frontiers | Frontiers in Genome Editing

TYPE Original Research
PUBLISHED 21 July 2023
DOI 10.3389/fgeed.2023.1167093

Check for updates

OPEN ACCESS

EDITED BY
Beau Alward,
University of Houston, United States

REVIEWED BY
Scott Alan Juntti,
University of Maryland, United States
Erica Westerman,
University of Arkansas, United States
James Gagnon,
The University of Utah, United States

*CORRESPONDENCE
Andres Bendesky,
a.bendesky@columbia.edu

†Present address: Madison R. Lichak, Department of Ecology and Evolutionary Biology, Princeton University, New York, NJ, United States

‡These authors have contributed equally to this work

RECEIVED 15 February 2023
ACCEPTED 07 July 2023
PUBLISHED 21 July 2023

CITATION
Palmiotti A, Lichak MR, Shih P-Y, Kwon YM and Bendesky A (2023), Genetic manipulation of betta fish.
Front. Genome Ed. 5:1167093.
doi: 10.3389/fgeed.2023.1167093

COPYRIGHT
© 2023 Palmiotti, Lichak, Shih, Kwon and Bendesky. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Genetic manipulation of betta fish

Alec Palmiotti[1,2‡], Madison R. Lichak[1,2†‡], Pei-Yin Shih[1,2], Young Mi Kwon[1,2] and Andres Bendesky[1,2]*

[1]Zuckerman Mind Brain Behavior Institute, Columbia University, New York, NY, United States, [2]Department of Ecology, Evolution and Environmental Biology, Columbia University, New York, NY, United States

Betta splendens, also known as Siamese fighting fish or "betta," is a freshwater fish species renowned for its astonishing morphological diversity and extreme aggressive behavior. Despite recent advances in our understanding of the genetics and neurobiology of betta, the lack of tools to manipulate their genome has hindered progress at functional and mechanistic levels. In this study, we outline the use of three genetic manipulation technologies, which we have optimized for use in betta: CRISPR/Cas9-mediated knockout, CRISPR/Cas9-mediated knockin, and Tol2-mediated transgenesis. We knocked out three genes: *alkal2l, bco1l,* and *mitfa,* and analyzed their effects on viability and pigmentation. Furthermore, we knocked in a fluorescent protein into the *mitfa* locus, a proof-of-principle experiment of this powerful technology in betta. Finally, we used Tol2-mediated transgenesis to create fish with ubiquitous expression of GFP, and then developed a bicistronic plasmid with heart-specific expression of a red fluorescent protein to serve as a visible marker of successful transgenesis. Our work highlights the potential for the genetic manipulation of betta, providing valuable resources for the effective use of genetic tools in this animal model.

KEYWORDS

betta fish, genetic manipulation, CRISPR, Tol2 system, genome editing, transgenesis methodologies, knock in, knock out

1 Introduction

The Siamese fighting fish (*Betta splendens*) or more commonly "betta," is a species of freshwater fish known for its vibrant and diverse colors and fin morphologies, as well as its exceptional aggressive behavior. Native to Southeast Asia, betta have experienced a long history of domestication and selective breeding beginning at least 400 years ago (Kwon et al., 2022). Due to their morphological and phenotypic diversity, relatively compact genome size (~430 Mb), their ease of growth in the laboratory, and their amenability to behavioral and neurobiological experimentation, betta have become an increasingly popular organism for scientific study (Lichak et al., 2022). The recent publication of high-quality reference genomes of both domesticated and wild betta has allowed for in-depth genetic and genomic analyses of sex determination, phenotypic traits such as pigmentation, fin shape and aggression, and of the evolutionary relationships between *Betta* species (Fan et al., 2018; Prost et al., 2020; Wang et al., 2021; 2022; Kwon et al., 2022; Yue et al., 2022; Zhang et al., 2022).

Establishing genetic tools in betta will advance their use as a powerful experimental system to study developmental processes and behavioral traits. Genetic manipulation of betta is facilitated by their reproductive biology: betta fertilize externally and produce clutches of ~250 eggs, each with a relatively large diameter of 1 mm (Valentin et al., 2015; Lichak et al., 2022). This enables the microinjection of zygotes in a similar manner to

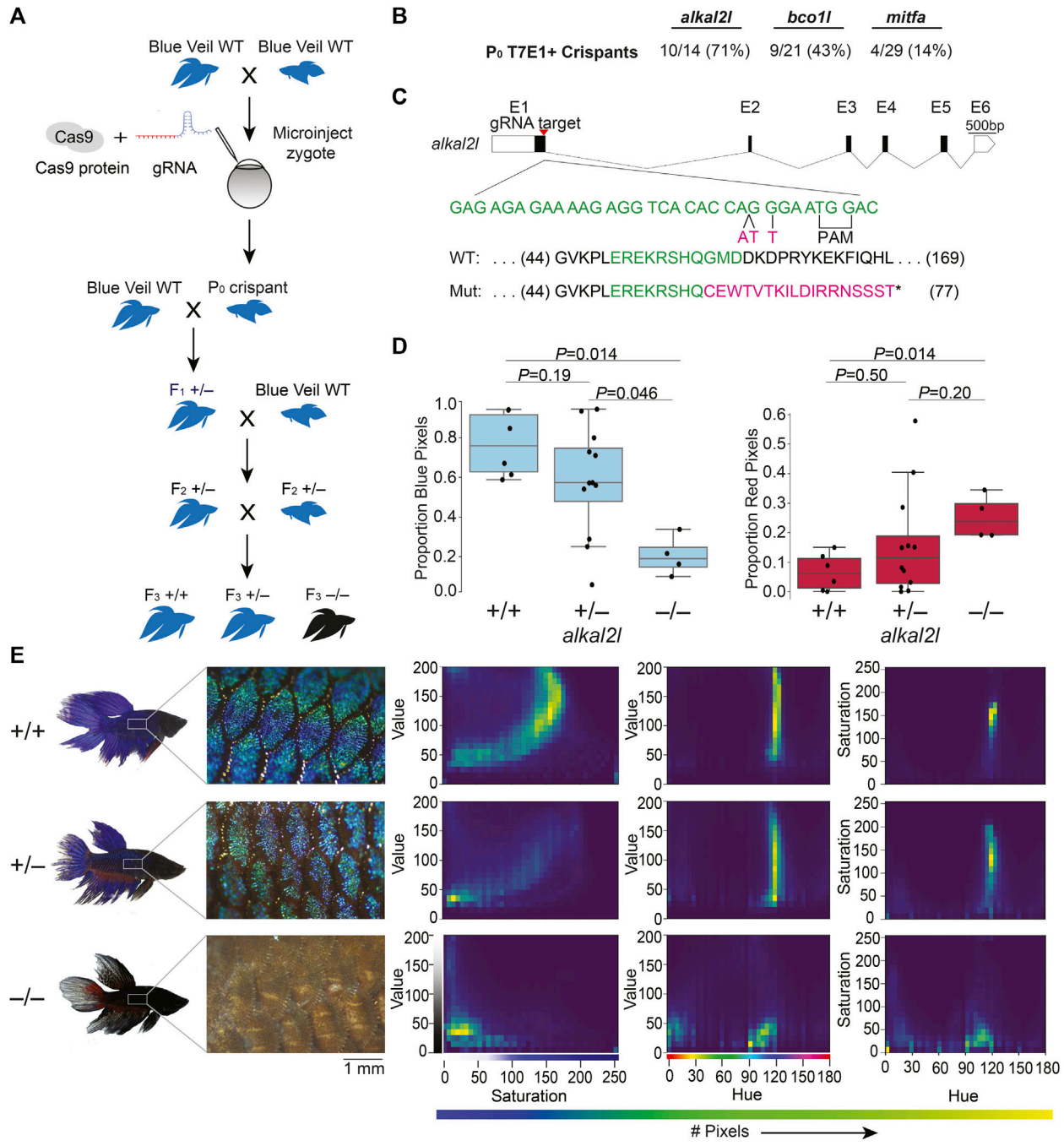


FIGURE 1
CRISPR/Cas9-mediated knockout in betta **(A)** Experimental scheme for generating knockout betta. **(B)** Efficiency of knockout generation in P0, as determined by a T7EI assay. Numbers denote the fraction and percent of T7EI+ P0 individuals. **(C)** Schematic of *alkal2l* gene and the location of the gRNA target (top); wildtype DNA sequence in green and CRISPR/Cas9-induced mutations in magenta (middle); wildtype protein sequence (green shows the amino acids encoded by the wildtype sequence shown in the middle) and mutant sequence due to a frameshift (magenta shows the new mutant amino acids and * denotes a stop codon). **(D)** Boxplots showing proportion of blue (left) and red (right) pixels in *alkal2l* F3 crispants, according to genotype. *p*-values by Mann-Whitney test adjusted by Bonferroni correction. Boxes denote the interquartile range and whiskers the 5th and 95th percentiles, with a line at the median. **(E)** Representative images of betta individuals according to *alkal2l* genotype (left). Microscope images of the side of the body showing scales and the iridophore cover (center); the color of the microscope images appears different from the whole-body photographs on the left, due to the use of bright incident light under the microscope. Joint per-pixel quantification of hue, saturation and value by genotype (right).

methods for genetic manipulation of zebrafish (*Danio rerio*) and medaka (*Oryzias latipes*). Nevertheless, the asynchronous egg fertilization resulting from a protracted mating process that can last many hours, coupled with a short interval between fertilization and cell division, as well as a thick chorion, constitute significant challenges for genetic manipulation (Valentin et al., 2015; Lichak

et al., 2022). Although a few studies have successfully utilized these tools, their reported success rates have been low, and none have reported germline transmission (Wang et al., 2021; 2022; Yue et al., 2022). Furthermore, although plasmid DNA microinjection has been used to make transgenic commercial betta (GloFish® Betta), the more controlled and efficient Tol2-based approach has not been performed. Tol2-mediated transgenesis has advantages over simple plasmid injections, such as improved rates of germline transmission and higher frequency of single copy integration (Kawakami, 2007).

In this study, we established effective protocols for three different genetic manipulation technologies in betta: CRISPR/Cas9 knockout, CRISPR/Cas9 knockin, and Tol2 transgenesis. We use CRISPR/Cas9 to knockout three genes, *ALK and LTK-ligand 2-like* (*alkal2l*), *beta-carotene oxygenase like-1* (*bco1l*), and *melanocyte inducing transcription factor a* (*mitfa*) and provide a phenotypic analysis of pigmentation and viability of *alkal2l* and *bco1l* germline crispants. We also report the first successful CRISPR/Cas9-mediated knockin, as well as the first use of Tol2-based transgenesis in betta. These methods, along with examples of their applications, highlight the potential for further adoption of genetic tools to study betta.

2 Results

2.1 CRISPR/Cas9 knockout editing of betta

We chose *alkal2l* and *bco1l* as targets for CRISPR/Cas9 editing, based on a previous genetic mapping study from our lab that identified variation in these genes as likely contributing to blue or red coloration in ornamental betta (Kwon et al., 2022). *alkal2l* alleles present in blue fish are associated with an increase in the proportion of blue coloration covering the body, and a concomitant decrease in the proportion of red, whereas *bco1l* alleles likely modulate red hue (Kwon et al., 2022). Although the molecular mechanisms responsible for this phenotypic variation have not been studied in betta, *alkal2l* is necessary in zebrafish for the differentiation of iridophores that contribute to blue iridescence (Mo et al., 2017), and *bco1l* cleaves β-carotene, a red-orange molecule, into two molecules of all-trans retinal (Kwon et al., 2022). As a positive control of the induction of CRISPR/Cas9-mediated deletions, we also targeted a third gene, *mitfa,* using a guide RNA that has been validated in betta (Wang et al., 2021). We injected 1–4 celled embryos with a ribonucleoprotein (RNP) injection mix consisting of Cas9 protein and guide RNA (gRNA; Figure 1A; Supplementary Figures S1, S4A). We interrupted betta 1–2 h after they began mating, as this maximizes the number of embryos at the 1–4 cell stage, and injected these embryos using a very sharp needle made with a micropipette puller and beveled using scissors, which facilitates the penetration of the thick chorion (see Methods and Supplementary Figure S2). We then allowed surviving embryos to develop for 5 weeks (when they are about 12 mm in length (Lichak et al., 2022)) before extracting their DNA from a fin clip and genotyping them using a T7 endonuclease I (T7EI) assay (Supplementary Figure S3A). Ten of the fourteen (71%) *alkal2l*-targeted fish for which PCR amplification of a genomic region surrounding the cleavage site was successful, showed T7EI activity indicative of Cas9 cutting activity. The rate of CRISPR/Cas9-mediated mutagenesis was lower for *bco1l* (9/21; 43%), and lowest for *mitfa* (4/29; 14%), yet at sufficiently high rates to be a practical approach (Figure 1B). Because betta *mitfa* knockout phenotypes have been reported previously (although only in the P0 generation (Wang et al., 2021)), we focused on the *alkal2l* and *bco1l* mutants for further study.

We crossed the P0 crispants to wildtype betta of the same color and fin type as the parents of the P0 animals (*bco1l* crispants to red veiltail fish and *alkal2l* crispants to blue veiltail fish) and genotyped the F1 offspring using a T7EI assay. One out of two *bco1l* and two out of two *alkal2l* P0 crispants we crossed transmitted a mutant allele. The *bco1l* P0 crispant transmitted a mutant allele to 3% (5/149) of its F1 offspring, whereas the *alkal2l* P0 crispants transmitted to 47% (8/17) and 80% (8/10) of their F1 offspring. The varying levels likely reflect the degree of mosaicism in the CRISPR/Cas9-induced mutations in the P0 crispants.

Next, we sequenced the targeted loci of 15 T7EI+ F1 offspring (5 *bco1l*, 8 *alkal2l)* to characterize the CRISPR-induced mutations. Sequence analysis revealed a variety of mutations in the F1 generations of *bco1l* and *alkal2l.* We chose one F1 crispant per gene that carried frameshift mutations leading to an early stop codon. We then crossed each of these F1 crispants in a second outcross to wildtype betta to obtain F2 generations containing heterozygotes genetically identical at the targeted loci (Figure 1A). Finally, we crossed two F2 heterozygous sibling pairs to each other to obtain a set of F3 animals in which we could compare homozygous wildtype, heterozygous, and homozygous mutant animals within a clutch.

For *bco1l,* the mutation we chose was a T insertion that led to a frameshift coupled to a premature stop codon, 21 amino acids downstream of the indel (Supplementary Figure S4B). Sequencing of the *bco1l* F3 individuals revealed that homozygous mutants were absent from the brood (19 +/+, 36 +/− and 0 −/−, $p = 6 \times 10^{-5}$ by Fisher's exact test compared to the expected 1:2:1 Mendelian ratio). This deficit of homozygous mutants could be due to lethality or to a prezygotic effect of the mutation in the eggs or sperm. Our P0 crispant was a female and her F1 offspring that sired the F2 generation was a male, demonstrating that the mutation can transmit through both eggs and sperm and that inheriting two mutant copies of *bco1l* is lethal. Image analyses of the color distributions and hues of *bco1l* +/+ vs. +/− adult fish revealed no significant difference in the proportion of red and orange pixels, or of the joint hue, saturation and value (HSV) distributions (Supplementary Figures S4C,D). Therefore, allelic variation at *bco1l* between red and blue ornamental betta is unlikely to be composed of null alleles, and future reciprocal hemizygosity assays will evaluate the involvement of *bco1l* in red–blue fish variation.

For *alkal2l,* we chose a CRISPR-induced mutation consisting of a two base-pair insertion coupled with a nearby single nucleotide mismatch. This mutation led to a frameshift associated with a premature stop codon 18 amino acids downstream of the indel (Figure 1C). Next, we analyzed the impact of *alkal2l* in the F3 generation containing +/+, +/−, and −/− genotypes (Figures 1D,E; Supplementary Figure S3B). We found a significant difference in the proportion of blue covering the body of *alkal2l* −/− compared to the other two genotypes, with −/− animals having only 25% as much blue cover as +/+ animals (Figure 1D). Heterozygotes were

10.3389/fgeed.2023.1167093

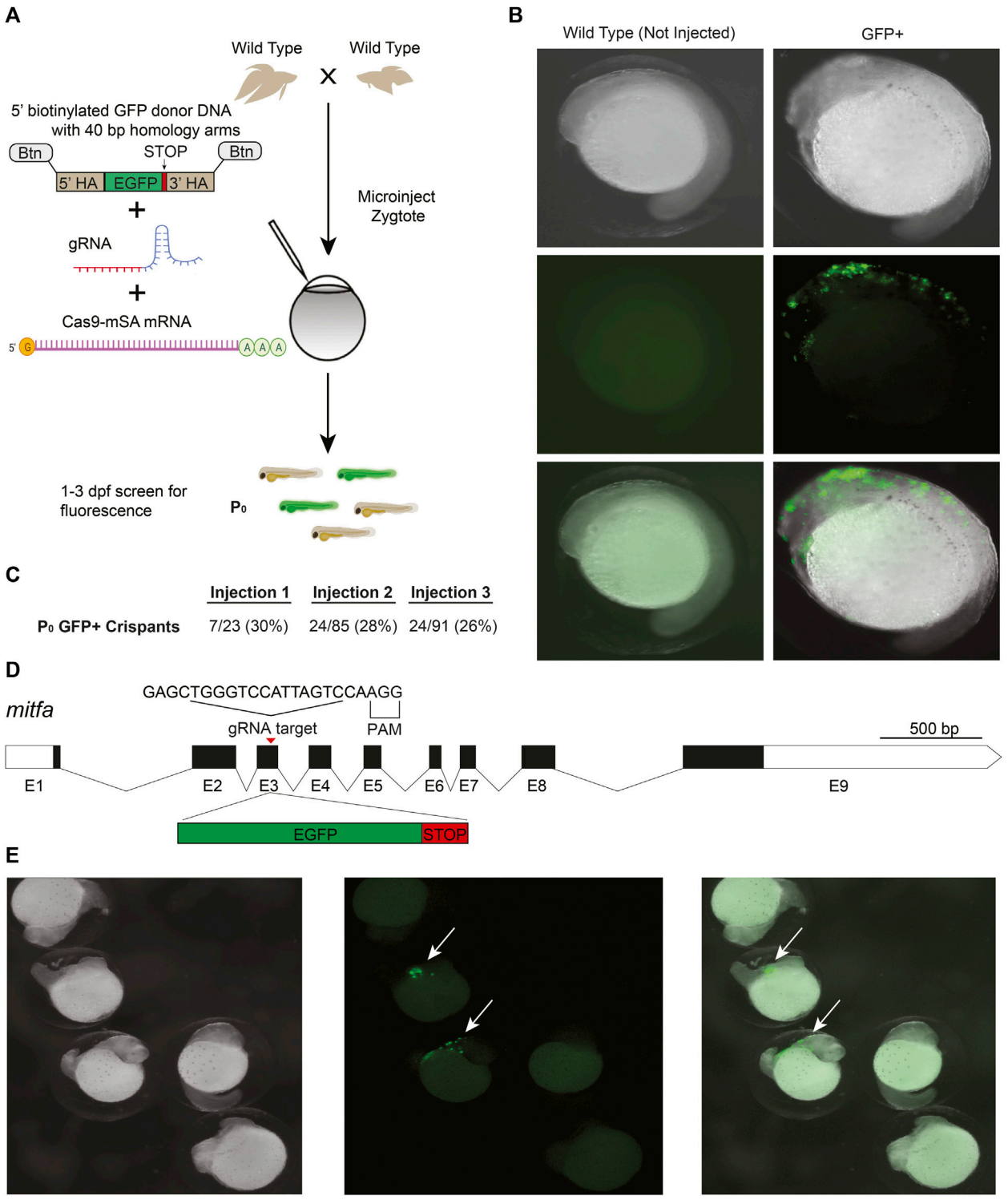


FIGURE 2
CRISPR/Cas9-mediated knockin in betta **(A)** Experimental scheme for generating knockin betta. **(B)** Brightfield (top), fluorescence only (middle) and merged images of wild type (non-injected) betta and P0 GFP+ betta at 24 hpf. **(C)** Efficiency of knockin generation across injection of three different clutches. Numbers denote the fraction and percent of GFP+ individuals **(D)** Schematic of *mitfa* gene and location of the gRNA target for insertion of GFP. **(E)** Brightfield (left), fluorescence only (center) and merge (right) of 24-hpf injected embryos with clear GFP expression (arrow) and no expression, even at low magnification.

not different from wildtype, indicating that the *alkal2l* mutation is recessive (Figure 1D). The extent of blue iridescence in *alkal2l* −/− was reduced in both the body and the fins, yet the reduction was more prominent in the body, consistent with previous findings showing that *alkal2l* mRNA is present at 210× higher levels in the body compared to the fins (Kwon et al., 2022). Consistent with the reduction in blue coloration, the proportion of red surface was 5× more extensive in homozygous mutants compared to wild type fish (Figure 1D). Microscope images of the skin surface revealed that the blue iridescence mediated by iridophores is almost entirely absent on the body of *alkal2l* −/− animals, but mostly normal on the body of +/− mutants, suggesting that one functional *alkal2l* allele is sufficient for iridescence in betta (Figure 1E). To better visualize differences in coloration, we plotted the joint distributions of per-pixel HSV (Figure 1E). The joint distributions of +/+ and +/− are similar, whereas −/− is an outlier. For hue vs. saturation, in both +/+ and +/− animals, the majority of pixels have a hue value of ~120 (corresponding to blue on a 180-degree hue color space), and these blue pixels range in saturation, with most falling between 100 and 200. The same graph for −/− shows a variety of different hues ranging from 0 to 120 with a saturation mostly below 50, indicating fewer blue pixels and less vibrant colors (Figure 1E). Furthermore, pixels in the −/− distributions are aggregated at much lower values than in the +/+ and +/− distributions, confirming that −/− mutants are darker and supporting the observation that knocking out *alkal2l* increases the visibility of deeper pigment layers that include melanophores (melanin-containing cells). In addition to increasing the visibility of deeper layers, the absence of *alkal2l* might also induce some skin cells to develop a different color fate. Together, these findings indicate that our method for obtaining CRISPR/Cas9 knockouts in betta is efficient and effective, that a homozygous knockout of *bco1l* is lethal, and that *alkal2l* is critical for blue coloration.

2.2 CRISPR/Cas9 knockin editing of betta

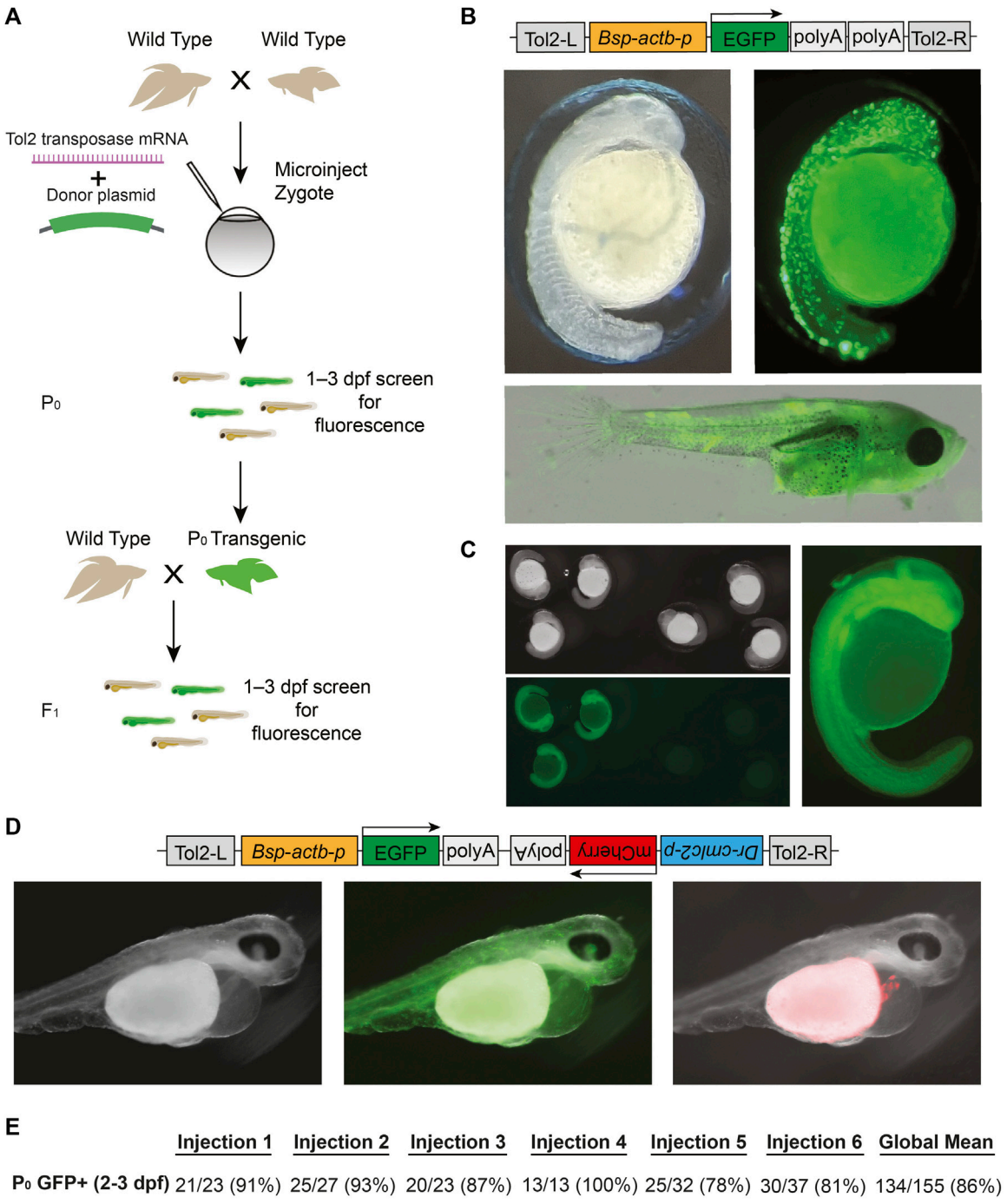
CRISPR/Cas9 knockin, which relies on the homology directed repair (HDR) pathway to integrate exogenous DNA into the host genome, has become the preferred method for the precise single-copy integration of various genome modifications such as fluorescent proteins, mutations relevant to specific disease models, and for epitope tagging (Hisano et al., 2015; Danner et al., 2017; Lino et al., 2018; Watakabe et al., 2018; Ranawakage et al., 2020). Recent advances in knockin technology, such as the use of 5′-biotinylated long homology arms, streptavidin-tagged Cas9, and *in vivo* linearization of the donor plasmid, have been shown to increase the efficiency of knockins in zebrafish and mammalian cells (Gutierrez-Triana et al., 2018; Wierson et al., 2020). To develop knockin technology for betta, we adopted a recent efficient, cloning-free CRISPR/Cas9-mediated knockin protocol that has been used in medaka (Seleit et al., 2021).

We targeted an insertion of GFP into *mitfa*, as this should lead to GFP expression in the precursors of melanophores (melanin-containing cells) early in development, facilitating visual screening (Figure 2D). To increase single-copy integration events and decrease the likelihood of donor DNA concatemerization (Gutierrez-Triana et al., 2018), we used primers with 5′-biotin modifications to amplify GFP with the homology arms by PCR (Seleit et al., 2021). We used monomeric streptavidin-tagged Cas9 mRNA (Cas9-mSA) to enhance Cas9 binding to the biotinylated donor constructs (Seleit et al., 2021). We then injected a mix consisting of *mitfa* gRNA (the same one we used for making knockouts), 5′-biotinylated donor DNA, and Cas9-mSA mRNA into 1–4 cell embryos (Figure 2A). As early as 18 h post fertilization (hpf) and more clearly at 24 hpf, robust GFP expression could be easily identified under the microscope in cells whose location is consistent with that of neural crest cells, which give rise to melanophores (Silver et al., 2006) (Figures 2B,E). The onset (at ~18 hpf) and disappearance (from ~48 to 96 hpf) of GFP expression is also consistent with the developmental timing of *mitfa* expression in zebrafish, where it is no longer expressed in differentiated melanophores at 96 hpf (Lister et al., 1999). In three separate injections, the percentage of GFP+ fish were: 30%, 28%, and 26% (Figure 2C). These findings indicate successful CRISPR/Cas9-mediated knockins in betta, although we have yet to grow these fish to adulthood to test for transmission and sequence the knockin alleles.

2.3 Tol2-mediated transgenesis in betta

Tol2 transgenesis, which relies on the autonomous Tol2 transposable element first discovered in medaka (Kawakami et al., 1998), has become the predominant mechanism for transgenesis in zebrafish and in other vertebrates due to its high efficiency and simplicity of transgene construct design (Kawakami et al., 1998; Urasaki et al., 2006; Kawakami, 2007). The standard protocol for Tol2 transgenesis involves the co-injection of a plasmid that contains Tol2 transposon repeats flanking the DNA sequence to be randomly integrated in the genome, along with synthetic transposase mRNA. As a first step towards adopting Tol2 transgenesis in betta, we opted for a promoter that would drive ubiquitous expression of GFP, facilitating successful screening under the microscope. To that end, we cloned the 4.2 kb region upstream of the betta β-actin (*actb*) transcription start site in front of GFP in a Tol2 plasmid to make pAB-1 (Figure 3B). We then injected a mix consisting of pAB-1 and transposase mRNA into 1–4 celled betta embryos (Figure 3A). In injections of six independent clutches, robust GFP expression indicative of transgenesis was apparent in the majority of embryos, with percentages ranging from 78% to 100% (Figure 3E). Expression was visible starting at 20 hpf. During the juvenile developmental period of 10–30 dpf, GFP expression was especially prominent in muscle, but was also visible in a variety of tissues including the brain, heart, bladder and intestine (Figure 3B). GFP expression remained visible into adulthood, suggestive of genomic integration rather than transient expression from the injected plasmid. To assess the integration of the transgene into the germline, we crossed five P0 transgenics with variable bodily extents of GFP expression to wild type betta. Of these five crosses, one generated offspring with GFP expression (26/153 larvae; 16%), indicating germline transmission of the transgenic construct (Figure 3C). It is likely that broader GFP expression in the P0 increases the probability of transmission of the transgene, but we have not measured this yet.

10.3389/fgeed.2023.1167093



	Injection 1	Injection 2	Injection 3	Injection 4	Injection 5	Injection 6	Global Mean
P0 GFP+ (2-3 dpf)	21/23 (91%)	25/27 (93%)	20/23 (87%)	13/13 (100%)	25/32 (78%)	30/37 (81%)	134/155 (86%)

FIGURE 3
Tol2-mediated transgenesis in betta **(A)** Experimental scheme for generating transgenic betta. **(B)** Schematic of pAB-1 *actb*:EGFP Tol2 plasmid (top), brightfield image of 24-hpf injected embryo (left) and fluorescent images of 24-hpf embryo (right) and ~30-dpf betta (bottom). **(C)** Brightfield (top left) and fluorescent bottom left, right) images of *actb*:EGFP F1 transgenics. **(D)** Schematic of pAB-16 Tol2 bicistronic plasmid (top) and images of 3-dpf injected embryos using bright field (left) and GFP filter (middle) and mCherry filter (right). **(E)** Efficiency of transgenic generation across injections of six different clutches. Numbers denote the fraction and percentage of GFP+ P0 individuals.

Next, to make a bicistronic plasmid expressing both GFP ubiquitously, as well as mCherry in the heart to serve as an additional visible marker, we cloned the 268-bp region upstream of zebrafish *cmlc2* (*myl7*) (Huang et al., 2003) in front of mCherry into pAB-1 to make pAB-16 (Figure 3D). We injected pAB-16 into 1–4 celled embryos (N = 211) using the same injection mix components as above. At 3 dpf, we observed both ubiquitous GFP expression, as well as robust mCherry expression in the

heart of injected embryos (36 out of 72 surviving embryos; Figure 3D). At 3 dpf, but not at 1 and 2 dpf, mCherry expression in the heart was bright enough that it could be used to quickly screen for transgenesis. Together, these results indicate that Tol2-mediated transgenesis is a highly effective method for generating transgenics in betta. Our constructs also offer effective visual markers of transgenesis, and constitute useful starting points for cloning future transgenes.

3 Discussion

We established protocols for three genetic modification technologies in betta—CRISPR/Cas9 knockout, CRISPR/Cas9 knockin, and Tol2 transgenesis—each with efficiencies on par with current methodologies developed for more established model organisms such as zebrafish and medaka. CRISPR/Cas9 knockout has been previously used to study betta, yet our P0 mutagenesis efficiencies for the three genes we targeted (14%, 43%, and 71%) are markedly higher than the efficiency reported by others (4/2000; 0.2%) when targeting *dmrt1* in betta (Wang et al., 2022). CRISPR mutagenesis efficiencies using our methods are comparable to those seen in zebrafish, where single gRNA editing efficiencies initially ranged from 25% to 60% (Hwang et al., 2013; Liu et al., 2019), and more recent reports using multi-guide RNP complexes report editing efficiencies ranging from 80% to 100% (Klatt Shaw and Mokalled, 2021; Kroll et al., 2021). By developing high-efficiency CRISPR mutagenesis protocols for betta, we have been able to study two genes—*alkal2l* and *bco1l*—that we had found to be linked to red–blue variation among ornamental betta (Kwon et al., 2022). We find that *alkal2l* is required for blue iridescence, consistent with its role in iridophore development in zebrafish (Mo et al., 2017). A loss-of-function mutation of *alkal2l* recessively leads to loss of blue coloration; these *alkal2l* mutant fish are instead darker and redder.

Tetrapods have two beta-carotene oxygenase genes: *bco1* and *bco2*. In addition, betta fish, like other fish, have a third *bco* gene: *bco1l* (Kwon et al., 2022). Interestingly, we find that *bco1l* homozygous mutations are lethal, indicating that its function is not redundant with that of *bco1* nor *bco2*. This lethality is possibly a result of a lack of retinal, the product of β-carotene cleavage by BCO1L. There was no alteration in coloration nor lethality in *bco1l* heterozygous mutants, indicating that a single functional copy of *bco1l* is enough for viability and red coloration. Together, these results suggest that genetic variation in *bco1l* segregating among betta may affect red coloration through more subtle effects—such as a gain of function or changes in spatiotemporal patterns of gene expression—than simple reduction of function.

Our P0 knockin targeting efficiency (28%) is consistent with efficiencies in medaka (11%–59%) using the same method (Seleit et al., 2021). The efficiency of germline transmission of the knockin alleles, however, remains to be determined. While here we use streptavidin-tagged Cas9 and 5′-biotinylated donor DNA, it is not clear if these modifications to Cas9 and the donor DNA improves knockin efficiency and decreases toxicity (Seleit et al., 2021). A recent report suggests that 5′-AmC6 modified donor DNA is more effective than 5′-biotinylation for CRISPR/Cas9 knockin in zebrafish (Mi and Andersson, 2023). It will be useful to determine if these observations hold in betta.

Our targeting efficiency for making transgenics using the Tol2 transposon system (86%) is close to that of zebrafish (~100%) (Urasaki et al., 2006). To facilitate the visual screening of transgenic animals, we have furthermore developed a bicistronic vector that drives expression of a fluorescent protein in the heart, along with a separate molecular payload of interest. Altogether, the methods we developed constitute powerful tools to manipulate the betta genome. These tools will enable modern approaches to study the genetic mechanisms underlying the astonishing morphological diversity of betta, as well as its unique aggressive behavior.

4 Materials and methods

4.1 Husbandry

Animals were maintained according to our standard procedures (Lichak et al., 2022). All animal protocols were approved by the Columbia University Institutional Animal Care and Use Committee. The night prior to injections, we set up mating pairs according to published protocols (Lichak et al., 2022). We added each mating pair to a 5 L acrylic tank filled to a height of 10 cm with system water. Each tank contained one terra cotta pot, one acrylic plant, and a floating piece of dried Indian almond leaf (*Terminalia catappa,* from SunGrow) to facilitate bubble nest building and allow for convenient egg collection. During mating, tanks were kept at 28–30°C using heating mats (VivoSun, Ontario, CA), and were covered with acrylic to maintain high levels of humidity (~90%) relative to the betta facility (~25%). The following morning, we checked if fish were mating, and allowed for 1–2 h of undisturbed mating before collecting eggs for microinjection. Eggs were then collected by carefully inverting the almond leaf containing the bubble nest and then gently washing the nest containing the eggs into a 250 mL beaker filled with system water from our recirculating fish rack.

4.2 Microinjection

Prior to the day of injection, we made microinjection needles and injection molds, and prepared working concentrations of components of microinjection mixes. To make microinjection needles we used a micropipette puller (Model P-1000, Sutter Instrument, Novato, CA) to pull 4-inch long thin wall filamented glass capillary tubes (TW100F-4, World Precision Instruments, Sarasota, FL) into injection needles using custom settings: Heat: 480, Pull: 90, Velocity: 80, Delay: 120, Pressure: 230, Ramp: 467. We then used micro-scissors (15000-08, Fine Science Tools, Foster City, CA) to cut the pulled needles at a ~45° angle under a dissection microscope so as to bevel the needle to facilitate piercing of the thick chorion (Supplementary Figure S2). The diameter of the tip after cutting is 8 µm. We then stored the pulled needles in a 150 mm petri dish on mounting putty (XTREME Putty, Tombow, Suwanee, GA) until use. We made injection molds to hold embryos for injection by dissolving 1 g agarose in 50 mL E3 medium containing 0.0005% methylene blue, pouring the hot agarose solution into a petri dish,

then placing the dry injection mold (TU-1, Adaptive Science Tools, Worcester, MA) into the solution when it reached 55°C. A 60× solution of E3 medium can be made by combining 8.7 g NaCl, 0.4 g KCl, 1.45 g $CaCl_2{\cdot}2H_2O$, 2.45 g $MgCl_2{\cdot}6H_2O$ and bringing to 500 mL using reverse osmosis water, followed by pH adjustment to 7.2 using NaOH. We removed the mold when the agarose solution solidified and stored the plate containing the mold at 4°C.

In advance of injections, we also resuspended CRISPR RNAs (crRNAs) (Alt-R® CRISPR-Cas9 crRNA XT, 2 nmol, Integrated DNA Technologies, Coralville, IA) and tracrRNA (1072532, Alt-R® CRISPR-Cas9 tracrRNA, 5 nmol, Integrated DNA Technologies) to a working concentration of 100 µM using TE buffer (1 mM EDTA, 10 mM Tris-Cl; pH 8.0). We then assembled the gRNA on ice by combining 1.5 µL 100 µM crRNA, 1.5 µL 100 µM tracrRNA, and 47 µL Duplex buffer (11-05-01-03, Integrated DNA Technologies), and then heating the solution to 95 °C for 5 min before removing from heat and letting the solution cool passively to room temperature (21–22°C) on benchtop. These mixes were stored at −70°C until injection. We also diluted Cas9 protein (1081058, Alt-R® S.p. Cas9 Nuclease V3, 10 µg/µL, Integrated DNA Technologies) by combining 0.5 µL stock Cas9 protein with 9.5 µL Cas9 buffer (20 mM HEPES, 150 mM KCl, pH 7.5) to a final concentration of 500 ng/µL. This was stored at −20°C until injection. On the day of injection, while fish mated and prior to collecting eggs, we prepared the RNP microinjection mix. For CRISPR/Cas9 knockouts we combined 2 µL 3 µM gRNA and 2 µL 500 ng/µL Cas9 protein and incubated at 37°C for 10 min before placing on ice and adding ~0.2 µL 0.05% phenol red for visualization while injecting.

The CRISPR/Cas9 knockin microinjection mix consisted of 2 µL 3 µM gRNA, 1.5 µL 100 ng/µL 5′ biotinylated donor DNA, 1.5 µL 400 ng/µL Cas9-mSA mRNA, and ~0.2 µL 0.05% phenol red. The gRNA was assembled in the same manner as described above for our knockouts. The 5′-biotinylated donor DNA was made by amplifying GFP followed by a stop codon from a plasmid using primers with biotinylated 40-bp overhangs that annealed to the host genome surrounding the cleavage site in frame. We confirmed the expected size of the PCR product by running an aliquot on a gel, and then used the DNA Clean and Concentrator Kit (D4209, Zymo Research, Irvine, CA) to purify the PCR reaction before diluting to a final concentration of 100 ng/µL in nuclease free water. To prepare Cas9-mSA mRNA, we linearized PCS2+Cas9-mSA (Addgene plasmid #103882 (Gu et al., 2018)); using NotI-HF (R3189S, New England Biolabs, Waltham, MA) and then transcribed it using mMessage mMachine SP6 (AM1340, Invitrogen, Waltham, MA). We then purified mRNA using the RNA clean and concentrator kit (R1017, Zymo Research), quantified it using a NanoDrop One (13-400-519, Thermo Fisher Scientific, Waltham MA), and made 400 ng/µL aliquots, which we stored at −70°C until use.

The Tol2 transgenesis microinjection mix consisted of 0.5 µL 250 ng/µL plasmid DNA, 0.5 µL 250 ng/µL transposase mRNA, 2.5 µL 0.4M KCl, 1 µL ddH_2O, ~0.2 µL 0.05% phenol red. To prepare Tol2 transposase mRNA, we linearized pCS2-zT2TP (Suster et al., 2011) (a gift from Koichi Kawakami via Jamie Gagnon) using NotI-HF and then transcribed it using mMessage mMachine SP6. We then purified mRNA using the RNA clean and concentrator kit and made 250 ng/µL aliquots quantified using the NanoDrop One, which we stored at −70°C until use.

After collecting the eggs from the mating tanks, we used a glass Pasteur pipette broken at the taper and flamed smooth, to transfer eggs from the 250 mL beaker to our microinjection mold. To prevent the eggs from moving around in the mold, we removed most of the water leaving just enough to prevent the eggs from drying. We then loaded 1 µL of the microinjection mix into the beveled needle using a P2.5 pipette by slowly adding the mix to the back of the needle and then allowing capillary action to bring the mix to the needle tip. After ensuring the micromanipulator (M-152, Narishige, Setagaya, Tokyo) was in the proper position (able to move freely in all directions), we loaded the needle into the mount of the micromanipulator. After visually confirming under the microscope (Zeiss Stemi 508) most embryos were not past the 1–4 cell stage, we microinjected ~1 nL microinjection mix into each embryo (either the single cell or one of the dividing cells) using a FemtoJet 4i microinjector (Eppendorf, Hamburg, Germany) (Supplementary Figure S1). After each round of injections, we used a plastic squeeze bottle filled with E3 medium to gently wash the injected embryos into a glass petri dish, and then placed the petri dish in a 28°C incubator until day 5 post fertilization. The incubator has a clear door from where it gets ambient light. At day 5, we transferred the fry into 1 L of water in a temperature (26–27°C) and light controlled (14 h light, 10 h dark) room and began feeding rotifers as described (Lichak et al., 2022). More details on husbandry can be found in (Lichak et al., 2022). Survival rates of injected animals are shown in Supplementary Figure S5.

4.3 Imaging

We photographed sexually mature fish inside a plastic tank using a Canon EOS RP with a Macro 100 mm lens (Macro Lens EF 100mm, Canon, Tokyo, Japan), as described (Kwon et al., 2022). All photos were taken inside a photo studio tent with white light emitting diodes. In each photo, a color card and ruler were visible, and a mirror was placed on the outside of the tank facing the fish to facilitate flaring, allowing for visualization of fins and consistency amongst photos. Before each imaging session, we calibrated the camera using the white balance (CT24-23–1424) on the 24ColorCard (CameraTrax).

4.4 Color analysis

We analyzed fish coloration in terms of hue, saturation and value (brightness) as in (Kwon et al., 2022). We used a custom Python script (Supplementary Code) to automatically threshold each image and determine the hue, saturation, and value (HSV) of each pixel of the body. We then calculated the joint distributions of HSV and the fraction of blue and red pixels. On a 0–180 hue scale, we defined blue as the range from 110 to 130 and red as the range from 0 to 30 (see hue scale bar in Figure 1E). For the *bco1l* color analysis, we differentiated orange and red pixels by defining red as the range from 0 to 15, and orange as the range from 15 to 30.

Primer name	Sequence	Purpose
Alkal_WT_MM_Fwd	AGAAAAGAGGTCACACCAGGAA	For genotyping *alkal2l* F3s using discriminative primers
Alkal_Mut_MM_Rev	AGAAAAGAGGTCACACCAATAT	For genotyping *alkal2l* F3s using discriminative primers
Alkal_Rev_(Wt/Mut)	ACAGACCAGATGTTAAGAGCTCA	For genotyping *alkal2l* F3s using discriminative primers
Mitfa_Donor_Fwd	/5Biosg/ CTTGGAGTCAAGTTACAATGAAGATGTCCTTG*GGGTGAGCAAGGGCGAGGAGCT*	For amplifying GFP (italicized) from plasmid with 5′ biotinylated overhangs annealing to *mitfa* locus surrounding cleavage site (nonitalicized)
Mitfa_Donor_Rev	/5Biosg/ CTTGGAGTCAAGTTACAATGAAGATGTCCTTG*CTTGTACAGCTCGTCCATGC*	For amplifying GFP (italicized) from plasmid with 5′ biotinylated overhangs annealing to *mitfa* locus surrounding cleavage site (nonitalicized)
Mitfa_T7_Fwd	AACCCCACAATAGACAAATGC	For genotyping mosaic and heterozygous *mitfa* crispants using T7EI
Mitfa_T7_Rev	CAAACGCAAAACACAAAAAGA	For genotyping mosaic and heterozygous *mitfa* crispants
Mitfa_crRNA	GAGCTGGGTCCATTAGTCCA	For targeting exon 3 of mitfa
Alkal2l_crRNA	AAAGAGGTCACACCAGGGAA	For targeting exon 1 of *alkal2l*
Bco1l_crRNA	GTTCATGTAGTTGATCTCAG	For targeting exon 4 of *bco1l*
Alkal2l_T7_Fwd	AGGAGTTGGAGCATGTGGAC	For genotyping mosaic and heterozygous *alkal2l* crispants using T7EI
Alkal2l_T7_Rev	GTGGTGATGTTCGTTGTTCG	For genotyping mosaic and heterozygous *alkal2l* crispants using T7EI
Bco1l_T7_Fwd	CAGATCCCTGCCAGAACATT	For genotyping mosaic and heterozygous *bco1l* crispants using T7EI
Bco1l_T7_Rev	TGAGCCTGTGGTTCACTGAC	For genotyping mosaic and heterozygous *bco1l* crispants using T7EI
Junct_Fwd_Mitfa	TGGGTGGTAAAACAGGAAGC	For junction PCR. Fwd primer is in *mitfa* locus
Junct_Rev_GFP	TCTCGTTGGGGTCTTTGCTC	For junction PCR. Rev primer is in GFP
V_C1_Fwd	TGGACCATGGTGAGCAAGGGC	For linearizing Tol2-*ubi*-EGFP and replacing *ubi* with the *actb* promoter
V_C1_Rev	ATAGCTTGGCGTAATCATGGTCATAGC	For linearizing Tol2-*ubi*-EGFP and replacing *ubi* with the *actb* promoter
F_C1_Fwd	TATGACCATGATTACGCCAAGCTAT*GTATAAACTCCTGTCTGGAAG*	For amplifying *actb* (italicized) from betta genomic DNA with overhangs annealing to tol2 backbone (nonitalicized)
F_C1_Rev	TCCTCGCCCTTGCTCACCATGGTCCA*GGTTTACTGTAAGTAAAAGAAAG*	For amplifying *actb* (italicized) from betta genomic DNA with overhangs annealing to tol2 backbone (nonitalicized)
V_C2_Fwd	CGTAAAACGACGGCCAGTGA	For linearizing Tol2-*actb*-GFP
V_C2_Rev	GGCTATTTTGGACTGTGCTGC	For linearizing Tol2-*actb*-GFP
F_C2_Fwd	TATTATACATAGTTGATAATTCACTGGCCGTCGTTTTACGGATAGGCCCCTTAC GTACGC	For amplifying *clmlc2*-mCherry-sv40 from pBH mcs
F_C2_Rev	ATTCATCAGCAGCTGCGAGCAGCACAGTCCAAAATAGCCGGGACAGATCTCGAG CTCAAG	For amplifying *clmlc2*-mCherry-sv40 from pBH mcs

4.5 Genotyping

We used various methods for genotyping crispants. P0, F1, and F2 crispants were fin clipped at 5 weeks (when fish are ~1.2 cm long) and genotyped by a T7EI assay (Mashal et al., 1995; Sentmanat et al., 2018) (M0302, New England Biolabs). In this assay, heteroduplex DNA (arising from heterozygosity and/or from cellular mosaicism of the CRISPR-induced mutation) amplified by PCR is susceptible to cutting by T7 endonuclease I. The results of the assay are visualized through gel electrophoresis (Supplementary Figure S3A).

We extracted DNA using the Quick-DNA Microprep Kit (D3021, Zymo Research). We Sanger sequenced *alkal2l* and *bco1l* F1s and used CRISP-ID (Dehairs et al., 2016) to aid with identifying indels. To genotype *alkal2l* +/− × +/− crosses of the allele consisting of a 2-bp insertion plus a 1-bp mismatch, we used discriminative primers (Supplementary Figure S3B) annealing specifically to mutant and wild type alleles. To increase specificity of these primers to their respective alleles, we introduced an extra mismatch (Chen and Schedl, 2021). To genotype the *bco1l* +/− × +/− crosses of the 1-bp insertion allele, we Sanger sequenced the progeny.

4.6 Cloning

To clone the Tol2-based transgenesis plasmids, we first amplified the 4,165-bp region upstream of the transcription start site of the betta *actb* gene using Q5 polymerase (M0491S, New England Biolabs) from an ornamental betta. We then replaced the ubiquitin promoter in the pDestTol2pA2_ubi:EGFP plasmid [Addgene plasmid #27323 (Mosimann et al., 2011)]; with the amplified *actb* promoter region using bacterial *in vivo* assembly (García-Nafría et al., 2016). We then purified plasmids for injection using the ZymoPURE plasmid miniprep kit (D4210, Zymo Research). We fully sequenced all plasmids prior to injection to confirm intended sequences. To create the bicistronic plasmid, we first linearized *actb*:EGFP and then amplified a 1.3-kb region from pBH-mcs (a gift from Michael Nonet) containing the 268-bp of the zebrafish *cmlc2* (*myl7*) promoter region, mCherry, and the polyadenylation and large T antigen signals. We then carried out all subsequent steps as described above.

4.7 Primer sequences

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Ethics statement

The animal study was reviewed and approved by the Columbia University Institutional Animal Care and Use Committee.

Author contributions

AP, ML, and YK contributed to conceptualization of the research project, investigation, data curation, methodology development and data analysis. AB and P-YS contributed to conceptualization of the research project, methodology development and supervision. AP (original draft) and AB wrote the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the following grants to AB: Searle Scholarship, Sloan Foundation Fellowship, and National Institutes of Health R34NS116734 and R35GM143051.

Acknowledgments

Michael Nonet, Jamie Gagnon, and Koichi Kawakami shared reagents. David Ng provided technical assistance. Joshua Barber, Christoph Gebhardt and Claire Everett provided advice for microinjections and fish care.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgeed.2023.1167093/full#supplementary-material

References

Chen, J., and Schedl, T. (2021). A simple one-step PCR assay for SNP detection. *MicroPublication Biol.* 2021. doi:10.17912/micropub.biology.000399

Danner, E., Bashir, S., Yumlu, S., Wurst, W., Wefers, B., and Kühn, R. (2017). Control of gene editing by manipulation of DNA repair mechanisms. *Mamm. Genome.* 28, 262–274. doi:10.1007/s00335-017-9688-5

Dehairs, J., Talebi, A., Cherifi, Y., and Swinnen, J. V. (2016). CRISP-ID: Decoding CRISPR mediated indels by sanger sequencing. *Sci. Rep.* 6, 28973. doi:10.1038/srep28973

Fan, G., Chan, J., Ma, K., Yang, B., Zhang, H., Yang, X., et al. (2018). Chromosome-level reference genome of the Siamese fighting fish *Betta splendens*, a model species for the study of aggression. *GigaScience* 7, giy087. doi:10.1093/gigascience/giy087

García-Nafría, J., Watson, J. F., and Greger, I. H. (2016). IVA cloning: A single-tube universal cloning system exploiting bacterial *in vivo* assembly. *Sci. Rep.* 6, 27459. doi:10.1038/srep27459

Gu, B., Posfai, E., and Rossant, J. (2018). Efficient generation of targeted large insertions by microinjection into two-cell-stage mouse embryos. *Nat. Biotechnol.* 36, 632–637. doi:10.1038/nbt.4166

Gutierrez-Triana, J. A., Tavhelidse, T., Thumberger, T., Thomas, I., Wittbrodt, B., Kellner, T., et al. (2018). Efficient single-copy HDR by 5' modified long dsDNA donors. *eLife* 7, e39468. doi:10.7554/eLife.39468

Hisano, Y., Sakuma, T., Nakade, S., Ohga, R., Ota, S., Okamoto, H., et al. (2015). Precise in-frame integration of exogenous DNA mediated by CRISPR/Cas9 system in zebrafish. *Sci. Rep.* 5, 8841. doi:10.1038/srep08841

Huang, C. J., Tu, C. T., Hsiao, C. D., Hsieh, F. J., and Tsai, H. J. (2003). Germ-line transmission of a myocardium-specific GFP transgene reveals critical regulatory elements in the cardiac myosin light chain 2 promoter of zebrafish. *Dev. Dyn.* 228, 30–40. doi:10.1002/dvdy.10356

Hwang, W. Y., Fu, Y., Reyon, D., Maeder, M. L., Tsai, S. Q., Sander, J. D., et al. (2013). Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat. Biotechnol.* 31, 227–229. doi:10.1038/nbt.2501

Kawakami, K., Koga, A., Hori, H., and Shima, A. (1998). Excision of the tol2 transposable element of the medaka fish, *Oryzias latipes*, in zebrafish, *Danio rerio*. *Gene* 225, 17–22. doi:10.1016/s0378-1119(98)00537-x

Kawakami, K. (2007). Tol2: A versatile gene transfer vector in vertebrates. *Genome Biol.* 8 (1), S7. doi:10.1186/gb-2007-8-s1-s7

Klatt Shaw, D., and Mokalled, M. H. (2021). Efficient CRISPR/Cas9 mutagenesis for neurobehavioral screening in adult zebrafish. *G3* 11, jkab089. doi:10.1093/g3journal/jkab089

Kroll, F., Powell, G. T., Ghosh, M., Gestri, G., Antinucci, P., Hearn, T. J., et al. (2021). A simple and effective F0 knockout method for rapid screening of behaviour and other complex phenotypes. *eLife* 10, e59683. doi:10.7554/eLife.59683

Kwon, Y. M., Vranken, N., Hoge, C., Lichak, M. R., Norovich, A. L., Francis, K. X., et al. (2022). Genomic consequences of domestication of the Siamese fighting fish. *Sci. Adv.* 8, eabm4950. doi:10.1126/sciadv.abm4950

Lichak, M. R., Barber, J. R., Kwon, Y. M., Francis, K. X., and Bendesky, A. (2022). Care and use of Siamese fighting fish (*Betta splendens*) for research. *Comp. Med.* 72, 169–180. doi:10.30802/AALAS-CM-22-000051

Lino, C. A., Harper, J. C., Carney, J. P., and Timlin, J. A. (2018). Delivering CRISPR: A review of the challenges and approaches. *Drug Deliv.* 25, 1234–1257. doi:10.1080/10717544.2018.1474964

Lister, J. A., Robertson, C. P., Lepage, T., Johnson, S. L., and Raible, D. W. (1999). Nacre encodes a zebrafish microphthalmia-related protein that regulates neural-crest-derived pigment cell fate. *Development* 126, 3757–3767. doi:10.1242/dev.126.17.3757

Liu, K., Petree, C., Requena, T., Varshney, P., and Varshney, G. K. (2019). Expanding the CRISPR toolbox in zebrafish for studying development and disease. *Front. Cell Dev. Biol.* 7, 13. doi:10.3389/fcell.2019.00013

Mashal, R. D., Koontz, J., and Sklar, J. (1995). Detection of mutations by cleavage of DNA heteroduplexes with bacteriophage resolvases. *Nat. Genet.* 9, 177–183. doi:10.1038/ng0295-177

Mi, J., and Andersson, O. (2023). Efficient knock-in method enabling lineage tracing in zebrafish. *Life Sci. Alliance* 6, e202301944. doi:10.26508/lsa.202301944

Mo, E. S., Cheng, Q., Reshetnyak, A. V., Schlessinger, J., and Nicoli, S. (2017). Alk and Ltk ligands are essential for iridophore development in zebrafish mediated by the receptor tyrosine kinase Ltk. *Proc. Natl. Acad. Sci. U. S. A.* 114, 12027–12032. doi:10.1073/pnas.1710254114

Mosimann, C., Kaufman, C. K., Li, P., Pugach, E. K., Tamplin, O. J., and Zon, L. I. (2011). Ubiquitous transgene expression and Cre-based recombination driven by the *ubiquitin* promoter in zebrafish. *Dev. Camb. Engl.* 138, 169–177. doi:10.1242/dev.059345

Prost, S., Petersen, M., Grethlein, M., Hahn, S. J., Kuschik-Maczollek, N., Olesiuk, M. E., et al. (2020). Improving the chromosome-level genome assembly of the Siamese fighting fish (*Betta splendens*) in a University master's course. *G3* 10, 2179–2183. doi:10.1534/g3.120.401205

Ranawakage, D. C., Okada, K., Sugio, K., Kawaguchi, Y., Kuninobu-Bonkohara, Y., Takada, T., et al. (2020). Efficient CRISPR-Cas9-mediated knock-in of composite tags in zebrafish using long ssDNA as a donor. *Front. Cell Dev. Biol.* 8, 598634. doi:10.3389/fcell.2020.598634

Seleit, A., Aulehla, A., and Paix, A. (2021). Endogenous protein tagging in medaka using a simplified CRISPR/Cas9 knock-in approach. *eLife* 10, e75050. doi:10.7554/eLife.75050

Sentmanat, M. F., Peters, S. T., Florian, C. P., Connelly, J. P., and Pruett-Miller, S. M. (2018). A survey of validation strategies for CRISPR-Cas9 editing. *Sci. Rep.* 8, 888. doi:10.1038/s41598-018-19441-8

Silver, D. L., Hou, L., and Pavan, W. J. (2006). The genetic regulation of pigment cell development. *Adv. Exp. Med. Biol.* 589, 155–169. doi:10.1007/978-0-387-46954-6_9

Suster, M. L., Abe, G., Schouw, A., and Kawakami, K. (2011). Transposon-mediated BAC transgenesis in zebrafish. *Nat. Protoc.* 6, 1998–2021. doi:10.1038/nprot.2011.416

Urasaki, A., Morvan, G., and Kawakami, K. (2006). Functional dissection of the Tol2 transposable element identified the minimal cis-sequence and a highly repetitive sequence in the subterminal region essential for transposition. *Genetics* 174, 639–649. doi:10.1534/genetics.106.060244

Valentin, F. N., do Nascimento, N. F., da Silva, R. C., Fernandes, J. B. K., Giannecchini, L. G., and Nakaghi, L. S. O. (2015). Early development of *Betta splendens* under stereomicroscopy and scanning electron microscopy. *Zygote Camb. Engl.* 23, 247–256. doi:10.1017/S0967199413000488

Wang, L., Sun, F., Wan, Z. Y., Yang, Z., Tay, Y. X., Lee, M., et al. (2022). Transposon-induced epigenetic silencing in the X chromosome as a novel form of *dmrt1* expression regulation during sex determination in the fighting fish. *BMC Biol.* 20, 5. doi:10.1186/s12915-021-01205-y

Wang, L., Sun, F., Wan, Z. Y., Ye, B., Wen, Y., Liu, H., et al. (2021). Genomic basis of striking fin shapes and colors in the fighting fish. *Mol. Biol. Evol.* 38, 3383–3396. doi:10.1093/molbev/msab110

Watakabe, I., Hashimoto, H., Kimura, Y., Yokoi, S., Naruse, K., and Higashijima, S.-I. (2018). Highly efficient generation of knock-in transgenic medaka by CRISPR/Cas9-mediated genome engineering. *Zool. Lett.* 4, 3. doi:10.1186/s40851-017-0086-3

Wierson, W. A., Welker, J. M., Almeida, M. P., Mann, C. M., Webster, D. A., Torrie, M. E., et al. (2020). Efficient targeted integration directed by short homology in zebrafish and mammalian cells. *eLife* 9, e53968. doi:10.7554/eLife.53968

Yue, G. H., Wang, L., Sun, F., Yang, Z., Shen, Y., Meng, Z., et al. (2022). The ornamental fighting fish is the next model organism for genetic studies. *Rev. Aquac.* 14, 1966–1977. doi:10.1111/raq.12681

Zhang, W., Wang, H., Brandt, D. Y. C., Hu, B., Sheng, J., Wang, M., et al. (2022). The genetic architecture of phenotypic diversity in the Betta fish (*Betta splendens*). *Sci. Adv.* 8, eabm4955. doi:10.1126/sciadv.abm4955

Supplementary Material

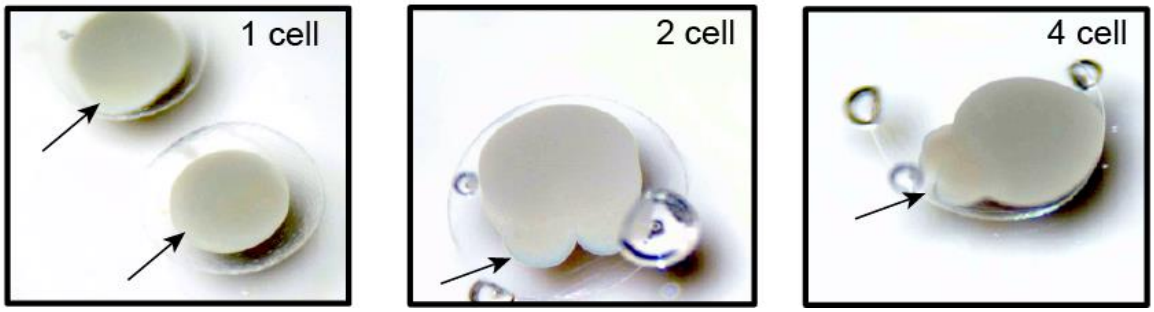
Genetic manipulation of betta fish

Alec Palmiotti[1,2]†, Madison R Lichak[1,2,3] †, Pei-Yin Shih[1,2], Young Mi Kwon[1,2], Andres Bendesky[1,2]*

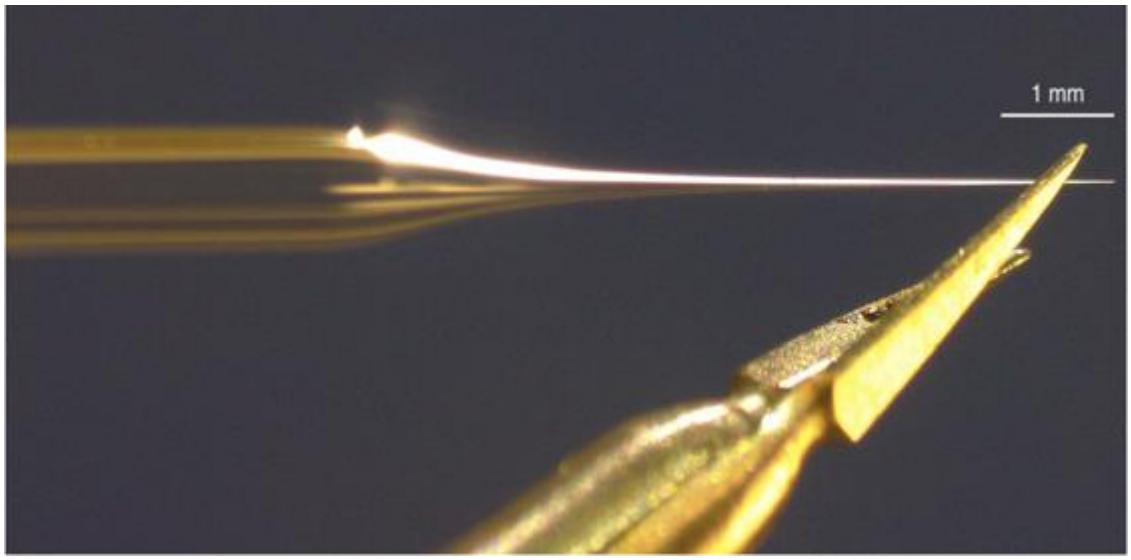
[1]Zuckerman Mind Brain Behavior Institute, Columbia University, NY, USA.
[2]Department of Ecology, Evolution and Environmental Biology, Columbia University, NY, USA.
[3]Present address: Department of Ecology and Evolutionary Biology, Princeton University, NJ, USA.

† Equal contribution

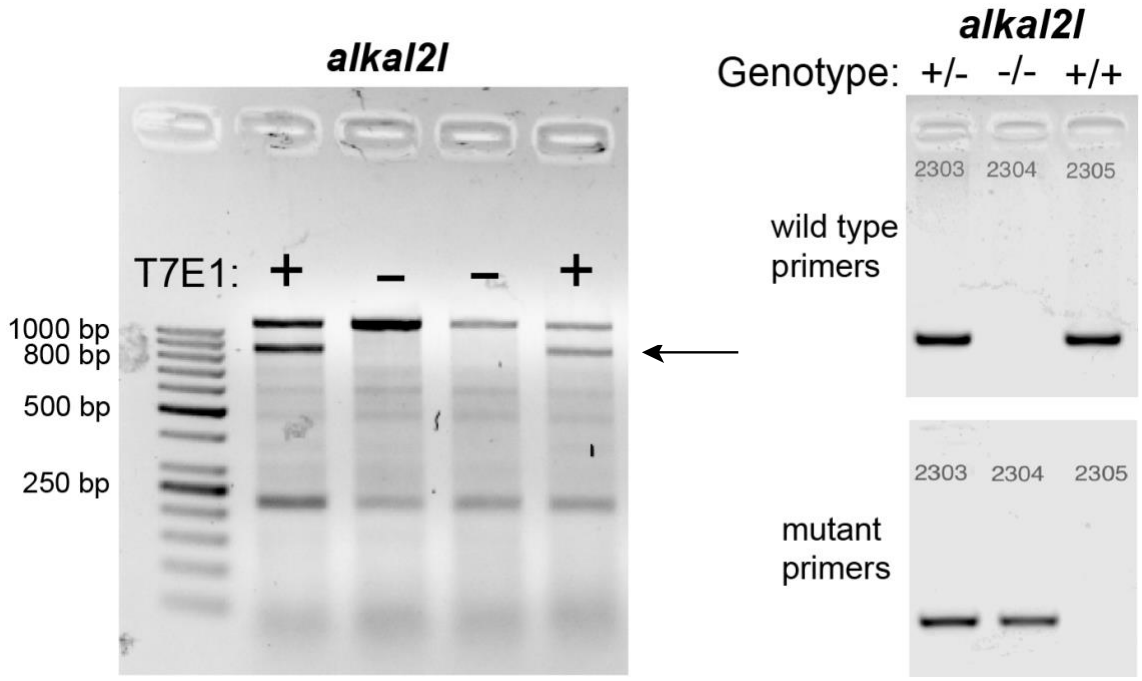
*** Correspondence:** *a.bendesky@columbia.edu*



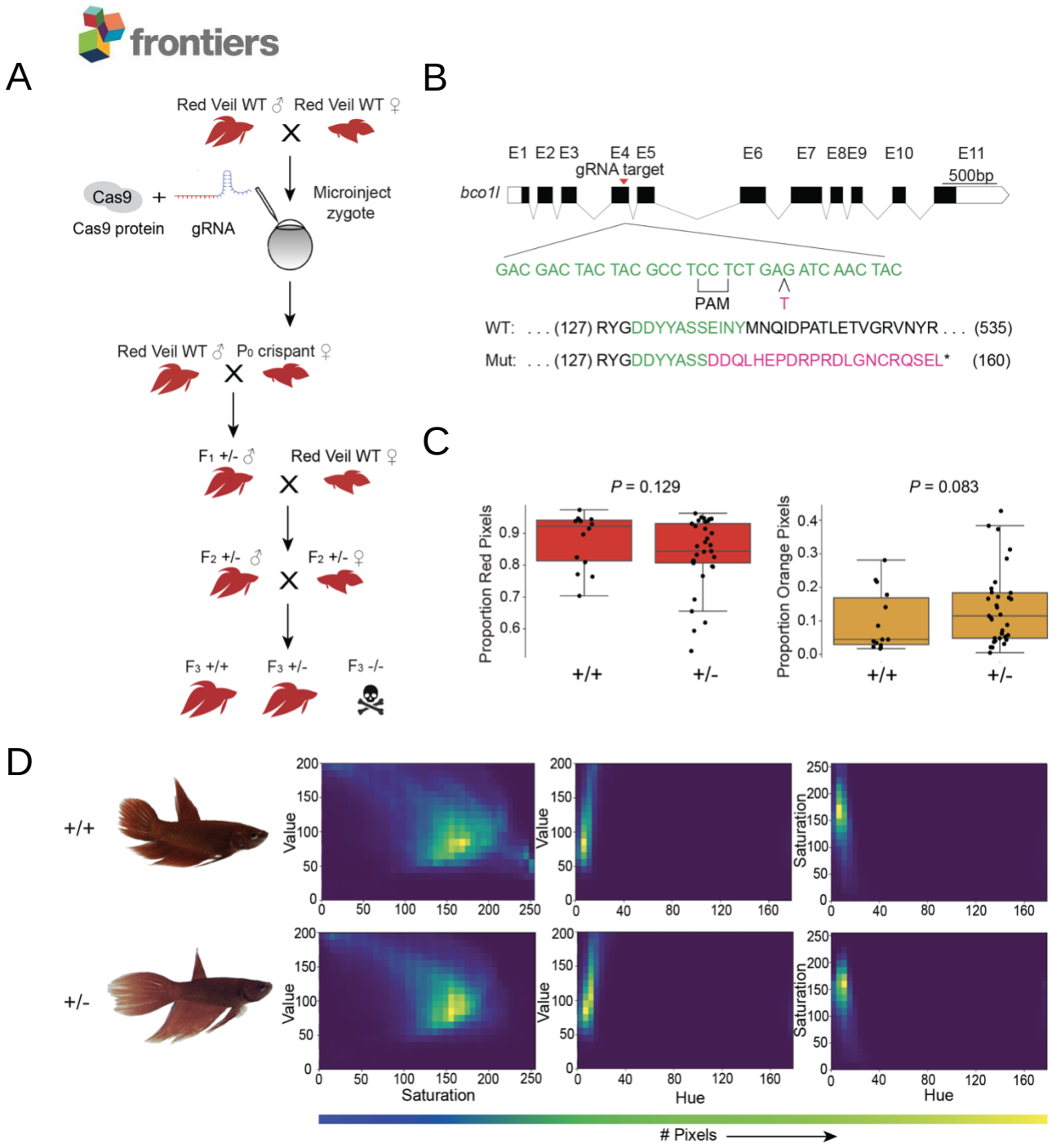
Supplementary Figure 1. Images of single-celled embryo (left), two-celled embryo (middle) and four-celled embryo (right).



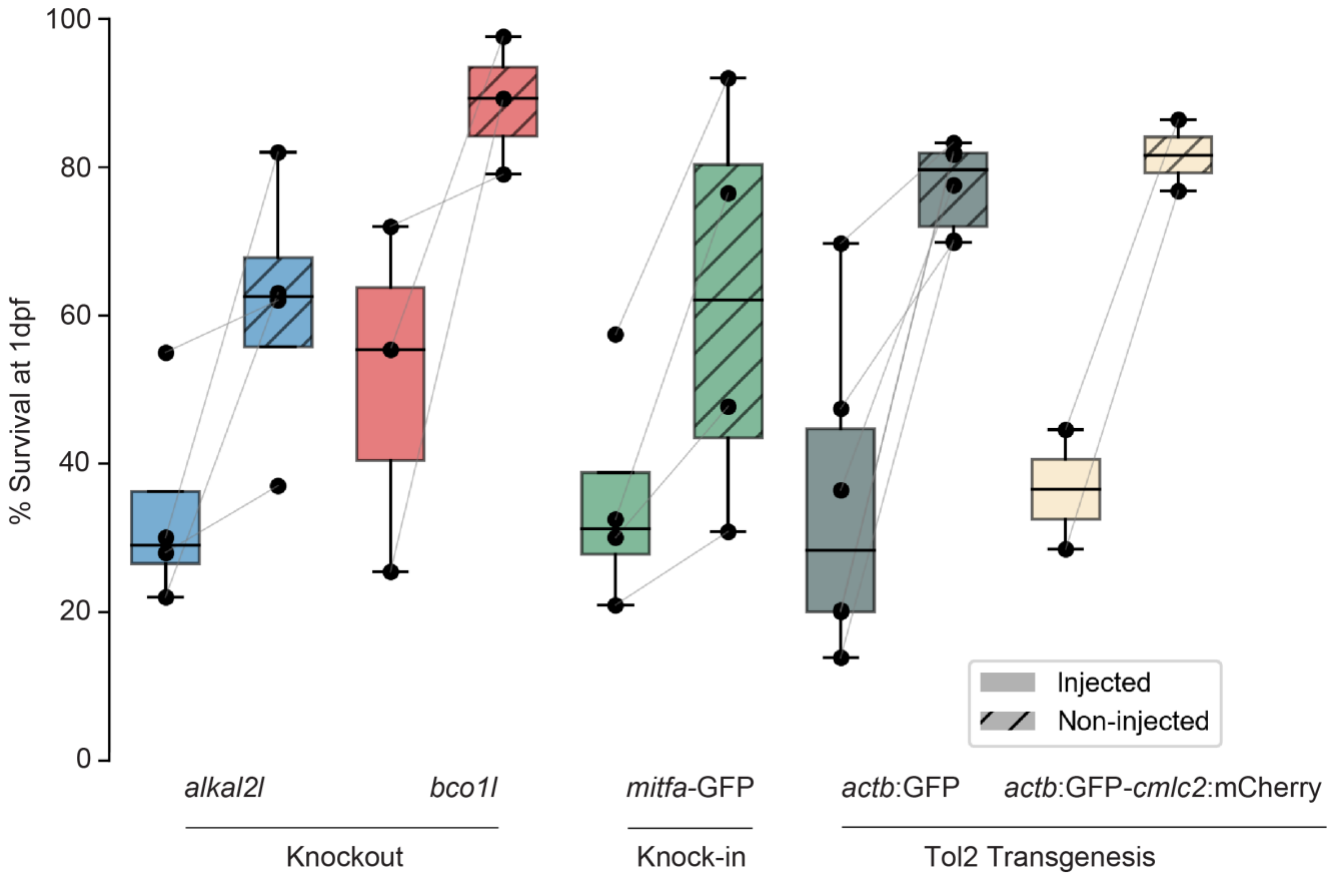
Supplementary Figure 2. Photograph of injection needle being cut ~0.5 mm from the tip at ~45º. The diameter of the tip is ~8 µm.



Supplementary Figure 3. (A) Representative T7EI assay for genotyping mosaic and heterozygous *alkal2l* crispants. Arrow points to extra band indicative of positive T7EI assay. + and – indicate presence or absence of heteroduplex DNA. **(B)** Representative discriminative primer assay for genotyping F_3 *alkal2l* +/+, +/– and –/– animals. 2303, 2304 and 2305 are fish IDs.



Supplementary Figure 4. (A) Experimental scheme for generating *bco1l* knockout in betta **(B)** Top: schematic of *bco1l* gene and target location of the gRNA; middle: wildtype DNA sequence in green and CRISPR/Cas9-induced mutations in magenta; bottom: wildtype protein sequence (green shows the amino acids encoded by the wildtype sequence shown in the middle) and mutant sequence due to a frame shift (magenta shows the new mutant amino acids and * denotes a stop codon). The numbers in parenthesis on the right side indicate the total length of the proteins. **(C)** Boxplots showing proportion of red (left) and orange (right) pixels in *bco1l* crispants, according to genotype. *P* values by Mann-Whitney test. Boxes denote the interquartile range and whiskers the 5th and 95th percentiles, with a line at the median **(D)** Joint HSV distributions of *bco1l* (+/+) and (+/-) animals.



Supplementary Figure 5. Viability at 1 dpf for various genetic manipulation techniques. Each point represents the injection of a single brood and is connected to the non-injected control of the same brood via gray lines.

Supplementary Code. *alkal2l* color analysis python3 code

```
import os
from glob import glob
import cv2 as cv
import numpy as np
import pandas as pd
from matplotlib import pyplot as plt
import seaborn as sns

img_fh = '2427_Homo_C1.tiff'
img = cv.imread(img_fh,cv.IMREAD_COLOR)
img_hsv = cv.cvtColor(img, cv.COLOR_BGR2HSV)
plt.imshow(cv.cvtColor(img_hsv, cv.COLOR_HSV2RGB))
plt.show()
hue_hist = cv.calcHist([img_hsv],[0],None,[256],[0,256])
sat_hist = cv.calcHist([img_hsv],[1],None,[256],[0,256])
val_hist = cv.calcHist([img_hsv],[2],None,[256],[0,256])
plt.plot(hue_hist)
plt.plot(sat_hist)
plt.plot(val_hist)
plt.legend(['hue','sat','value'])
plt.show()
mask = cv.inRange(img_hsv, np.array([0,0,0]), np.array([255,255,200]))
plt.imshow(mask)
plt.show()
res = cv.bitwise_and(img_hsv, img_hsv, mask=mask)
plt.imshow(cv.cvtColor(res, cv.COLOR_HSV2RGB))
plt.show()
hue_hist = cv.calcHist([img_hsv],[0],mask,[256],[0,256])
sat_hist = cv.calcHist([img_hsv],[1],mask,[256],[0,256])
val_hist = cv.calcHist([img_hsv],[2],mask,[256],[0,256])
plt.plot(hue_hist)
plt.plot(sat_hist)
plt.plot(val_hist)
plt.legend(['hue','sat','value'])
plt.show()
h = np.arange(180, dtype=np.uint8).reshape(1, 180, 1)
s = (np.ones(180, dtype=np.uint8)*255).reshape(1, 180, 1)
v = s.copy()
hs = np.append(h, s, axis=2)
hsv = np.append(hs, v, axis=2)
hsv = np.repeat(hsv, 20, axis=0)
rgb = cv.cvtColor(hsv, cv.COLOR_HSV2RGB)

plt.figure(figsize=(6,3.2))
plt.imshow(rgb)
plt.yticks([])
# plt.xticks([0,20,30,40, 60, 120, 180], fontsize=15)
plt.tight_layout(pad=0.3)
plt.show()
np.sum(hue_hist[90:140])/np.sum(hue_hist)
hues_in_mask = img_hsv[:,:,0][mask == 255]
sats_in_mask = img_hsv[:,:,1][mask == 255]
vals_in_mask = img_hsv[:,:,2][mask == 255]
```

```
plt.hist2d(hues_in_mask.ravel(), sats_in_mask.ravel(), bins=(40, 20))
plt.xlim(0,179)
plt.ylim(0,255)
plt.xlabel('Hue')
plt.ylabel('Saturation')
plt.show()
plt.hist2d(hues_in_mask.ravel(), vals_in_mask.ravel(), bins=(40, 20))
plt.xlim(0,179)
plt.ylim(0,200)
plt.xlabel('Hue')
plt.ylabel('Value')
plt.show()
plt.hist2d(sats_in_mask.ravel(), vals_in_mask.ravel(), bins=(40, 20))
plt.xlim(0,255)
plt.ylim(0,200)
plt.xlabel('Saturation')
plt.ylabel('Value')
plt.show()
print('Average')
print('hue',np.mean(hues_in_mask))
print('sat',np.mean(sats_in_mask))
print('val',np.mean(vals_in_mask))

# %%

tiffs = glob(os.path.join('Alkal_Data/','*.tiff'))
tiffs

def analyze_tiff(tiff):
    img_fh = tiff
    results_file_prefix = os.path.basename(img_fh).removesuffix('.tiff')
    img = cv.imread(img_fh,cv.IMREAD_COLOR)
    img_hsv = cv.cvtColor(img, cv.COLOR_BGR2HSV)
    fig, ax = plt.subplots(3,1)
    # Plot original image
    ax[0].imshow(cv.cvtColor(img_hsv, cv.COLOR_HSV2RGB))
    mask = cv.inRange(img_hsv, np.array([0,0,0]), np.array([255,255,200]))
    # Plot mask
    ax[1].imshow(mask)
    res = cv.bitwise_and(img_hsv, img_hsv, mask=mask)
    # Plot maksed image
    ax[2].imshow(cv.cvtColor(res, cv.COLOR_HSV2RGB))
    fig.suptitle(results_file_prefix)
    plt.savefig(results_file_prefix+'masked_image.jpg')
    plt.close()

    # Plot hue histograms
    hue_hist = cv.calcHist([img_hsv],[0],mask,[256],[0,256])
    sat_hist = cv.calcHist([img_hsv],[1],mask,[256],[0,256])
    val_hist = cv.calcHist([img_hsv],[2],mask,[256],[0,256])
    fig, ((ax1, ax2), (ax3, ax4)) = plt.subplots(2, 2)
    ax1.plot(hue_hist)
    ax1.plot(sat_hist)
    ax1.plot(val_hist)
    ax1.legend(['hue','sat','value'])
```

```
    # Plot joint distributions
    hues_in_mask = img_hsv[:,:,0][mask == 255]
    sats_in_mask = img_hsv[:,:,1][mask == 255]
    vals_in_mask = img_hsv[:,:,2][mask == 255]

    hue_sat_hist2d, _,_,_ = ax2.hist2d(hues_in_mask.ravel(),
sats_in_mask.ravel(), bins=(40, 20), range=[[0,180],[0,255]])
    ax2.set_xlim(0,179)
    ax2.set_ylim(0,255)
    ax2.set_xlabel('Hue')
    ax2.set_ylabel('Saturation')

    hue_val_hist2d, _,_,_ = ax3.hist2d(hues_in_mask.ravel(),
vals_in_mask.ravel(), bins=(40, 20), range=[[0,180],[0,200]])
    ax3.set_xlim(0,179)
    ax3.set_ylim(0,200)
    ax3.set_xlabel('Hue')
    ax3.set_ylabel('Value')

    sat_val_hist2d, _,_,_ = ax4.hist2d(sats_in_mask.ravel(),
vals_in_mask.ravel(), bins=(40, 20), range=[[0,255],[0,200]])
    ax4.set_xlim(0,255)
    ax4.set_ylim(0,200)
    ax4.set_xlabel('Saturation')
    ax4.set_ylabel('Value')

    fig.suptitle(results_file_prefix)
    fig.tight_layout()
    plt.savefig(results_file_prefix+'300dpidistributions.jpg.', dpi=300)
    plt.close()

    # Fraction blue pixels
    blue_pixels = np.sum(hue_hist[110:130])/np.sum(hue_hist)

    # Fraction redish pixels
    redish_pixels = np.sum(hue_hist[0:30])/np.sum(hue_hist)

    # Mean HSV
    mean_hue = np.mean(hues_in_mask)
    mean_sat = np.mean(sats_in_mask)
    mean_val = np.mean(vals_in_mask)

    return results_file_prefix, blue_pixels, redish_pixels, mean_hue,
mean_sat, mean_val, \
        hue_sat_hist2d, hue_val_hist2d, sat_val_hist2d

# Dictionary to hold results:
results_dict = {'Fish_ID':[], 'prop_blue_pixels':[],
'prop_redish_pixels':[],\
                'mean_hue':[], 'mean_sat':[], 'mean_val':[]}
pixel_dist_dict = {'hue_sat_hist2d':{'Wt':[], 'Het':[], 'Homo':[]}, \
                   'hue_val_hist2d':{'Wt':[], 'Het':[], 'Homo':[]}, \
                       'sat_val_hist2d':{'Wt':[], 'Het':[], 'Homo':[]}}

for tiff in tiffs:
```

```
    results_file_prefix, blue_pixels, redish_pixels, mean_hue, mean_sat,
mean_val, \
        hue_sat_hist2d, hue_val_hist2d, sat_val_hist2d = analyze_tiff(tiff)
    results_dict['Fish_ID'].append(results_file_prefix)
    results_dict['prop_blue_pixels'].append(blue_pixels)
    results_dict['prop_redish_pixels'].append(redish_pixels)
    results_dict['mean_hue'].append(mean_hue)
    results_dict['mean_sat'].append(mean_sat)
    results_dict['mean_val'].append(mean_val)
    results_df = pd.DataFrame.from_dict(results_dict)
    results_df
    # results_df.to_csv('results.csv')

    geno = results_file_prefix.split('_')[1]
    pixel_dist_dict['hue_sat_hist2d'][geno].append(hue_sat_hist2d)
    pixel_dist_dict['hue_val_hist2d'][geno].append(hue_val_hist2d)
    pixel_dist_dict['sat_val_hist2d'][geno].append(sat_val_hist2d)


genotype = [i.split('_')[1] for i in results_df['Fish_ID']]
results_df['genotype'] = genotype
results_df.head()

# Import necessary modules
from scipy import stats

# Plot results
order = ['Wt','Het', 'Homo']

# Perform one-way ANOVA test and post-hoc Tukey test on prop_blue_pixels data
F, p = stats.f_oneway(results_df[results_df['genotype'] ==
'Wt']['prop_blue_pixels'], results_df[results_df['genotype'] ==
'Het']['prop_blue_pixels'], results_df[results_df['genotype'] ==
'Homo']['prop_blue_pixels'])
if p < 0.05:
    from statsmodels.stats.multicomp import pairwise_tukeyhsd
    tukey = pairwise_tukeyhsd(endog=results_df['prop_blue_pixels'],
groups=results_df['genotype'], alpha=0.05)
    print(tukey.summary())

# Perform Mann Whitney U test on prop_blue_pixels and Prop_redish_pixels data
from scipy.stats import mannwhitneyu
homo_prop_blue = results_df[results_df['genotype']=='Wt']['prop_blue_pixels']
wt_prop_blue = results_df[results_df['genotype']=='Het']['prop_blue_pixels']
U1, p = mannwhitneyu(homo_prop_blue, wt_prop_blue, alternative='less')
print(p)

wt_prop_red = results_df[results_df['genotype']=='Wt']['prop_redish_pixels']
het_prop_red =
results_df[results_df['genotype']=='Het']['prop_redish_pixels']
U1, p = mannwhitneyu(wt_prop_red, het_prop_red, alternative='less')
print(p)

homo_prop_red =
results_df[results_df['genotype']=='Homo']['prop_redish_pixels']
```

```
het_prop_red =
results_df[results_df['genotype']=='Het']['prop_redish_pixels']
U1, p = mannwhitneyu(homo_prop_red, het_prop_red, alternative='greater')
print(p)

homo_prop_red =
results_df[results_df['genotype']=='Homo']['prop_redish_pixels']
wt_prop_red = results_df[results_df['genotype']=='Wt']['prop_redish_pixels']
U1, p = mannwhitneyu(homo_prop_red, wt_prop_red, alternative='greater')
print(p)


# Plot boxplots
sns.boxplot(x='genotype', y='prop_blue_pixels', data=results_df, order =
order, color="skyblue", fliersize=0)
sns.stripplot(x='genotype', y='prop_blue_pixels', data=results_df, order =
order, jitter=True, size=6, color='black')
plt.xlabel('Genotype')
plt.ylabel('Proportion Blue Pixels')
plt.title("Proportion Blue Pixels by Genotype")
plt.savefig(results_file_prefix+'BluishPixelBoxPlots.jpg.', dpi=300)
plt.show()

# Plot boxplots
order = ['Wt','Het', 'Homo']
sns.boxplot(x='genotype', y='prop_redish_pixels', data=results_df,
order=order, color="crimson", fliersize=0)
sns.stripplot(x='genotype', y='prop_redish_pixels', data=results_df, order =
order, jitter=True, size=6, color='black')
plt.xlabel('Genotype')
plt.ylabel('Proportion Red Pixels')
plt.title("Proportion Red Pixels by Genotype")
plt.savefig(results_file_prefix+'RedishPixelBoxPlots.jpg.', dpi=300)
plt.show()

distributions = ['hue_sat_hist2d']
genotypes = ['Wt', 'Het', 'Homo']

for d in distributions:
    # Set figure size to 10 inches wide and 12 inches tall
    fig, ax = plt.subplots(3,1, figsize=(12,10))
    for j,g in enumerate(genotypes):
        aggregate = np.zeros(pixel_dist_dict[d][g][0].shape)
        for i in pixel_dist_dict[d][g]:
            aggregate += i
        # Set x and y tick positions
        ax[j].set_xticks(np.arange(0, 181, 30))
        ax[j].set_yticks(np.arange(0, 256, 50))
        # Set x and y tick labels
        ax[j].set_xticklabels(['0', '30', '60', '90', '120', '150', '180'])
        ax[j].set_yticklabels(['0', '50', '100', '150', '200', '250'])
        ax[j].set_title("Hue vs Saturation", fontsize=20)
        ax[j].set_xlabel("Hue", fontsize=16, color="black")
        ax[j].set_ylabel("Saturation", fontsize=16, color="black")
        # Set the x and y axis limits
        ax[j].imshow(np.rot90(aggregate), extent=[0,180,0,255])
```

```
    fig.tight_layout()
    plt.savefig(d+'_joint-distributions.png', dpi=300)
    plt.show()

distributions = ['hue_val_hist2d']
genotypes = ['Wt', 'Het', 'Homo']

for d in distributions:
    # Set figure size to 10 inches wide and 12 inches tall
    fig, ax = plt.subplots(3,1, figsize=(12,10))
    for j,g in enumerate(genotypes):
        aggregate = np.zeros(pixel_dist_dict[d][g][0].shape)
        for i in pixel_dist_dict[d][g]:
            aggregate += i
        # Set x and y tick positions
        ax[j].set_xticks(np.arange(0, 181, 30))
        ax[j].set_yticks(np.arange(0, 201, 50))
        # Set x and y tick labels
        ax[j].set_xticklabels(['0', '30', '60', '90', '120', '150', '180'])
        ax[j].set_yticklabels(['0', '50', '100', '150', '200'])
        ax[j].set_title("Hue vs Value", fontsize=20)
        ax[j].set_xlabel("Hue", fontsize=16, color="black")
        ax[j].set_ylabel("Value", fontsize=16, color="black")
        # Set the x and y axis limits
        ax[j].imshow(np.rot90(aggregate), extent=[0,180,0,200])
    fig.tight_layout()
    plt.savefig(d+'_joint-distributions.png', dpi=300)
    plt.show()

distributions = ['sat_val_hist2d']
genotypes = ['Wt', 'Het', 'Homo']

for d in distributions:
    # Set figure size to 10 inches wide and 12 inches tall
    fig, ax = plt.subplots(3,1, figsize=(12,10))
    for j,g in enumerate(genotypes):
        aggregate = np.zeros(pixel_dist_dict[d][g][0].shape)
        for i in pixel_dist_dict[d][g]:
            aggregate += i
        # Set x and y tick positions
        ax[j].set_xticks(np.arange(0, 256, 50))
        ax[j].set_yticks(np.arange(0, 201, 50))
        # Set x and y tick labels
        ax[j].set_xticklabels(['0', '50', '100', '150', '200', '250'])
        ax[j].set_yticklabels(['0', '50', '100', '150', '200'])
        ax[j].set_title("Saturation vs Value", fontsize=20)
        ax[j].set_xlabel("Saturation", fontsize=16, color="black")
        ax[j].set_ylabel("Value", fontsize=16, color="black")
        # Set the x and y axis limits
        ax[j].imshow(np.rot90(aggregate), extent=[0,255,0,200])
    fig.tight_layout()
    plt.savefig(d+'_joint-distributions.png', dpi=300)
    plt.show()
```