode 8, which had a mutation in the GFP coding sequence, we did not see significant variability in expression levels either in the wing disc. Subtle differences seen in this figure most likely result from differences in larval staging. Similar consistent native (unstained) GFP expression of the TaG-EM constructs was seen in crosses with Mhc-Gal4 (described above).

(3) As the authors mentioned in the manuscript, multiple barcodes for one experimental condition would be a better experimental design. Could the authors suggest a recommended number of barcodes for each experiential condition? 3? 4? Or more?

See response to Reviewer #3, point number 1 above.

(3b) Also, it would be great if the authors could provide a short discussion on the cost of such TaG-EM method. For example, for the phototaxis assay, if it is much more expensive to perform TaG-EM as compared to manually scoring the preference index by videotaping, what would be the practical considerations or benefits of doing TaG-EM over manual scoring?

While this will vary depending on the assay and the scale at which one is conducting experiments, we have added an analysis of labor savings for the larval gut motility assay (Supplemental Figure 8). We have also added the following text to the Discussion describing some of the trade-offs to consider in assessing the potential benefit of incorporating TaG-EM into behavioral measurements:

“While the utility of TaG-EM barcode-based quantification will vary based on the number of conditions being analyzed and the ease of quantifying the behavior or phenotype by other means, we demonstrate that TaG-EM can be employed to cost-effectively streamline labor-intensive assays and to quantify phenotypes with small effect sizes (Figure 4, Supplemental Figure 8).”

Recommendations for the authors:

While recognising the potential of the TaG-EM methodology, we had a few major concerns that the authors might want to consider addressing:

As stated above, we are grateful to the reviewers and editor for their thoughtful comments. We have addressed many of the points below in our responses above, so we will briefly

respond to these points and where relevant direct the reader to comments above.

(1) We were concerned about the efficacy of TaG-EM in assessing more complex behaviours than oviposition and phototaxis. We note that Barcode #14 cannot be reliably detected in oviposition experiment. This suggests that the BC 14 fly line might have additional mutations in the atp2 chromosome arm that affects this behavior. Perhaps other barcode lines also have unknown mutations and would cause issues for other untested behaviors. One possible solution is to back-cross all 20 lines with the same genetic background wild-type flies for >7 generations to make all these lines to have the same (or very similar) genetic background. This strategy is common for aging and behavior assays.

See response to Reviewer #1 and Reviewer #2, item 10, above.

(2) We were unable to assess the drop-out rates of the TaG-EM barcode from the sequencing. The barcode detection rate is low (Fig S9 and Fig 5F, J and N). This would be a considerable drawback (relating to both experimental design and cost), if a large proportion of the cells could not be assigned an identity.

See comments above addressing this point.

(3) The effectiveness of TaG-EM scRNA-seq on the larvae gut is not very effective - the cells are not well annotated, the barcodes seem not to have labelled expected cell types (ECs and EEs), and there is no validation of the Gal4 drivers in vivo.

See previous comments. We have addressed specific comments above on data processing and annotation, included a visualization of the overall effectiveness of labeling, added a protocol and data on enriched TaG-EM barcode libraries, and have added detailed characterization of the Gal4 drivers in the larval gut (Figure 6, Supplemental Figures 17-21).

(4) A formal assessment of the cost-effectiveness would be an important consideration in broad uptake of the methodology.

While this is difficult to do in a comprehensive manner given the breadth of potential applications, we have included estimates of labor savings for one of the behavioral assays that we tested (Supplemental Figure 8). We have also included a discussion of some of the factors that would make TaG-EM useful or cost-effective to apply for behavioral assays (see response to Reviewer #3, comment 3b, above). We have also added the following text to the discussion to address the cost considerations in applying TaG-EM for scRNA-Seq:

“For single cell RNA-Seq experiments, the cost savings of multiplexing is roughly the cost of a run divided by the number of independent lines multiplexed, plus labor savings by also being able to multiplex upstream flow cytometry, minus loss of unbarcoded cells. Our experiments indicated that for the specific drivers we tested TaG-EM barcodes are detected in around one quarter of the cells if relying on endogenous expression in the gene expression library, though this fraction was higher (~37%) if sequencing an enriched TaG-EM barcode library in parallel (Figure 6, Supplemental Figures 18-19).”

(5) Similarly, a formal assessment of the effect of the insertion on the variability in GFP expression and the behaviour needs to be documented.

See responses to Reviewer #1, Reviewer #2, item 9, and Reviewer #3, item 2 above.

Reviewer #1 (Recommendations For The Authors):

(in no particular order of importance)

• L84-85: *the authors should either expand, or remove this statement. Indeed, lack of replicates is only true if one ignores that each cell in an atlas is indeed a replicate. Therefore, depending on the approach or question, this statement is inaccurate.*

This sentence was meant to refer to experiments where different experimental conditions are being compared and not to more descriptive studies such as cell atlases. We have revised this sentence to clarify.

“Outside of descriptive studies, these costs are also a barrier to including replicates to assess biological variability; consequently, a lack of biological replicates derived from independent samples is a common shortcoming of single-cell sequencing experiments.”

• L103-104: *this sentence is unclear.*

We have revised this sentence as follows:

“Genetically barcoded fly lines can also be used to enable highly multiplexed behavioral assays which can be read out using high throughput sequencing.”

• *In Fig S1 it is unclear why there are more than 20 different sequences in panel B where the text and panel A only mention the generation of 20 distinct constructs. This should be better explained.*

The following text was added to the Figure legend to explain this discrepancy:

“Because the TaG-EM barcode constructs were injected as a pool of 29 purified plasmids, some of the transgenic lines had inserts of the same construct. In total 20 unique lines were recovered from this round of injection.”

• *It would be interesting to compare the efficiency of TaG-EM driven doublet removal (Fig 5A) with standard doublet-removing software (e.g., DoubletFinder, McGinnis et al., 2019).*

We have done this comparison, which is now shown in Supplemental Figure 15.

• *I would encourage the authors to check whether barcode representation in Fig S13 can be correlated to average library size, as one would expect libraries with shorter reads to be more likely to include the 14-bp barcode and therefore more accurately recapitulate TaG-EM barcode expression.*

These are not independent sequencing libraries, but rather data from barcodes that were multiplexed in a single flow sort, 10x droplet capture, and sequencing library. Thus, there must be some other variable that explains the differential recovery of these barcodes.

• *Fig 4A should appear earlier in the paper.*

We have moved Figure 4A from the previous manuscript (a schematic showing the detailed design of the TaG-EM construct) to Figure 1A in the revised version.

Reviewer #2 (Recommendations For The Authors):

Minor:

(1) There is a typo for Fig S13 figure legends: BC1, BC1, BC3... should be BC1, BC2, BC3.

Fixed.

Reviewer #3 (Recommendations For The Authors):

Comments to authors:

(1) It would be great if the authors could provide an additional explanation on how these 29 barcode sequences were determined.

Response: This information is in the Methods section. For the original cloned plasmids:

“Expected construct size was verified by diagnostic digest with _Eco_RI and _Apa_LI. DNA concentration was determined using a Quant-iT PicoGreen dsDNA assay (Thermo Fisher Scientific) and the randomer barcode for each of the constructs was determined by Sanger sequencing using the following primers:

SV40_post_R: GCCAGATCGATCCAGACATGA

SV40_5F: CTCCCCTGAACCTGAAACA”

For transgenic flies, after DNA extraction and PCR enrichment (details also in the Methods section):

“The barcode sequence for each of the independent transgenic lines was determined by Sanger sequencing using the SV40_5F and SV40_PostR primers.”

(2) Why did the authors choose myr-GFP as the backbone instead of nls-GFP if the downstream application is to perform sequencing?

We initially chose myr::GFP as we planned to conduct single cell and not single nucleus sequencing and myr::GFP has the advantage of labeling cell membranes which could facilitate the characterization or confirmation of cell type-specific expression, particularly in the nervous system. However, we have considered making a version of the TaG-EM construct with a nuclear targeted GFP (thereby enabling “NucEM”). In the Discussion, we mention this possibility as well as the possibility of using a second nuclear-GFP construct in conjunction with TaG-EM lines is nuclear enrichment is desired:

“In addition, while the original TaG-EM lines were made using a membrane-localized myr::GFP construct, variants that express GFP in other cell compartments such as the cytoplasm or nucleus could be constructed to enable increased expression levels or purification of nuclei. Nuclear labeling could also be achieved by co-expressing a nuclear GFP construct with existing TaG-EM lines in analogy to the use of hexameric GFP described above.”

Minor comments:

(1) Line 193, Supplemental Figure 4 should be Supplemental Figure 5

Fixed.

(2) Scale bars should be added in Figure 4, Supplemental Figures 6, 7, and 8A.

We have added scale bars to these figures and also included scale bars in additional Supplemental Figures detailing characterization of the gut driver lines.

| (3) Were Figure 4C and Supplemental Figure 7 data stained with a GFP antibody?

No, this is endogenous GFP signal. This is now noted in the Figure legends.

| (4) Line 220, specify the three barcode lines (lines #7, 8, 9) in the text.

Added this information.

Same for Lines 251-254. Line 258, which 8 barcode Gal4 line combinations?

| (5) Line 994, typo: (BC1, BC1, BC3, and BC7)-> (BC1, BC2, BC3, and BC7)

Fixed.

| (6) Figure 5 F, J and N, add EC-Gal4, EB-Gal4, and EE-Gal4 above each panel to improve readability.

We have added labels of the cell type being targeted (leftmost panels), the barcode, and the marker gene name to Figure 6 C-N.

<https://doi.org/10.7554/eLife.88334.2.sa0>